

DEVELOPMENT OF THE HIGH RESOLUTION MELT (HRM)  
METHOD TO DETECT THE HLA-B\*58:01 ALLELE IN ORDER TO  
PREVENT ALLOPURINOL-ASSOCIATED DRUG  
HYPERSENSITIVITY.

ANUSHA DEVI NAWOOR

Thesis submitted to the University of Nottingham for the degree of  
Doctor of Philosophy (PhD)

Supervisors : Dr Then Sue Mian & Prof Ting Kang Nee

Submission date: 24th August 2020

Expected Graduation date: July 2022

## Abstract

The HLA-B\*58:01 allele was identified as a genetic marker for allopurinol-induced severe cutaneous adverse drug reactions (SCARs) in gout patients. Malaysia has a high frequency of 10.4% of the HLA-B\*58:01 allele in allopurinol-induced SCARs patients. However, the strength of association of the HLA-B\*58:01 allele to allopurinol-induced SCARs still need to be further validated in Malaysian pharmacogenetics studies. This project aims to develop a new, cost-effective, user-friendly and rapid method of screening for this allele by using the High Resolution Melt (HRM) method. The HRM method was used for its ability of reference curve-based targeted genotyping, where a positive control's melt curve is used as a reference for screening of unknown samples. A gout cohort (n=145) and a healthy volunteers cohort (n=145), matched for age, gender and ethnicity, were used for the HRM screening. The HRM method showed a sensitivity of 0%, specificity of 57%, positive predictive value of 0% and negative predictive value of 57%, due to the presence of significant limiting factors. Several significant limitations were met in this study, starting with the slow sample collection, low number of SCARs samples, positive control's heterozygosity, low primer specificity and the high level of polymorphism in the HLA-B\*58:01 allele. Moreover, the Sanger sequencing and NGS methods were used to validate the HRM method and delve into the complexity of the HLA-B alleles' role in Malaysians. The newer theory of the presence of numerous HLA-B alleles as pharmacogenetic markers in populations was also investigated. HLA-B\*58:01 was seen as a strong genetic marker in mild allopurinol-induced hypersensitivities and SCARs by Sanger sequencing. HLA-B\*58:01 positive samples identified by Sanger sequencing had high frequencies of 32.1% for two different alleles; HLA-B\*58:01:01 and HLA-B\*35:01:01. Moreover, healthy volunteer samples showed high frequencies of 28.5% for HLA-B\*58:01:01 and HLA-B\*35:01:01 alleles. Hence, all these aforementioned HLA-B alleles are identified as potential pharmacogenetic markers in Malaysia. Next Generation Sequencing (NGS) was performed on 6 gout samples and 3 Malaysian specific SNPs (rs11423052, rs151341211 and rs9279154) were identified for the HLA-B\*58:01 allele. Future studies need to focus on SNPs amplification in order to fully exploit the HRM method's strengths as a screening method.

## Table of contents

Abstract.....	2
CHAPTER 1: INTRODUCTION .....	16
1.1 Research rationale .....	17
1.2 Research approaches .....	19
1.3 Significance of study .....	21
1.4 Hypothesis of research .....	21
1.5 Aims and Objectives.....	22
CHAPTER 2: LITERATURE REVIEW.....	23
2.1 Gout .....	24
2.1.1 Discovery of gout and monosodium urate (MSU) crystals .....	24
2.1.2 Epidemiology of gout .....	25
2.1.3 Pathogenesis of hyperuricaemia.....	32
2.1.4 Clinical manifestation of gout.....	40
2.1.5 Diagnosis of gout.....	46
2.1.6 Risk factors for gout .....	52
2.1.7 Treatment of gout.....	56
2.2 Severe Cutaneous Adverse Reactions (SCARs) .....	65
2.2.1 Introduction .....	65
2.2.2 Classification of SCARs .....	67
2.2.3 Immunological SCARs mechanisms.....	75
2.2.4 SCARs treatment .....	77
2.2.5 Outcome and sequelae .....	79
2.2.6 ADRs and SCARs reports in Malaysia .....	81
2.3 Pharmacogenomics.....	86
2.3.1 Introduction .....	86
2.3.2 Human Leukocyte Antigen (HLA) genes.....	87
2.3.3 Pathogenesis of HLA induced SCARs.....	90
2.3.4 Pharmacogenomics success stories .....	93
2.3.5 Pharmacogenomics in Malaysia.....	94
2.3.6 HLA-B*58:01's role in allopurinol-induced SCARs .....	96
2.4 Existing methods for HLA allele screening.....	119
2.4.1 General HLA allele screening methods .....	119
2.4.2 HLA-B*58:01 specific screening methods.....	120

2.5 Cost-effectiveness analysis of HLA-B*58:01 screening.....	125
2.5.1 Introduction .....	125
2.5.2 Positive CEA studies .....	126
2.5.3 Negative CEA studies .....	128
2.5.4 CEA in Malaysia.....	130
2.6 New approach: The High Resolution Melt (HRM) method .....	133
2.6.1 Introduction .....	133
2.6.2 HLA screening by the HRM method.....	137
2.6.3 Multiple HLA allele theory .....	140
CHAPTER 3: METHODOLOGY .....	143
3.1 Introduction to Methodology section.....	144
3.1.1 Operational flowchart.....	145
3.2 Sample size calculation .....	146
3.2.1 Introduction .....	146
3.2.2 Sample size calculation .....	146
3.3 Sample collection.....	148
3.3.1 Consent form design and ethics approval .....	149
3.3.2 Volunteer recruitment and eligibility.....	150
3.3.3 Blood sample collection.....	151
3.3.4 Patient medical record examination.....	152
3.4 Extraction of genomic DNA.....	153
3.4.1 Preparation of reagents for DNA extraction:.....	153
3.4.2 DNA extraction protocol .....	155
3.5 Primer Design.....	159
3.6 Polymerase Chain Reaction (PCR).....	162
3.6.1 Standard PCR protocol .....	162
3.6.2 Multiplex PCR with IC.....	164
3.6.3 PCR Clean-up.....	165
3.6.4 Determination of DNA concentration and purity .....	166
3.7 Cloning of the positive control.....	168
3.7.1 Introduction .....	168
3.7.2 Preparation of cloning materials and reagents .....	170
3.7.3 LB agar preparation.....	170
3.7.4 Cloning protocol.....	172
3.8 High Resolution Melt (HRM) method .....	183
3.8.1 Introduction .....	183

3.8.2 HRM mix preparation.....	183
3.8.3 HRM run set up .....	185
3.8.4 HRM evaluation .....	186
3.9 Next Generation Sequencing .....	189
3.9.1 Introduction .....	189
3.9.2 NGS samples selection .....	190
3.9.3 NGS Sample DNA amplification and preparation .....	190
3.9.4 NGS Library Preparation .....	197
3.9.5 Bioinformatics Analysis .....	211
CHAPTER 4: RESULTS.....	221
4.1 Sample collection, DNA extraction, demographic data and clinical data.....	222
4.2 Polymerase Chain Reaction (PCR).....	228
4.2.1 Single PCR with P2 and P3.....	228
4.2.2 Multiplex PCR with IC.....	230
4.3 Cloning of the positive control.....	232
4.3.1 Gel Electrophoresis results for cloning .....	232
4.3.2 Sanger sequencing for cloning of exons .....	234
4.4 High Resolution Melt (HRM) method's results .....	235
4.4.1 Overview of HRM screening .....	235
4.4.2 Standard HRM screening result .....	236
4.4.3 Establishing a trend in the HRM results.....	241
4.5 Sanger Sequencing results .....	257
4.5.1 Sanger sequencing on 28 ADR samples .....	257
4.5.2 Further analysis done on HRM results.....	267
4.5.3 Combined HRM analysis and Sanger sequencing results for healthy volunteers.....	274
4.6 Evaluation of the HRM method .....	284
4.6.1 Sensitivity, specificity, positive predictive value and negative predictive value .....	284
4.6.2 Statistical Analysis.....	285
4.7 NGS Result compilation .....	289
4.7.1 Overall NGS Run results .....	289
4.7.2 Sequencing Quality Control .....	289
4.7.3 Standard method of NGS analysis for all SNPs .....	292
4.7.4 Analysis of NGS done on 6 gout samples.....	295
4.7.5 Result compilation and analysis the 96 NGS samples.....	302
CHAPTER 5: DISCUSSION .....	304
5.1 Building the backbone of this study.....	305

5.1.1 Demographic data.....	305
5.1.2 Severe Cutaneous Adverse Drug Reactions (SCARs).....	308
5.1.3 Optimization of DNA extraction.....	313
5.1.4 Optimization of Polymerase Chain reaction (PCR) and Cloning .....	316
5.1.5 Conclusion.....	320
5.2 Core findings: Evaluation of the HRM method .....	321
5.2.1 Decoding the HRM results .....	321
5.2.2 The HRM method: reference curve-based targeted genotyping.....	322
5.2.3 Validation and verification of the HRM method via Sanger sequencing on the gout cohort .....	325
5.2.4 Combining Sanger sequencing and HRM results to find a specific trend in results .....	333
5.2.5 Validation and verification of the HRM method via Sanger sequencing on the volunteer cohort.....	336
5.2.6 HRM screening method's performance evaluation.....	341
5.2.7 Conclusion.....	344
5.3 Core Findings: Validation of multiple pharmacogenetic markers presence by NGS .....	345
5.3.1 NGS foundation.....	345
5.3.2 NGS method evaluation.....	346
5.3.3 NGS on allopurinol-induced SCARs samples from gout cohort. ....	349
5.3.4 Single Nucleotide Polymorphisms' role in HLA-B alleles .....	351
5.3.5 Conclusion.....	351
5.4 Limitations of study.....	352
CHAPTER 6: FINAL CONCLUSION AND FUTURE WORK .....	355
6.1 Final conclusion.....	356
6.2 Future work.....	359
REFERENCES.....	361
APPENDICES .....	389

# List of Figures

Figure 2-1 Prevalence of gout worldwide in seven representative countries, shown against patients' age (part a) and gender (part b). Part a shows a linear increase in gout prevalence with age and part b shows that gout patients are mostly male in all countries (Kuo et al, 2015) .....	29
Figure 2-2 Flow diagram showing the conversion of purines to serum urate and ultimately to an excess leading to crystal formation and chronic tophaceous gout (Richette & Bardin, 2010). .....	33
Figure 2-3 Pathogenesis of hyperuricaemia is depicted, along with the risk factors affecting this whole process (Ragab et al., 2017) .....	33
Figure 2-4 Pathway for the progression of hyperuricaemia to tophaceous gout, along with its depicted inflammation response (Dalbeth et al., 2016).....	36
Figure 2-5 Clinical manifestations of gout are shown, along with the different stages of the disease and the risk factors involved. Normouricaemia progresses to hyperuricaemia, followed by MSU crystal deposition, which in turn leads to gout flares, chronic gouty arthritis and tophaceous gout (Dalbeth et al, 2019). .....	41
Figure 2-6 Figure A depicts the deposition of uric acid crystal in the metatarsophalangeal (MTP) joint, also called podagra. Figure B shows the tophi formed in the right big toe and the finger interphalangeal joints in a gout patient (Roddy, 2011).....	43
Figure 2-7 Flowchart depicting the right main recommendations for diagnosis of gout. Adapted from EULAR 2019 recommendations (Dincer et al, 2002; Richette et al, 2020). .....	47
Figure 2-8 Detection of MSU crystals in the synovial fluid of a gout patient, by using light microscopy (polarized light) and a first-order compensating filter. Black arrows show the presence of large, bright needle-shaped MSU crystals, while white arrows show presence of brick-shaped calcium pyrophosphate dihydrate crystals. (Dalbeth et al, 2019).....	49
Figure 2-9 The three examples of ultrasonography shown are (a) Intra-articular tophus in MTP joint, (b) double contour sign in the hyaline cartilage and (c) synovial effusion/hypertrophy and crystal formation (indicated by 2 red arrows) in the tendon (Ragab et al, 2017). .....	50
Figure 2-10 DECT showing two view points for MSU deposits (red deposits) in the tendon of a gout patient. (Ragab et al, 2017) .....	51
Figure 2-11 General principles of gout management with pharmacological and non-pharmacological guidelines based on EULAR recommendations (Zhang et al., 2006). .....	57
Figure 2-12 Simplified flowchart from the EULAR recommendations for management of hyperuricaemia in gout patients (Ragab et al., 2017). .....	59
Figure 2-13 The mechanism of action of allopurinol is depicted here, where it inhibits xanthine oxidase action as seen by the back T symbol. Barred lines shows where inhibition is	

taking place and arrows shows activation or consequences. Adapted from American Society of Haematology.....	62
Figure 2-14 EULAR recommendation for the management of flares in patients with gout (Ragab et al., 2017). ....	<b>Error! Bookmark not defined.</b>
Figure 2-15 List of recommendations for diet, lifestyle modification and non-pharmacological management of gout (Ragaab et al., 2017). ...	<b>Error! Bookmark not defined.</b>
Figure 2-16 Classification of ADRs into Type A and Type B and further into immune-mediated and non-immune-mediated reactions. SCARs such as SJS/TEN, HSS/DRESS and AGEP are found in the Type 4: Delayed T-cell-mediated reaction box on the top right corner (Bharadwaj et al., 2012).....	66
Figure 2-17(a) Patient with dusky red confluent macular lesions, along with tense bullae and large areas of dermoepidermal detachment (shown by asterisk). Figure 1 (b) shows the typical distribution of skin and mucous membrane lesions in SJS/TEN. Figure 1 (c) shows overt mucous membrane involvement and moderate skin detachment with crusts in an SJS case (Chung et al., 2008).....	69
Figure 2-18(A) Patient with DRESS showing scaly patch with papules on his forearm and Figure 2.18 (B) shows desquamation of his soles with petechiae visible. (Choudhary et al., 2013) .....	74
Figure 2-19 Schematic diagram of the different mechanisms of action for immune stimulation of the Major Histocompatibility Complex (MHC) involved in SCARs development. Figure 2.19 A & B depict the hapten and pro-hapten models while C and D show the p-i and the altered protein repertoire concept (Nguyen et al., 2019). ....	92
Figure 2-20 This figure shows the potential mechanisms which are involved in the allopurinol-induced SCARs. The hapten concept, p-i concept and the altered repertoire model are depicted. The interaction of allopurinol and oxypurinol with the HLA0B*58:01 molecule can be clearly seen here (McDonagh et al., 2014).....	98
Figure 2-21 The characteristic HRM melting profile principle where an increase in temperature causes the denaturation of dsDNA to ssDNA. Fluorescence is shown to be inversely proportional to temperature rise as per the HRM principle. Figure A shows the original fluorescence data, while figure B shows the derived normalized melt curve after background fluorescence subtraction (Reed, 2007).....	135
Figure 3-1 Complete operational flowchart, showing all the different methods used in sequential order (DIY). ....	145
Figure 3-2 Flowchart showing the planning process before sample collection and the flow of sample collection done in hospitals (DIY). ....	148
Figure 3-3 RBC lysis was carried out first, where the steps 1 to 3 were repeated 3 times until a white pellet was left. (DIY) .....	157
Figure 3-4 DNA extraction step which was done after RBC lysis, started with a white pellet and ended with the extracted DNA, dissolved in TE buffer. (DIY).....	158
Figure 3-5 Flow of step-wise PCR optimisation done for all primers designed.....	163

Figure 3-6 First part of cloning process depicted, starting from ligation, transformation, plating, spreading and finally blue-white colony formation. (DIY).....	175
Figure 3-7 Streaking of the white colonies by using the quadrant streak technique. (DIY)..	176
Figure 3-8 Subculture of the successful white clones in LB broth for further increase in copy number and subsequent extraction of plasmids. (DIY) .....	178
Figure 3-9 General flow of the NGS method, starting from samples collection to the end step of running all the samples on the MiSeq machine. ....	189
Figure 3-10 The simplified flow of the seven stages for sample amplification and preparation. ....	191
Figure 3-11 Breakdown of the steps along with their hands-on time and running time required, in order to effectively fit all of them in the first half of the NGS preparation day. ....	191
Figure 3-12 NGS library preparation workflow, which was carried out and repeated for every single sample.....	198
Figure 3-13 Breakdown of each step with the required time for hands-on preparation and the actual running time. ....	198
Figure 3-14- Last step for the NGS preparation and the final running of all the libraries on the MiSeq system and following result generation. ....	199
Figure 3-15 Simplified flowchart of the whole bioinformatics pathway used to analyze the NGS run results generated.....	213
Figure 3-16 Breakdown of the full bioinformatics analysis pathway with the input format, software used and output generated.....	214
Figure 3-17 The full timeline of the designed screening kit as implemented in a hospital setting for HLA-B*58:01 allele testing before drug prescription to gout patients.....	<b>Error!</b>
	<b>Bookmark not defined.</b>
Figure 3-18 The Block Diagram of the HLA-B*58:01 screening system. The NI USB-6003 device was used to record fluorescence signal and control the temperature of thermal electric module. The device connected to the PC via universal serial bus for analysis. ...	<b>Error!</b>
	<b>Bookmark not defined.</b>
Figure 3-19 The six well PCR block which was designed using Autodesk Fusion 360. The block was designed to fit the 0.1 mL, low profile, PCR tube snugly. ...	<b>Error! Bookmark not defined.</b>
Figure 3-20 Block diagram of the PTC10K-CH feedback control system. The system maintained the temperature of PCR block by comparing the output voltage of NI USB 6003, $V_i(t)$ against the NTC thermistor output voltage, $V_t(t)$ .....	<b>Error! Bookmark not defined.</b>
Figure 3-21 The side view of the thermal cycler chamber. The PCR block was located at the centre of heating chamber with the 6 grooves for the PCR tubes.	<b>Error! Bookmark not defined.</b>
	<b>Bookmark not defined.</b>
Figure 3-22 Single channel fluorescence detection chamber for the HLA-B*58:01 screening system. The LED was located on the side of the PCR block to excite the fluorescence dye. The silicon photomultiplier was located on top of PCR tube to detect the fluorescence light. ....	<b>Error! Bookmark not defined.</b>

Figure 3-23 The developed difference melt curve plot on the GUI for the HLA-B\*58:01 screening test. The processed data was plotted on the black panel. The final analyzed result is shown on the bottom, right-hand side of the graphical user interface.**Error! Bookmark not defined.**

Figure 4-1 Amplification of exon 2 (Figure A) and exon 3 (Figure B) of the HLA-B\*58:01 allele. 100bp ladders are shown starting from the left, followed by the Non-template control (NTC), P2 product at 149bp (Figure A) and P3 product at 249bp (Figure B). .....228

Figure 4-2 Multiplex PCR done in triplicate for I2 and I3, where I2 combined IC+P2 primers and I3 combined IC+P3 primers. IC primers produced a product at around 423bp, P2 around 149bp and P3 around 249bp. ....230

Figure 4-3 Cloning of the exon 2 and 3 of the HLA-B\*58:01 allele into the pDrive vector from Qiagen and subsequent analysis by double restriction digestion by BamH1 and Xho1. Figure 4.3-A shows the presence of the digested target exon 2, which is less than 200bp, for plasmids 1 and 2. Figure 4.3-B shows 4 digested plasmids (~4000bp) with their target exon 3 around 250bp.....233

Figure 4-4 Derivative Melt plot for HLAG1 (green colour) and the positive control (blue colour). .....238

Figure 4-5 Normalized melting curve for HLAG1 (green colour) and the positive control (blue colour). .....240

Figure 4-6 Difference melting curve for HLAG1 (green colour) and the positive control (blue colour). .....240

Figure 4-7 Compilation of 3 graphs for QC check, where graphs A and B showed Data by cycle against %>=Q30 and FWHM respectively, followed by graph C showing Data by Lane against density. ....291

Figure 4-8 The gel electrophoresis results shows the six amplified DNA samples (S1, S2, S3, S4, S5 and S6) amplified by our developed PCR system in (a) and was compared to results from the Eppendorf Mastercycler nexus gradient machine in (b). Both gels show the 6 samples amplified at the correct target of 150bp (Chong et al., 2018).**Error! Bookmark not defined.**

Figure 4-9 The final difference melt curve method of the HLA-B\*58:01 screening system showing a positive and negative screening result. The positive and negative threshold were shown by the yellow dotted lines which were set at 45 and -45. Figure 8 (a) showed a maximum peak value exceeding the positive threshold value, thus showing a negative screening result. Figure 8 (b) showed a maximum and minimum peak value less than the threshold values, therefore resulting in a positive screening result.**Error! Bookmark not defined.**

# List of Tables

Table 2-1 Gout prevalence estimates in developed and developing countries. GP stands for general population (Kuo et al, 2015). .....	27
Table 2-2 Incidence recorded for seven main countries worldwide, along with different studies and their specific duration. Three entries have multiple incidence recorded in this table, recorded at different time frames in the same study (Kuo et al, 2015). .....	31
Table 2-3 Detailed clinical features to distinguish SJS, SJS/TEN overlap and TEN. Plus (+) sign indicates strength of confluence, where + is lower than ++ (Chung et al., 2008).....	70
Table 2-4 Seven independent risk factors used to classify SJS/TEN as part of SCORTEN scale. The scale of 0 or 1 shown in the last two columns are attributed according to the risk factors they fall in (Bastuji-Garin et al., 2000). .....	71
Table 2-5 Second part of SCORTEN estimation by using Table 1's initial calculated scores and linking it to an estimate of the possible mortality rate in terms of percentage (Bastuji-Garin et al., 2000). .....	71
Table 2-6 Overview of SCARS treatment, outcome, main sequelae and management after resolution. (RCT= Randomised controlled trial) (Chung et al., 2008).....	78
Table 2-7 Highest global allelic frequency mined for top fourteen hits and countries (Gonzalez-Galarza et al., 2020) .....	99
Table 2-8 Summary of studies reporting the association of HLA-B*58:01 and allopurinol-induced ADRs (including cADRs and SCARs) worldwide. CI: Confidence interval; ADRs: Adverse drug reactions; SCARs: Severe cutaneous adverse drug reactions; cADRs: Cutaneous adverse drug reactions; SJS: Stevens-Johnson syndrome; TEN: Toxic epidermal necrolysis (Wu et al., 2016; Chang et al., 2020) .....	104
Table 2-9 Simplified search results from the database for Malaysia HLA-B*58 and HLA-B*58:01 hits (Gonzalez-Galarza et al., 2020).....	112
Table 2-10 Detailed list of all the ethnolinguistic groups of Malaysia (The Joshua Project, joshuaproject.net, retrieved 2015).....	113
Table 2-11 Full table for all studies mined for the Malaysian population. (SSOP-sequence-specific oligonucleotide probes, SSP-sequence specific primer, SBT-sequence based typing) (adapted from Gonzalez-Galarza et al., 2020).....	116
Table 2-12 Existing HLA-B*58:01 screening methods with their specific machines/kit used, cost, run time and type of analysis required. ....	120
Table 3-1 Full primer details for the internal control and 2 pairs of primers used to amplify the HLA-B*58:01 allele, including the primer sequence, length, Tm and GC %.....	161
Table 3-2 Detailed volume of the ligation mixture components to be prepared. ....	172
Table 3-3 Components of the restriction digestion mixture to be added in two steps. ....	182
Table 3-4 Breakdown of the HRM reaction contents with their detailed volume used and the final concentration needed.....	184
Table 3-5 The fully detailed cycling steps for the initial PCR and following HRM step done for all samples during HLA-B*58:01 screening.....	184

Table 3-6 Calculation of sensitivity and specificity in a two-by-two table, where TP stands for True positive, FP for False positive, FN for False negative and TN for True negative. (Wong & Lim, 2011). .....	187
Table 3-7 Different statistical tests used paired to the different formulated hypotheses to be tested. ....	188
Table 3-8 Full details for the HLA-B forward and reverse primers used for the NGS run. ....	192
Table 3-9 Full details for the HLA-B amplification for the NGS run, including component breakdown, their volume and final concentration. ....	193
Table 3-10 Thermal cycler program for HLA-B amplification on the Veriti 96 well thermal cycler (Applied Biosystems, USA) .....	193
Table 3-11 Extract from the IEM index labelling sheet. ....	200
Table 3-12 Amplification conditions for libraries .....	202
Table 4-1 Summary of the demographic and clinical data for the 145 gout patients and 145 healthy control samples.....	222
Table 4-2 Seven samples identified with a smaller DNA concentration. ....	223
Table 4-3 Documentation of ADRs/SCARs observed in the 28 gout patients along with the percentage of patients affected. ....	225
Table 4-4 Twenty eight ADR/SCAR samples linked to their detailed allergic reactions, gender, ethnicity and age.....	226
Table 4-5 Sanger sequencing results done on sample HLAG26 for P2 and P3 primers. ....	230
Table 4-6 Sanger sequencing results for the 2 cloned plasmids containing the exon 2 insert. ....	234
Table 4-7 Sanger sequencing results for the 4 cloned plasmids containing the exon 3 insert. ....	234
Table 4-8 HRM results compilation generated on amplification, with C <sub>q</sub> and T <sub>m</sub> values. ....	237
Table 4-9 HRM normalized and difference melt curve compilation for all the 28 gout samples with ADRs. The positive control's duplicate in all the graphs are shown by the dark blue horizontal line, along with the blue downward peak. The gout samples compared to the positive control are all in different colours for easier comparison. ....	243
Table 4-10 HLA allele calling with the Sanger Sequencing results of the 28 gout samples with ADRs from NCBI and IMGT/HLA website. The final 2 HLA-B alleles for each individual were shown in the fourth column, where the highest hit was called and linked to the ADRs observed in the samples. HLA-B alleles are simplified to B* and HLA-C alleles to C*. Some samples will have several rows due to the presence of numerous and corresponding SNPs. ....	260
Table 4-11 Percentage of each HLA-B allele found in all the 28 ADR samples, as per Sanger sequencing. ....	265
Table 4-12 Comparison of possible HRM HLA-B*58:01 positive samples to HLA-B*58:01 positive samples in Sanger sequencing, for the 28 ADR samples. ....	266
Table 4-13 Standard SNP identification method by using base change and melting point of samples as variables. ....	267
Table 4-14 Base changes seen in the 9 possible positive samples, along with their linked HLA alleles and T <sub>m</sub> change.....	269

Table 4-15 Base changes seen in the 19 negative samples, along with their linked HLA alleles and T <sub>m</sub> change.....	270
Table 4-16 Compilation of all the different variables used to identify a trend in the HRM results for the 9 possible positive ADR samples. Sanger sequencing results from Table 4.11 were also added into 2 columns based on their databases. HRM difference melt curves and ADRs observed were shown in the last two columns.....	271
Table 4-17 Summary of HLA-B alleles present in the 14 sequenced normal samples. ....	276
Table 4-18 14 random normal samples Sanger sequencing results for P3 primer and exon 3. ....	277
Table 4-19 Fisher’s exact test for the association in the HLA-B*58:01 allele’s distribution between the healthy volunteer cohort and the gout cohort. ....	285
Table 4-20 Fisher’s exact test for HLA-B*58:01’s association with the different ethnic groups in the gout cohort. Pearson Chi-squared test was also carried out in parallel. ....	286
Table 4-21 Fisher’s exact test for HLA-B*58:01’s association with the different ethnic groups in the healthy volunteer cohort. Pearson Chi-squared test was also carried out in parallel. ....	287
Table 4-22 Test of Homogeneity of Variances for all the hypersensitivity observed.....	288
Table 4-23 One Way ANOVA carried out for HLA-B*58:01’s presence in the different types of allopurinol-induced hypersensitivity reactions seen. ....	288
Table 4-24 Per Read metrics extracted from BaseSpace, showing the yield, error rate and %≥Q30. ....	290
Table 4-25 Full details of how SNPs and HLA allele hits were analysed for sample HLAG12. ....	293
Table 4-26 SNPs identified from NGS results accompanied by HLA allele calling for sample HLAG12. ....	296
Table 4-27 SNPs identified from NGS results accompanied by HLA allele calling for sample HLAG19. ....	297
Table 4-28 SNPs identified from NGS results accompanied by HLA allele calling for sample HLAG26. ....	297
Table 4-29 SNPs identified from NGS results accompanied by HLA allele calling for sample HLAG28. ....	298
Table 4-30 SNPs identified from NGS results accompanied by HLA allele calling for sample HLAG29. ....	298
Table 4-31 SNPs identified from NGS results accompanied by HLA allele calling for the positive control. ....	299
Table 4-32 Comparison of the Sanger sequencing and NGS results for the 6 gout samples. Common alleles observed and the SCARs identified are also included for analysis purposes. ....	301
Table 4-33 Compilation of the 3 cohorts’ common HLA alleles and their percentage. ....	303

## **Abbreviations:**

ADR- Adverse Drug Reaction

AED- Anti-Epileptic-Drug

ALLO- Allopurinol

AHS- Allopurinol Hypersensitivity Syndrome

ABC- Abacavir

AFND- the Allele Frequency Net Database

BSA- body surface area

CBZ- Carbamazepine

CLC- Colchicine

cADR- cutaneous Adverse Drug Reaction

DRESS- Drug reaction with eosinophilia and systemic symptoms

FEB- Febuxostat

HLA- Human Leukocyte Antigen

LAMP- loop-mediated isothermal amplification

MPE- Maculopapular Eruption

MHCI- Major histocompatibility complex class I

NPV- Negative predictive value

PGx- Pharmacogenomics

PCR- Polymerase Chain Reaction

PPV- Positive predictive value

RFLP- Restriction Fragment Length Polymorphism

SCAR- Severe Cutaneous Adverse drug Reaction

SJS- Steven Johnson Syndrome

SNP- Single Nucleotide Polymorphism

SSP- Sequence specific priming

SSO- Sequence Specific Oligonucleotide Typing

TEN- Toxic Epidermal Necrolysis

T<sub>m</sub>- Melting point

## Acknowledgments

**“As we express our gratitude, we must never forget that the highest appreciation is not to utter words, but to live by them.” -John F. Kennedy**

I would like to express my sincere appreciation to all the people who played an important role throughout my whole PhD. Starting with my supervisor, Dr Then Sue Mian, who has supported and guided me from the start of this project. Thank you to Dr Then, for trusting me with your project. I hope I fulfilled a part of your vision for this research. A big thanks goes to Prof Ting Kang Nee, who's always been a constant source of support and inspiration. I would like to thank all the staff and laboratory technicians of the Department of Biomedical Sciences, The University of Nottingham Malaysia Campus. Thank you to Asma and Siti who have been a big support in lab since 2015.

Thank you to everyone in HUKM hospital who helped in sample collection. Starting with Prof Shahrir, our main collaborator, who helped deal with patients and collect blood samples. A special thanks goes to Mrs Azura in the Rheumatology department, for being my pillar of support and for the kindness that she showed me as she helped me adapt to the hospital environment. Thank you to all the people in UMBI for the help in DNA extraction, sample collection, storage and the NGS run. Thank you to Dr Azian and Zam, our collaborators, who supported this research from the start. My biggest thanks goes to Mrs Sheda in UMBI. She's been by my side in lab from the start in UMBI, helping me constantly with unwavering kindness, even in the most stressful times.

Thank you to Dr Liza from Hospital Putrajaya for being so welcoming and kind, throughout the whole sample collection process, even while we were juggling with countless patients every day. Without your help, this study wouldn't have achieved one of its most important targets.

Now, for the people closest to my heart, my friends and fellow PhD candidates; Grace, Rachel and May Zie. Thank you for your support, love and help for the past five years. We wouldn't have made it without each other as our pillars and our biggest supporters! A huge huge thank you to Grace, the best PhD senior, for helping me get through the tough HUKM times and for encouraging me from the start.

Big thanks to all my supportive and loving Nottingham friends.

A huge thanks to my family for believing in me from the start and for supporting me in all my endeavours. Thank you will never be enough for what you do for me.

Thank you to each and everyone from the bottom of my heart. So much blood, sweat and tears have contributed to make this research what it is today. I hope this thesis does it justice and that I live up to everyone's expectations.

# CHAPTER 1: INTRODUCTION

## 1.1 Research rationale

Gout is a purine metabolism disorder caused by depositions of urate crystals in joints, which is in turn brought about by long-lasting hyperuricemia (Richette & Bardin, 2010). The worldwide prevalence of gout is 0.1-10%, with an annual incidence of 0.3-6 cases per 1000 persons-years and has an incidence of 2-6 fold more in men, when compared to women. Gout's incidence in Malaysia was at 1% in 2016, totalling up to around 31,000 patients affected (Kuo et al., 2015). Allopurinol is the first-line, urate-decreasing therapy commonly used in gout patients due to its rapid lowering effect on serum urate levels and affordable cost over other drugs. Allopurinol and its metabolite oxypurinol inhibits xanthine oxidoreductase and leads to a decreased urate production and *de novo* purine synthesis (Day et al., 2017; Kwok & Kwong, 2013).

However, allopurinol use may lead to occurrences of adverse drug reactions (ADRs) in patients which may lead to mortality. This type of ADR is classified as Type B ADRs, also known as idiosyncratic ADRs, which are either immune-mediated or non-immune related. These ADRs can manifest as mild maculopapular eruptions (MPE), the life-threatening severe cutaneous adverse reactions (SCARs) which includes Stevens-Johnson syndrome/toxic epidermal necrolysis (SJS/TEN), drug reaction with eosinophilia and systemic symptoms (DRESS), as well as allopurinol hypersensitivity syndrome (AHS) (Stamp et al., 2016). Allopurinol is one of the highest causes of SCARs, accounting for around 5% of total SCARs cases. DRESS syndrome is characterized by skin rashes, eosinophilia, lymphadenopathy, fever and multi-organ dysfunction, with a 10% mortality rate in patients (Tsai & Yeh, 2010). SJS and TEN are characterized by cellular apoptosis, mucous membrane erosion, severe detachment of the epidermis and constitutional symptoms. SJS has a mortality rate of 5-10% with body surface area (BSA) detachment of less than 10%, while TEN has a mortality rate of 30-40% with BSA detachment of more than 30% (Bharadwaj et al., 2012; Roujeau et al., 1995).

The pathogenesis of all these hypersensitivity reactions are mediated by the immune system, more specifically by the Major histocompatibility complex class I (MHCI) molecules which consists of the Human Leukocyte Antigen B (HLA-B) alleles (Bharadwaj et al., 2012;

Chung et al., 2008). Hence, the pharmacogenomics (PGx) field was born, linking drugs' ADRs to HLA-B alleles and specific populations. PGx and its implications in medical science skyrocketed in the 2000s, along with the Human Genome Project's completion (Collins & McKusick, 2001). Several pharmacogenetic markers have been identified, starting with HLA-B\*57:01 in abacavir (ABC)-induced SCARs in HIV patients (Mallal et al., 2008) and followed by HLA-B\*15:02 in carbamazepine (CBZ)-induced SCARs in epileptic patients (Chung et al., 2004).

The HLA-B\*58:01 allele is the most well-established genetic link to allopurinol-induced hypersensitivity reactions (Jung et al., 2011). The mechanisms behind this genetic association follows the HLA-dependent T cell activation by the drug, whereby CD8+ T cells activation will be accompanied by cytotoxic effects on keratinocytes which ultimately results in the characteristic epidermal aggression and necrolysis seen in SCARs (Bidwell, 1994; McDonagh et al., 2014). Gout patients risk SCARs development 80-97 times more in the presence of the HLA-B\*58:01 allele (Choo et al., 2014). The genetic association between HLA-B\*58:01 and AH has been validated in numerous countries, including Taiwan (Hung et al., 2005), Thailand (Tassaneeyakul et al., 2009; Jantararoungtoun et al., 2014; Sukasem et al., 2016), China (Cao et al., 2012; Chiu et al., 2012), Hong Kong (Cheng et al., 2015), Japan (Kaniwa et al., 2008; Tohkin et al., 2013; Niihara et al., 2013), Korea (Jung et al., 2011; Kang et al., 2011), and Europe (Christallo et al., 2011; Lonjou et al. 2008). Malaysia was shown to have a high frequency of 10.4% for the HLA-B\*58:01 allele in SCARs, as recorded in the Allele Frequency Net Database (AFND) (González-Galarza et al., 2015). Therefore, it has been proposed that HLA-B\*58:01 screening prior to allopurinol prescription may effectively reduce incidences of allopurinol-induced ADRs and minimize morbidity and mortality (Jung et al., 2011).

## 1.2 Research approaches

The screening of HLA alleles started with serological testing, followed by Restriction Fragment Length Polymorphism (RFLP), Polymerase Chain Reaction (PCR), Sequence specific priming (SSP), Sequence Specific Oligonucleotide Typing (SSO) and loop-mediated isothermal amplification (LAMP). However, these methods had several disadvantages, namely, being time-consuming, inefficient, using large amounts of reagents and low specificity due to high HLA polymorphisms (Choo, 2007; Pozzi et al., 1999). The gold standard for HLA typing is DNA sequencing, but it is too time consuming, especially when immediate results are needed in hospitals for subsequent prescriptions (Bidwell, 1994).

This study introduces the use of the High Resolution Melt method (HRM) as a rapid and cost-effective screening method for the HLA-B\*58:01 allele, which overcomes the shortcomings of previous screening methods used. The HRM method has been extensively used for genotyping, single nucleotide polymorphism (SNP) scanning and HLA donor matching in organ transplantation (Reed, 2007). The HRM method can accurately differentiate between different DNA sequences based on their GC composition, melting point, DNA length and complementarity by using special DNA binding dyes. During the HRM method, temperature is increased in short increments of 0.008-0.2°C allowing detailed melting curve generation and analysis where single base changes are detected in otherwise identical DNA sequences (Reed, 2007). The HRM method can detect the slightest change in one DNA base, which will then be reflected in a change of the HRM melt curve shape. This theory is called the reference curve-based targeted genotyping, and it has been used in a few HLA studies only (Rani et al., 2018; Imperiali et al., 2015; Cui et al., 2013; Lundgren et al., 2012; Zhou et al., 2004). HRM also provides the advantages of a quick, qualitative screening method, with lower costs, high accuracy, sensitivity and specificity (Zhou et al., 2004). By using the unique features of the HRM method, a screening method for the HLA-B\*58:01 allele will be designed and optimised in this study, in order to decrease allopurinol-induced ADRs. Sanger sequencing, the gold standard for HLA typing, will be used to validate the HRM method developed and to confirm the presence of the HLA-B\*58:01 allele in positive samples identified.

Owing to the HLA alleles' high number of polymorphisms, even Sanger sequencing has its limits in HLA typing and in uncovering their full potential. Here, the Next Generation Sequencing (NGS) method comes into play to decipher the ambiguities and gaps left in the understanding of the HLA-B alleles' role in drug-induced SCARs in diseases and infections. The high resolution and high-throughput capability of NGS will allow us to significantly resolve the current HLA-B sequence coverage gap and fully characterize the multifaceted HLA-B genes (Wang et al., 2012). Moreover, NGS will also be used as the ultimate validation for both the HRM and Sanger sequencing methods. One of the biggest hurdles of the PGx community is the translation of research into clinical application (Chang et al., 2020).

### 1.3 Significance of study

Malaysia was shown to have a high frequency of 10.4% for the HLA-B\*58:01 allele in SCARs, as recorded in the AFND (González-Galarza et al., 2015). However, studies on the AFND were only anthropology and transplantation studies, without one single PGx study. Recently, a study was published, proving the role of the HLA-B\*58:01 allele in allopurinol-induced SCARs in Malaysia (Low et al., 2020).

HLA typing methods, sequencing and commercial kits are too expensive for Malaysians, showing the need for development of a low cost and effective screening method, thus allowing anyone to predict their reactions to drugs based on their unique genetic makeup. This project aims to develop a new screening procedure by exploiting the HRM method which offers numerous advantages over all existing HLA screening methods. Moreover the NGS method will be used to decode the HLA-B\*58:01 allele into depth to give us an insight of its complexity and involvement in allopurinol-induced SCARs in Malaysia. The possible presence of other HLA-B pharmacogenetic markers and SNPs will be investigated too and give us a deeper understanding about their mechanisms of action.

### 1.4 Hypothesis of research

The hypothesis of this research states that the HLA-B\*58:01 allele is strongly linked to allopurinol-induced SCARs in Malaysian gout patients. Moreover, the HRM method is hypothesized to be an accurate screening method to use in order to detect the HLA-B\*58:01 allele. Another hypothesis states that other HLA-B alleles might be linked to SCARs and other milder forms of ADRs in Malaysians. These hypotheses will be thoroughly investigated in this research.

## 1.5 Aims and Objectives

- 1) Design specific primers for amplification of exon 2 and 3 of the HLA-B\*58:01 allele and validate it by Sanger sequencing.
- 2) Optimize a standard and a multiplex Polymerase Chain Reaction (PCR) for amplification of the HLA-B\*58:01 allele and validate this by Sanger sequencing.
- 3) Develop and optimize the HRM method with the designed primer sets and the cloned positive control as reference genotype. Validate the HRM screening method with Sanger sequencing.
- 4) Evaluate the sensitivity, specificity, positive and negative predictive value of the HRM method.
- 5) Use NGS to further validate HRM and Sanger sequencing findings.

# CHAPTER 2: LITERATURE REVIEW

## 2.1 Gout

### 2.1.1 Discovery of gout and monosodium urate (MSU) crystals

Gout is a disease of antiquity, found in the skeletal remains of pharaohs in Egypt (2620 BC to 2480 BC and even in the fossil of dinosaurs such as the *Tyrannosaurus rex* (Hartung, 1957). It was also known as the 'disease of kings' with its long history dating back to around 3000 B.C. Gout is a purine metabolism disorder caused by depositions of urate crystals in joints which is in turn brought about by long-lasting hyperuricaemia. The prevalence of gout is constantly increasing and may be attributable to changes in diets, lifestyle and increased ageing population. Gout affects millions of people worldwide, and understanding the disease is key to prevention and cure (Richette & Bardin, 2010). Gout is nowadays extensively understood and properly treated as a true crystal deposition disease which results from the formation of monosodium urate (MSU) crystals. Uric acid (UA) was first discovered by Karl Wilhelm Scheele in 1776 (Scheele & Beddoes, 1786), followed by the discovery of W. H. Woolaston three years later which demonstrated the presence of sodium urate in his own tophus (Fraser, 1992). Sir Alfred Garrod first postulated uric acid as the cause of joint inflammation and he also devised the first qualitative test for uric acid, known as the "thread test" (Mellen et al, 2006). In 1897, MSU crystals were thought to be the cause of inflammation by Gustave Riehl and this was later proven by Wilhelm His Jr. MSU crystals were later linked to both inflammation and necrosis by Freudweiler who performed the histological study on rabbits and chickens injected with MSU crystals (Wortmann et al, 2006). The same experiment was repeated on a human volunteer in the early 1960s and showed an intense inflammatory response in the injected joint (Buchanan et al, 1965). Hence, following this series of experiments over time, MSU crystals were identified unequivocally as the pathogenic agent in gout. Gout's concept has been elusive for centuries and was only cemented after the aforementioned experiments by numerous scientists. This disease now has distinct phenotypic characteristics, such as, high serum urate concentration (hyperuricaemia), recurrent acute arthritis attacks, MSU crystals accumulation around joints, possible renal disease and uric acid urolithiasis (Firestein, 2009).

### 2.1.2 Epidemiology of gout

Measuring a disease's occurrence comprises of the prevalence and incidence, where definitions are the number of cases per population at a given time and the number of the new cases per population at a period of time, respectively. The worldwide prevalence of gout is 0.1-10%, with an annual incidence of 0.3-6 cases per 1000 persons-years. The incidence of gout in men is 2-6 fold more than in women. The worldwide incidence is increasing gradually due to poor dietary habits, sedentary lifestyle, increased obesity and metabolic syndrome. Western countries have a prevalence of 3-6% in men and 1-2% in women, and may even go up to 10% in some. Prevalence increases to 10% in elderly men and 6% in elderly women of more than 80 years old (Kuo et al, 2015).

Different country in the Asia-Pacific regions have varied prevalence, with Taiwan (upto 10.42%) being amongst the top countries worldwide with the highest prevalence of gout. China (6.10%) and Malaysia (1%) were found to have higher gout levels compared to Japan (0.51%), Thailand patients (<0.50%) and other Asian populations (Poór & Mituszova, 2003) as shown in Table 2.1. In 2016, 34 million cases of gout for people older than 15 years old were recorded worldwide and an approximate forecast for the year 2025 is of 37.9 million cases. Gout's incidence in Malaysia was at 1% in 2016, totalling up to around 31,000 patients affected (Kuo et al, 2015).

### 2.1.2.1 Prevalence

Global prevalence cannot be captured in one single estimate, as it varies from country to country and data is lacking in many regions. The highest prevalence (>10%) was recorded in Oceanian countries, specifically in Taiwanese aboriginals (Chang et al., 1997; Chou et al., 1998) and Maori people (Rose and Prior, 1963; Pascart et al., 2014). North America and Western Europe have a prevalence of around 1-4% as detailed in Table 2.1, while other countries are reported to have an even smaller prevalence, such as Guatemala, Iran, Turkey, etc (Obregón-Ponce et al, 2012; Davatchi et al., 2008; Cakir et al., 2012). Factors influencing the estimates of the prevalence of gout vary according to differences in sample number, age, gender distribution, geographic locations, sampling methods, gout definition used and ethnicities of study patients (Wijnands et al., 2015). Figure 2.1 shows two graphs with prevalence data recorded for seven main countries, against age (part a) and gender (part b). Figure 2.1 (a) shows that age-specific prevalence of gout is similar in the UK, USA and New Zealand, whereas Taiwan has the highest gout prevalence across all ages. However, irrespective of prevalence in different countries, the latter always increases with age and plateaus around 70 years old. Figure 2.1 (b) shows the higher gout prevalence in men compared to women in all the seven countries, in a ratio of around 4:1 (male: female) (Kuo et al, 2015).

Table 2-1 Gout prevalence estimates in developed and developing countries. GP stands for general population (Kuo et al, 2015).

Countries	Estimated prevalence (%)	Reference
<b>America</b>		
Mexico	0.30-0.40	65, 66
Cuba	0.30-0.40	67
Venezuela	0.30-0.40	68
Guatemala	0.01	20
Jamaica	0.80	69
<b>Canada</b>	3.00	37
<b>Europe</b>		
Greece	4.75	38
United Kingdom (UK)	2.49	39
Spain	2.49	40
Netherlands	2.49	41
Germany	1.49	42
France	0.90	43
Italy	0.91	44
Portugal	0.30	45
Czech Republic	0.30	46
<b>Australia and New Zealand</b>		
Australia (GP)	1.44	48
Pacific Islanders	7.63	49
Maori	6.06	49
European descent	3.24	49
<b>Asia</b>		
Japan	0.51	50
South Korea	0.40	51
China	5.10-6.10	52
Northern China	1.14	78

Singapore	4.10	53
Taiwan (GP)	4.92	54
Taiwanese aboriginals	10.42	14
Indonesia	1.70	70
Malaysia	1.00	26
Kuwait	0.80	70
Bangladesh	<0.50	71
India	<0.50	74
Northern India (Jammu)	0.82	82
Iran	<0.50	21
Pakistan	<0.50	75
Philippines	<0.50	27
Thailand	<0.50	76
Vietnam	<0.50	77

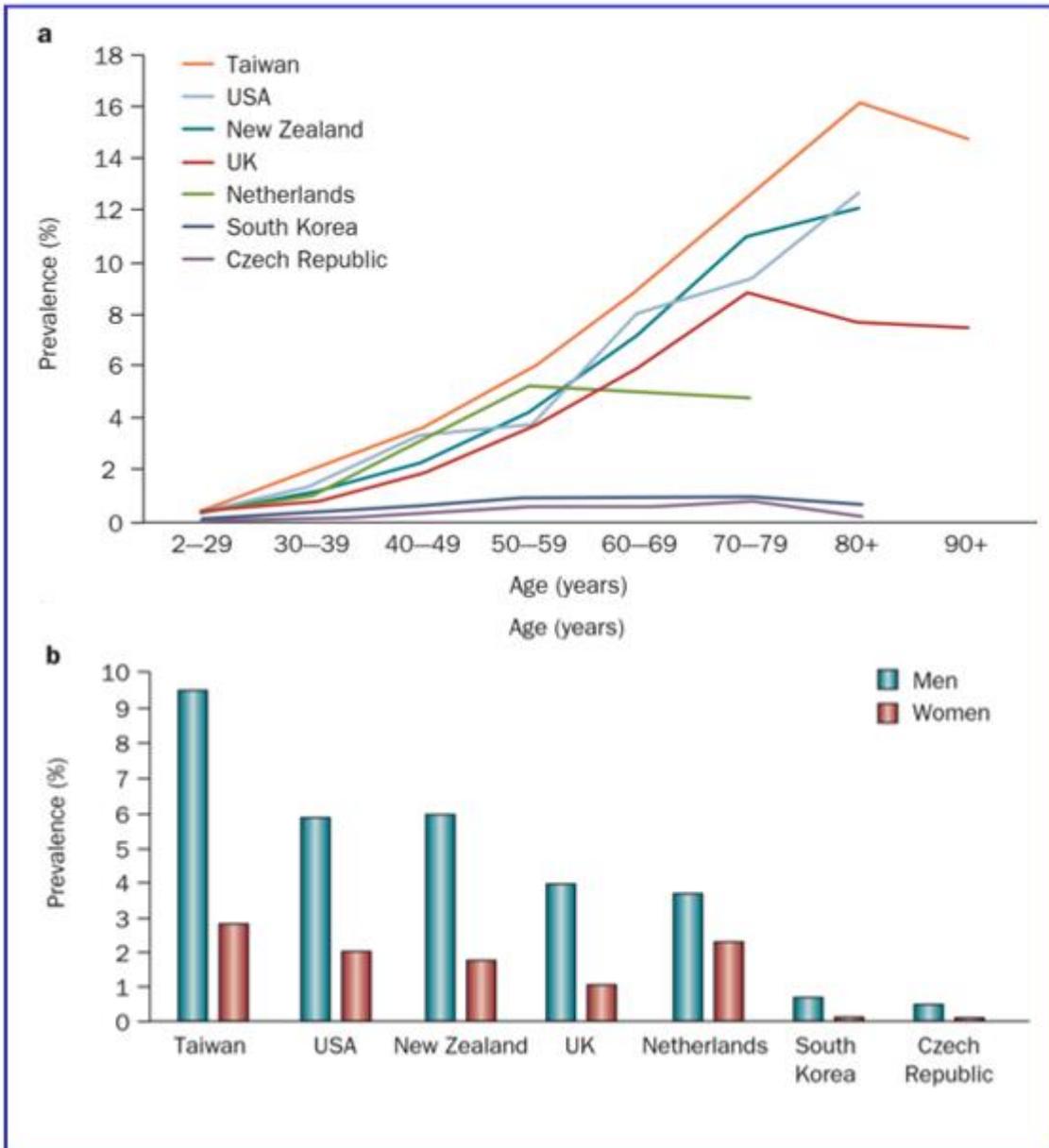


Figure 2-1 Prevalence of gout worldwide in seven representative countries, shown against patients' age (part a) and gender (part b). Part a shows a linear increase in gout prevalence with age and part b shows that gout patients are mostly male in all countries (Kuo et al, 2015)

### 2.1.2.2 Incidence

Studies on the incidence of gout are fewer compared to the global prevalence studies. An estimate for the general gout incidence is 2-6 fold higher in men than in women. Incidence is also known to increase with age and reaches a plateau around 70 years, similar to gout prevalence. Numerous studies have reported the incidence of gout in the USA, namely the Framingham study with 0.84 cases per 1000 person years (Abbott et al., 1988) and another study in Sudbury with 1.0 cases per 1000 persons, as shown in Table 2.2 (O'Sullivan et al., 1972).

However, with different studies there are different case definitions, age limits, sampling strategies and location, thus making it difficult to set a proper trend for incidence. Interestingly, when the Rochester Epidemiology Project used a case-definition of physician-diagnosed gout, along with an age and sex adjusted factor, the incidence decreased to half compared to the other three aforementioned studies. Moreover, an increase in incidence was also observed in the second part of the study done until 1996, as shown in Table 2.2 (Maynard et al., 2014). This increase in incidence with time was also seen in three other studies, namely the "USA Rochester Epidemiology Project, the UK Second and Third National Studies of Morbidity and the UK Clinical Practice Datalink (CPRD). The highest gout incidence recorded from Table 2.2 is from the Maori people of New Zealand (6.45), followed by Taiwan (2.74) and a UK Health Improvement Network study (2.68). The Singapore Chinese health cohort study reported an incidence rate of as 0.5 per 1000 person-years in men and 0.294 per 1000 person-years in women. However, there is no incidence yet reported for Malaysia and it is assumed to be the same as Singapore, due to their similar population demographics Hence, incidence rate will also vary from country to country and based on numerous other factors like gout definitions used, patients' age, sampling techniques used, length of study, etc.

Table 2-2 Incidence recorded for seven main countries worldwide, along with different studies and their specific duration. Three entries have multiple incidence recorded in this table, recorded at different time frames in the same study (Kuo et al, 2015).

Country	Study name or population	Incidence (cases per 1000 persons)	Duration of study
<b>USA</b>	Framingham study	0.84	1948-1980
	New England town study (Sudbury)	1.00	1964 (1 year)
	Atherosclerosis Risk Communities (ARIC) study	0.84	1987-2012
	Rochester Epidemiology Project	0.45 0.62	1977-1978 1995-1996
<b>UK</b>	Second and Third National Studies of Morbidity in General Practice	1.00	1971-1975
		1.40	1981-1982
	Clinical Practice Research Datalink (CPRD)	1.19-1.80	1990-1999
		1.36-1.77	1997-2012
		2.26	2012 (1 year)
Royal College of General Practitioners Weekly Returns Service (WRS)	1.12-1.35	1994-2007	
The Health Improvement Network (THIN)	2.68	2000-2007	
<b>Finland</b>	New England town study (Sudbury)	0.06	1974 (1 year)
<b>Italy</b>	Longitudinal Patient Database	0.95	2005-2009
<b>Czech Republic</b>	Ceske Budejovice and Cheb	0.41	2002-2003
<b>New Zealand</b>	Maori individuals	6.45	1962-1974
<b>Taiwan</b>	National Health Insurance Research Database (NHIRD)	2.74	2010 (1 year)
<b>Singapore</b>	The Singapore Chinese health cohort study	0.294-0.5	1993-1998

## 2.1.3 Pathogenesis of hyperuricaemia

### 2.1.3.1 Introduction

The last metabolite of the endogenous and dietary purine metabolism is uric acid. The latter is a weak acid (pH of 5.8) which is ionised to urate at physiological pH of 7.4. Urate crystals will start to deposit in joints when the serum uric acid (SUA) level increases above the normal threshold, as shown in Figure 2.2 below. The pathological threshold of hyperuricaemia is set at 6.8mg/dL (Dalbeth et al., 2019; McCarty and Hollander, 1961). Urate is mostly present as monosodium urate due to high sodium levels in the extracellular compartment. Monosodium urate crystals will form and precipitate when the urate levels exceed 380  $\mu\text{mol/L}$ . Urate is mostly produced in the liver, followed by the small intestines. Urate production depends on purine synthesis, salvage and degradation pathways, as depicted in Figure 2.3. Hyperuricaemia can be caused by an increased turnover of nucleic acid or an increase in ATP breakdown to AMP. The enzyme uricase mutated and absent in humans, thus preventing the degradation of uric acid to soluble allantoin and leaving urate close to its solubility limit (Richette & Bardin, 2010). Genetic polymorphisms in crucial urate transporters, such as human URAT1 transporter and the fructose SCL2A9 transporter, will decrease urate excretion (Tausche et al, 2009). Hyperuricaemia is caused in 90% of cases by the faulty renal excretion of uric acid and the rest of the 10% constitutes people who overproduce uric acid, as shown in Figure 2.3 (Richette & Bardin, 2010).

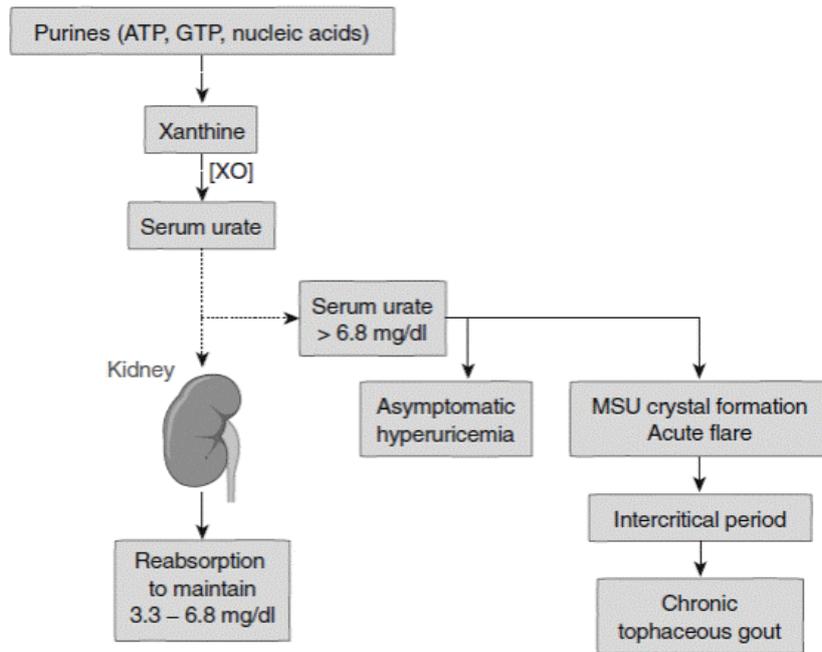


Figure 2-2 Flow diagram showing the conversion of purines to serum urate and ultimately to an excess leading to crystal formation and chronic tophaceous gout (Richette & Bardin, 2010).

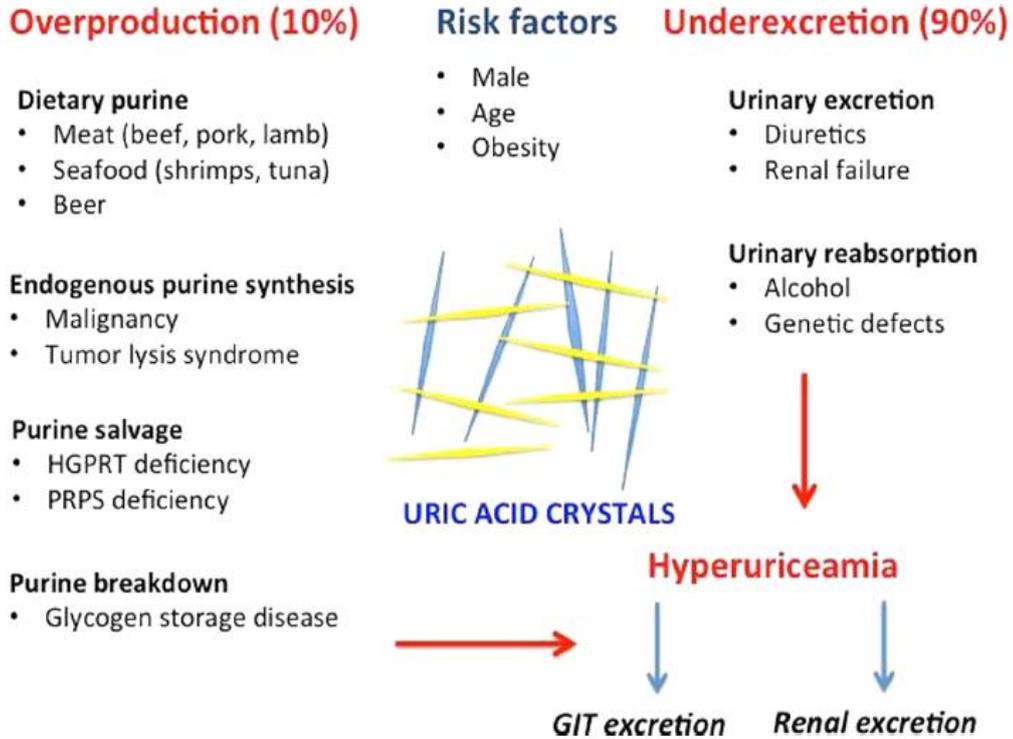


Figure 2-3 Pathogenesis of hyperuricaemia is depicted, along with the risk factors affecting this whole process (Ragab et al., 2017)

Monosodium urate crystals will initiate, increase and maintain a strong inflammatory response as shown in Figure 2.4. Patients with intra-articular depositions of MSU crystals develop an acute inflammatory response which leads to acute gout flares. The latter is caused by MSU crystals' interaction with resident macrophages, which in turn causes the formation and activation of the NLRP3 inflammasome (as shown in Figure 2.4) (Martinon et al., 2006). This process is enhanced by a microtubule-driven spatial co-localisation with mitochondria and involves  $\alpha$ -tubulin acetylation (Misawa et al., 2013). Caspase 1 (blue oval in Figure 2.4) is recruited by the inflammasome and processes pro-interleukin 1 $\beta$  into mature interleukin 1 $\beta$  (Martinon et al., 2006). Production of interleukin 1 $\beta$  also necessitates the presence of another signal besides the presence of MSU crystals, such as long-chain free fatty acids (Joosten et al., 2010).

This inflammatory response is enhanced by the activation of neutrophils and mast cells, causing the release of a host pro-inflammatory cytokine, chemokines, reactive oxygen species, prostaglandin E2 and lysosomal enzymes (Cronstein and Sunkureddi, 2013). The resolution phase of acute gouty inflammation is mediated by the introduction of anti-inflammatory cytokines, lipid mediators and aggregated neutrophil extracellular trap structures (bottom left of Figure 2.4) (Schauer et al., 2014). The tophi formed in gout is an organized chronic inflammatory granulomatous response to MSU crystals with the involvement of the innate and adaptive immune cells (Dalbeth et al., 2010). Aggregated neutrophil extracellular traps might also have a role in tophus formation by organizing MSU crystals in a non-inflammatory state and developing the crystals' core (Schauer et al., 2014). Infiltration of tophi in bones and joint then cause bone erosion and joint damage in gout (Dalbeth et al., 2009). Urate concentrations vary according to inheritance and environmental factors (Richette & Bardin, 2010). Deposition of crystals will be seen in joints, soft tissues and kidneys which ultimately manifests as arthritis, tophi, urate nephropathy and nephrolithiasis. Numerous factors contribute to deposition of crystals such as decreased body temperatures, variations in pH level and the degree of articular dehydration (Tausche et al, 2009).

Factors affecting the solubility of uric acid in joints are the synovial fluid pH, synovial components such as proteoglycans and collagen, water concentration and electrolytes level.

A level of SUA is achieved by the balance between its production and its excretion by the kidneys and gastrointestinal tract (GIT). SUA production comprises of diet purine intake and endogenous production by cellular turnover. 10% of gout cases can be attributed to increased SUA, while 90% of cases are caused by its renal under-excretion (Mandal & Mount, 2015). Serum uric acid (SUA) level is affected by age and gender. SUA is normally low in children, starts to increase and reaches normal levels during puberty. Men usually show higher SUA levels compared to women, but post-menopausal women can reach men's SUA level too. This shows the reason behind gout being a disease for middle-aged, elderly patients and post-menopausal women. Gout occurring in children and young adults are due to rare inborn errors in their purine metabolism. Enzymatic defects in the latter cause an increase in SUA levels, followed by consequent UA crystals production in kidneys and joints (Kamei et al, 2014).

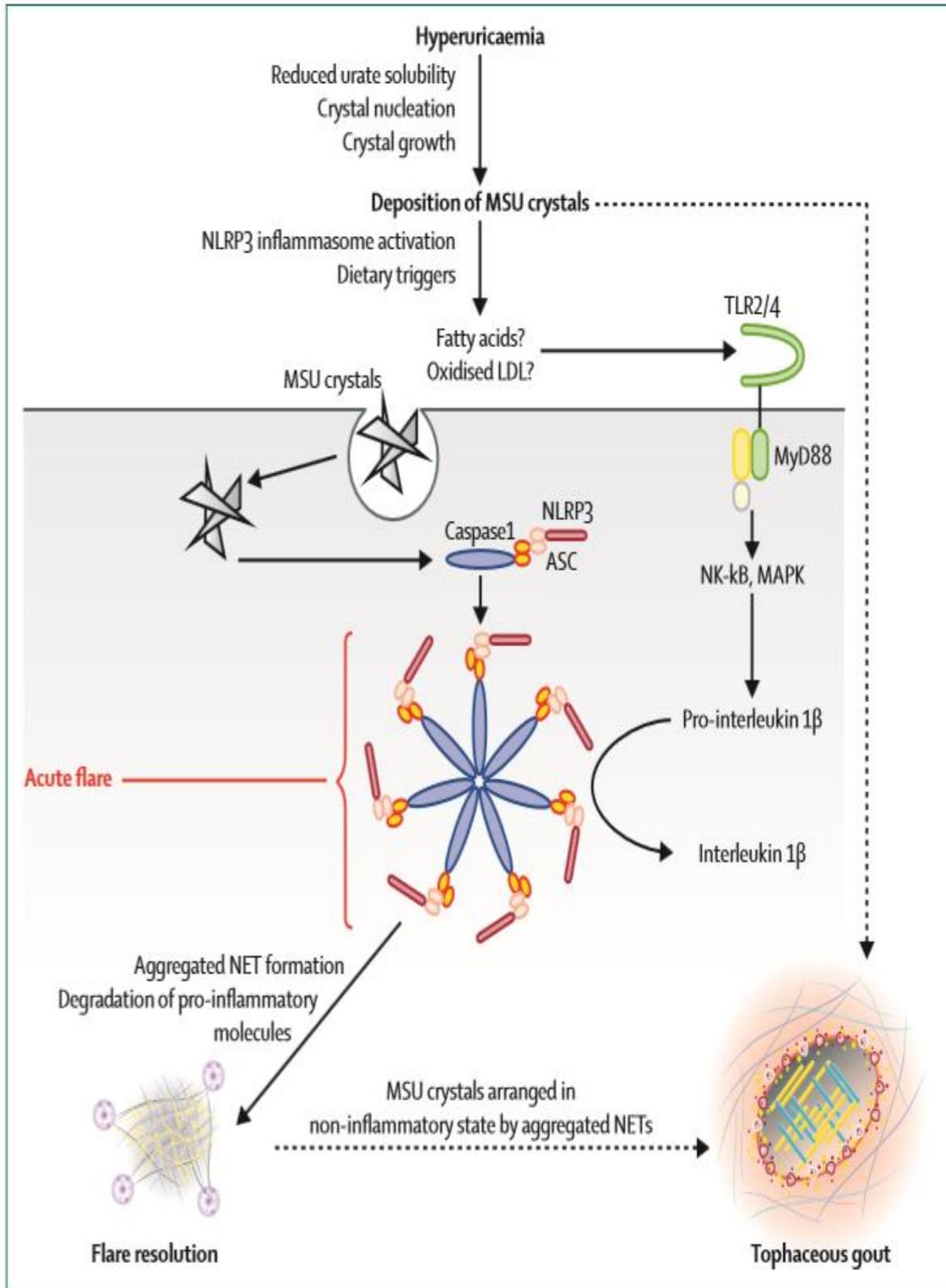


Figure 2-4 Pathway for the progression of hyperuricaemia to tophaceous gout, along with its depicted inflammation response (Dalbeth et al., 2016).

### 2.1.3.2 Overproduction of uric acid

Overproduction of UA in the purine metabolism is usually caused by a deficiency in enzymes. Lesch-Nyhan syndrome is caused by a lack of a specific enzyme in UA metabolism, known as hypoxanthine–guanine phosphoribosyltransferase. The latter is an inborn error of metabolism and a genetic X-linked recessive disorder, which can vary in severity based on specific mutations. Neurological abnormalities are seen in this disease, such as chorea, self-mutilation, dystonia, cognitive dysfunction, articular manifestations and compulsive injurious behavior. In absence of any treatment, patients may develop tophi, renal stones and renal failure (Torres & Puig, 2007). In youngsters, the enzyme phosphoribosyl pyrophosphate synthetase is known to be superactive and thus causes gout. The latter is an X-linked dominant inherited disorder with two clinical forms, namely, a severe early onset form in children and an early adult onset form. Neurological abnormalities are seen again, with hypotonia, sensorineural hearing loss and severe ataxia. A mild form of this disease presents itself as arthritis and UA renal stones. The two enzymatic disorders listed above make up less than 10% of UA overproduction cases (Reginato & Olsen, 2007). Overproduction of UA is often related to a patient's diet or dietary purine, as shown in Figure 2.3. The easiest way of increasing UA precursors in the body is the ingestion of purine rich food, be it cooked or processed, from animal or seafood origin. Another important risk factor for gout is alcohol consumption, especially beer intake. Detailed explanation for specific dietary intakes linked to a higher risk of gout will be given in the following section (Kanbara & Seyama, 2011; Towiwat & Li, 2015).

### **2.1.3.3 Endogenous urate production**

An increase in endogenous UA production is seen in cases of accelerated cellular turnover, such as in malignancies, inflammatory and hematological diseases, as depicted in Figure 2.4. Chemotherapy and tissue damage are other instances of increased purine production. UA production is enhanced by an increase in weight, obesity and leptin, which all contribute in elevating the risk of hyperuricaemia. Thus patients can exercise and lose weight in order to reduce SUA levels and gout risk (Bedir et al, 2003; Dessein et al, 2000; Emmerson, 1998; Mahmoud et al, 1998).

### **2.1.3.4 Declined uric acid excretion**

The kidneys are responsible for the bulk of urate excretion, while the GIT excretes one third of the remaining urate in our bodies. In cases where the transporter ABCG2 has a lower secretory function, the amount of UA excreted through the GIT decreases, thus leading to a rise in body SUA levels and a boost in urate renal elimination (Ichida et al., 2012; Mandal & Mount, 2015). Specific membrane transporters are needed to carry the insoluble UA crystals through cell membranes. The main type of transporters are, the urate transporter/channel (mainly URAT1) and the organic anion transporters (OAT1 and OAT3) (Enomoto & Endou, 2005; Mandal & Mount, 2015). URAT1 activity can be decreased by using uricosuric drugs, such as probenecid, benzbromarone and sulfinpyrazone, which in turn causes a decrease in UA reabsorption in the proximal tubules. Other drugs such as pyrazinamide, lactate and nicotinate, work by increasing urate reabsorption via URAT1 and increasing glomerular filtration and tubular reabsorption of UA, thus preventing any loss of UA in urine ( Cho et al., 2015; Tan et al., 2016). A dose-based response is also seen on URAT1's activity with other substances, where it can be inhibited or potentiated. Aspirin is one example, where a high dose has a uricosuric and cis-inhibition effect by blocking URAT1's action. The opposite occurs with a low aspirin dose having an anti-uricosuric and trans-stimulation effect on URAT1 (Tan et al., 2016). Some autosomal dominant disorders are also associated with a reduction in UA renal excretion. One example is mutations in the gene uromodulin, which usually expressed in the loop of henle and regulates water permeability. Mutations in the

latter results in a reduction in fractional excretion of UA and a following increase in SUA level (Han et al., 2013).

#### **2.1.3.5 Genes governing uric acid regulation**

The SLC22A12 gene, encoding URAT1 transporters on the renal tubules, is involved in UA excretion regulation. SLC2A9 encodes for a specific transporter protein present on renal tubules and regulates UA excretion. Both genes can have polymorphisms and reduce the level of UA excretion, thus increasing SUA levels. Other genes identified are the ABCG2 transporter gene in the kidney and GIT, two membrane transporters in the kidney (SLC17A1 & SLC17A3), SLC22A11, Carmil (LRRC16A), the glucokinase regulatory protein (GCKR) and PDZ domain containing 1 (PDZK1) genes (Kolz et al, 2009; Phipps-Green et al, 2016).

## 2.1.4 Clinical manifestation of gout

### 2.1.4.1 Introduction

Gout can manifest itself at varying degrees and stages. It is imperative to understand the different types of gout for proper clinical diagnosis and subsequent treatment. It usually starts with asymptomatic hyperuricaemia, followed by acute gout, an intercritical stage and chronic gout. Figure 2.5 below shows an overall picture of the parallel clinical manifestation stages of gout, along with the disease progression and risk factors linked to the disease's deterioration. On the appearance of hyperuricaemia, deposits of MSU crystal manifest in joints and the consequent gout flares, along with chronic and tophaceous gout are seen in patients. The main determining stages for this disease are the start of hyperuricaemia and MSU crystal deposits, which are unfortunately asymptomatic and can be exacerbated by numerous risk factors. The key element in gout treatment and management is the proper and timely detection of these two aforementioned stages. The different stages of disease progression are broken down and explained, followed by detection and diagnosis in following sections (Campion et al, 1987; Dalbeth et al, 2019).

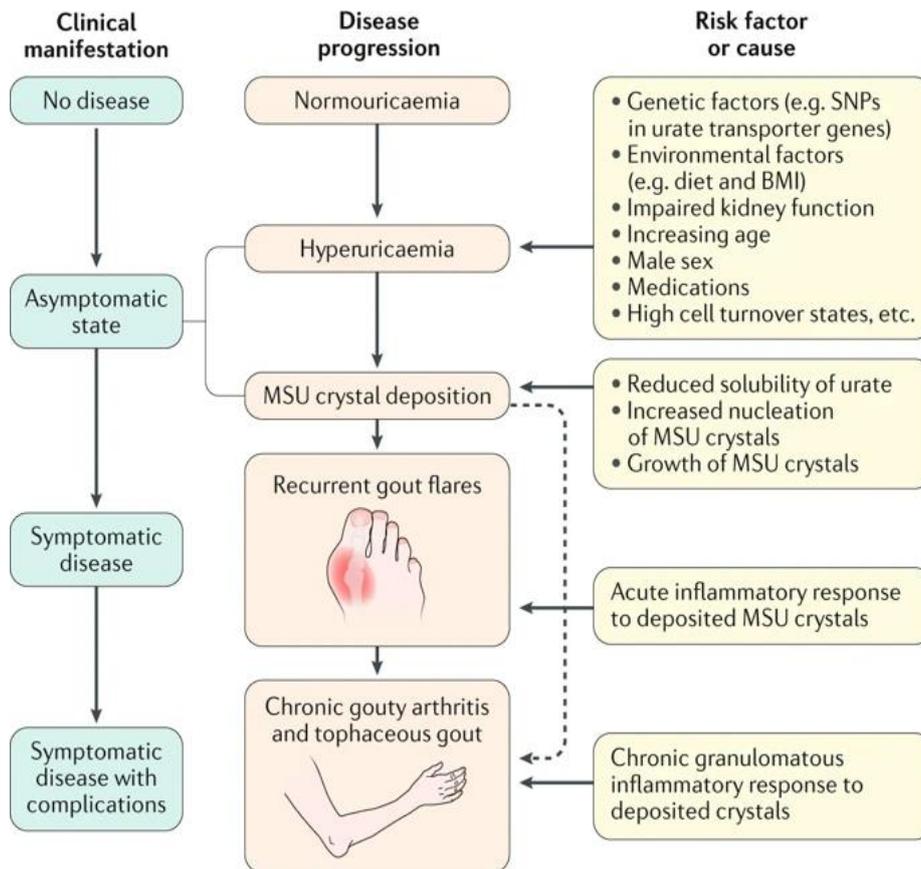


Figure 2-5 Clinical manifestations of gout are shown, along with the different stages of the disease and the risk factors involved. Normouricaemia progresses to hyperuricaemia, followed by MSU crystal deposition, which in turn leads to gout flares, chronic gouty arthritis and tophaceous gout (Dalbeth et al, 2019).

## **2.1.4.2 Types of gout**

### **2.1.4.2.1 Asymptomatic hyperuricaemia**

Asymptomatic hyperuricaemia is the first disease stage to occur, out of the four different stages of gout. This stage is usually suddenly discovered while checking a patient's SUA level as no other signs or symptoms are present. In some cases, hyperuricaemia alone can cause a sudden, acute gouty attack.

### **2.1.4.2.2 Acute gouty attack**

An acute attack is often monoarthritic and peaks after a few hours and goes to inflame joints with numerous signs of inflammation (Roddy, 2011). A maximum level of inflammation will be reached within 4 to 12 hours after the onset of the gout attack and this response will decrease and fade, even in the absence of treatment. The Ministry of Health in Malaysia (MOH) gathers data regularly in order to improve the management of gout in patients. Studies across hospitals show that gout attacks last an average of 2-3 weeks, with a complete resolution of inflammation. 60% of patients experience a second attack in the first year and 78% have a second attack in the second year. A small number of patients do not have any recurrence in around 10 years, amounting to around 7% (Malaysia Ministry of Health, 2008).

The first attack of gouty arthritis usually occurs in the small joints of the lower limb, particularly the metatarsophalangeal (MTP) joint, also known as podagra, shown in Figure 2.6 A. The affected joints will experience pain, hotness, tenderness, swelling, redness and loss of function. The knees and ankles are categorized as larger joints and usually do not show any skin signs and symptoms, but swelling and pain in those locations are very intense. Other usually affected areas are the tarsal and metatarsal joints, wrists, metacarpophalangeal joints (MCP) and the interphalangeal joints in the hands (shown in Figure 2.6 B), while the rarely involved areas are the hip, shoulder and spinal cord. Soft tissue inflammation is also possible and includes olecranon bursitis and Achilles tendonitis (Roddy, 2011). Untreated gout and postmenopausal women commonly experience the

involvement of more than one joint simultaneously. In cases where constitutional symptoms are present, such as malaise, fever and headache, it is treated and managed as septic arthritis, until proven otherwise (Perez-Ruiz et al, 2014).

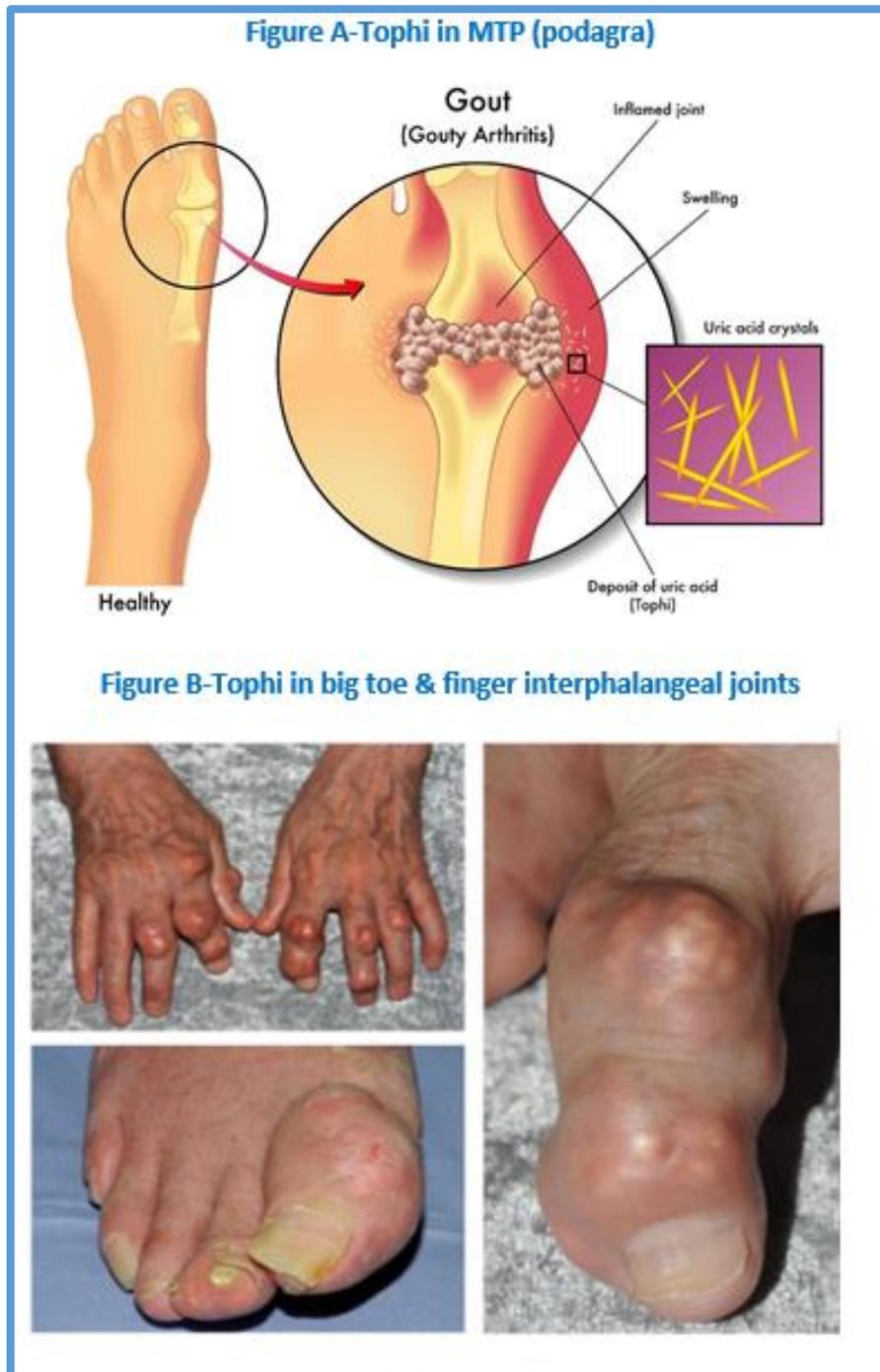


Figure 2-6 Figure A depicts the deposition of uric acid crystal in the metatarsophalangeal (MTP) joint, also called podagra. Figure B shows the tophi formed in the right big toe and the finger interphalangeal joints in a gout patient (Roddy, 2011).

#### **2.1.4.2.3 Intercritical period**

An acute gout attack will usually settle down and enter a remission phase within hours or days after use of drugs, such as NSAIDs or colchicine. This stage usually does not show any symptoms, but can be interrupted by new attacks in the absence of proper treatment and management of hyperuricaemia. The intercritical period can be lengthened successfully after the first attack but in the opposite case, attacks will become more frequent and severe (Pascual et al, 1999).

#### **2.1.4.2.4 Chronic tophaceous gout**

Untreated cases leads to joints' destruction, along with formation of palpable tophi. Accumulated uric acid will crystalize and accumulate in large amounts in joints to form a tophus in chronic untreated gout. Tophi appear like white chalky deposit macroscopically and may cause bony erosions when it reaches the bone. These tophi can be located even in the joints of ears and subcutaneous tissue and skin. The apparition of tophi themselves is a sign of uncontrolled and chronic gout. Tophi need to be differentiated from other nodules seen in rheumatoid arthritis, Bouchard's, lipomas and osteoarthritic Heberden's nodes. (Chhana & Dalbeth, 2015)

Tophaceous gout is more likely to appear in patients with polyarticular presentation, SUA levels more than 9 mg/d and in younger patients at 40 years or less. More severe cases of gout may show presence of chronic joint symptoms linked to joint damage or with apparent tophi. The clinical diagnosis of gout depends on the confirmation by a patient's medical history, a physical examination, laboratory tests and other tests. (Zhang et al, 2006)

#### 2.1.4.2.5 Gout presentation in elderly patients

The average age of onset for gout in elderly patients is more than 65 years, compared to a disease peak around mid forties for the younger population. A larger number of men are affected compared to women in the younger population, but this distribution changes for the elderly. Older men and women are equally affected by gout but women above 80 years old have a higher frequency of the disease. Gout presents itself as acute monoarthritis in younger patients, with the lower extremities affected and 60% of those cases affecting podagra. However, with age, this changes to more polyarticular involvement, with the upper and smaller extremities' involvement, specially the fingers. Polyarticular gout mimics rheumatoid arthritis where symmetrical small joints are involved and tophi show up on extensor tendon surfaces. Differentiating these two diseases gets more confusing due to low rheumatoid factors in both. Thus, synovial fluid analysis is imperative to remove any any confusion. Older patients develop tophi faster during the early stages of gout and in atypical locations, whereas tophi only show up after years of attacks in younger patients, especially in the elbows. Moreover, a higher association of gout to renal diseases, renal insufficiency and diuretic use is observed in the elderly. Youngsters show a greater association to obesity, hyperlipidaemia, hypertension and heavy alcohol use (Malaysia Ministry of Health, 2008).

## 2.1.5 Diagnosis of gout

### 2.1.5.1 Clinical diagnosis

Clinical diagnosis of gout is a key factor in treatment and prevention. This comprises of correctly identifying the type of gout a patient has via the different steps and stages of diagnosis methods available. Clinical diagnosis includes the physical examination by doctors and other more specific and sensitive methods like microscopic diagnosis of crystals, radiography and blood tests (Malik et al, 2009). In certain cases of atypical presentation of gout, for example in unconventional joint distribution and when multiple joints are affected, monosodium urate crystals (MSU) crystal identification is crucial for correct differentiation. The standard resulting diagnosis for gout is usually elevated SUA levels and common joints involved, such as podagra (Zhang et al, 2006). Formation of tophi, present more in late gout stages, are definitive gout indicators, especially when other types of nodules have been ruled out of the diagnosis (Atdjian & Fernandez-Madrid, 1981).

The European League Against Rheumatism (EULAR) put together an updated set of 8 recommendations for diagnosis of gout, which was honed and improved after several decades, shown in Figure 2.7 (Richette et al, 2020). EULAR states that the first key factor in identifying gout is to show the presence of MSU crystals in the synovial fluid or tophus aspirates and that hyperuricaemia alone isn't enough for diagnosis. In cases of acute arthritis or undiagnosed inflammatory arthritis, gout should not be ruled out as the cause and further investigations to prove the latter must be done. Imaging techniques are recommended in the absence of any other MSU crystal identification techniques as shown in recommendation five and six in Figure 2.6. The last two recommendations emphasizes on the importance of thoroughly investigating the numerous risk factors and comorbidities linked to gout. Physicians can use these recommendations to guide them for a proper and efficient gout diagnosis in patients (Richette et al, 2020).

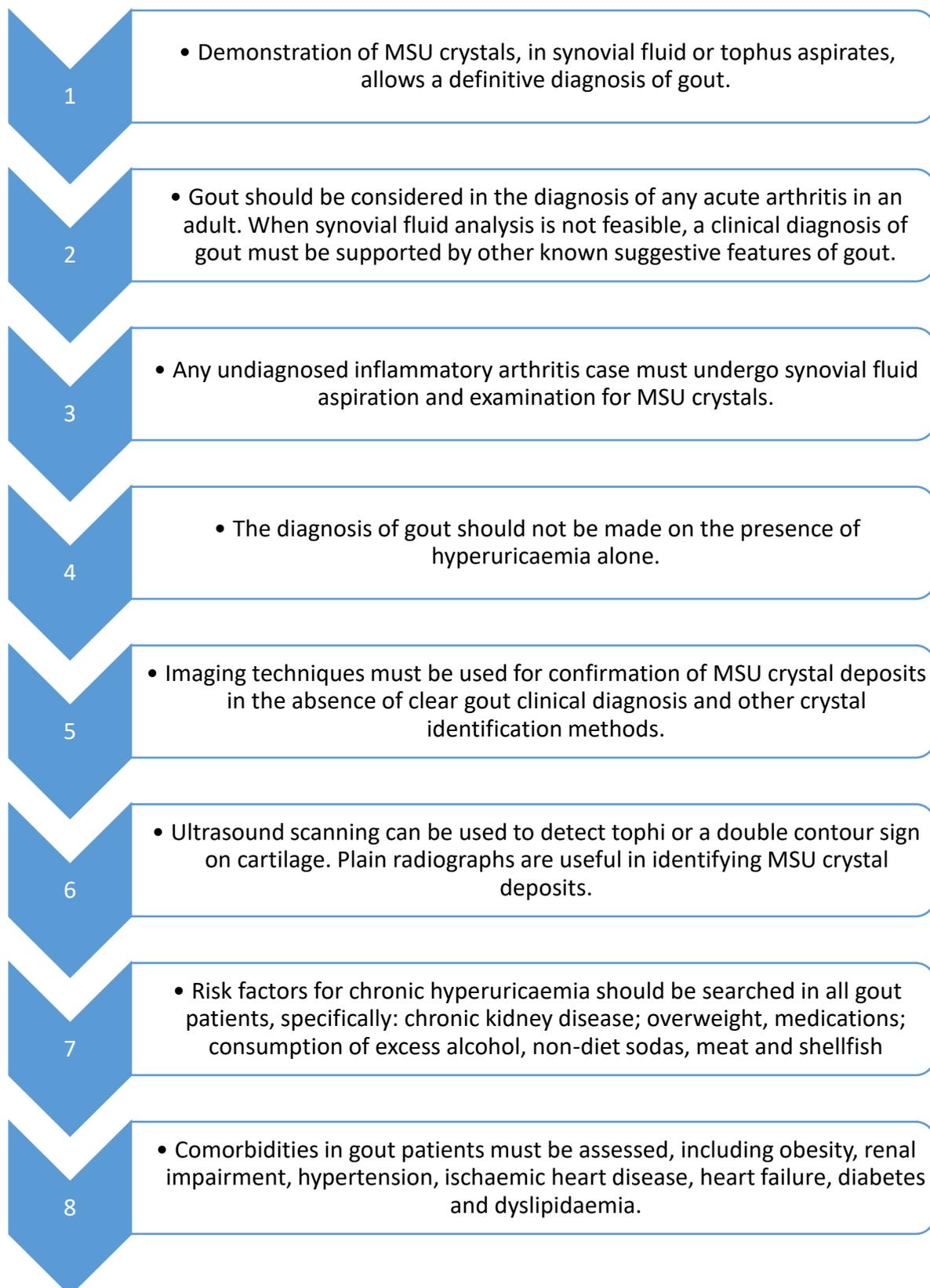


Figure 2-7 Flowchart depicting the right main recommendations for diagnosis of gout.

Adapted from EULAR 2019 recommendations (Dincer et al, 2002; Richette et al, 2020).

## 2.1.5.2 Types of diagnosis

### 2.1.5.2.1 Laboratory diagnosis

Non-rheumatologists have a common misconception of diagnosing gout based on hyperuricaemia alone. However the latter is asymptomatic and does not give a definitive gout diagnosis. Only 0.09% patients with SUA levels in between 7-7.9 mg/dL will develop gout yearly, further supporting this claim. Patients with SUA levels of 8-8.9 mg/dL may have a 0.4% chance of developing gout, while those with SUA level above 9 mg/dL have a 0.5% chance of probably getting gout (Dincer et al, 2002). Even though hyperuricaemia is a hallmark of gout, SUA levels are known to drop to normal levels during a gouty attack, hence further showing how hyperuricaemia is a weak marker for diagnosis (Bădulescu et al, 2014).

The gold standard for gout diagnosis is MSU crystal identification in synovial fluid aspirate by using the polarized light microscopy method, with better yields by using a compensator. A standard light microscope can also be used to single out MSU crystals and differentiate them from calcium pyrophosphate dehydrate (CPPD) crystals, as shown in Figure 2.8 below. MSU crystals are conveniently found in all different stages of gout, be it in gout attacks, the intercritical period or in the chronic stage (Underwood, 2006). Baseline investigations commonly done in Malaysian hospitals are a full differential blood count, serum creatinine level, blood glucose level, serum urate level, fasting lipid profile, urinalysis and in certain cases a 24 hour urinary excretion investigation (Malaysia Ministry of Health, 2008). Samples need to be analysed within 6 hours or within 24 hours when kept refrigerated, in order to prevent cellular dissolution and crystal disappearance (Pascual et al, 1989). Synovial fluid analysis also includes leukocyte count, culture, chemistry and sensitivity (Strasinger & Di Lorenzo, 2014). The uric acid level analysis in urine helps to assess the etiology of hyperuricaemia. Uric acid levels more than 800 mg/24 hour shows an increased UA production and a higher excretion level of UA (Kramer & Curhan, 2002).



Figure 2-8 Detection of MSU crystals in the synovial fluid of a gout patient, by using light microscopy (polarized light) and a first-order compensating filter. Black arrows show the presence of large, bright needle-shaped MSU crystals, while white arrows show presence of brick-shaped calcium pyrophosphate dihydrate crystals. (Dalbeth et al, 2019)

#### 2.1.5.2.2 Imaging for diagnosis and assessment (G5)

Ultrasonography (US) and Dual-Energy CT (DECT) are used to detect MSU crystals, evaluate symptoms and check atypical clinical manifestations (Filippucci et al, 2013; Reuss-Borst et al, 2014). Ultrasonography shows a double-contour sign, indicating MSU crystal deposition in hyaline cartilage, and tophi presence as shown in Figure 2.9 below (Taylor et al, 2015). US can detect crystal deposition, joint effusion, synovitis, differentiate between active and inactive synovitis, study cartilage, tendons, bony erosions and monitor disease evolution (shown in Figure 2.9 below)(Nestorova & Fodor, 2015).

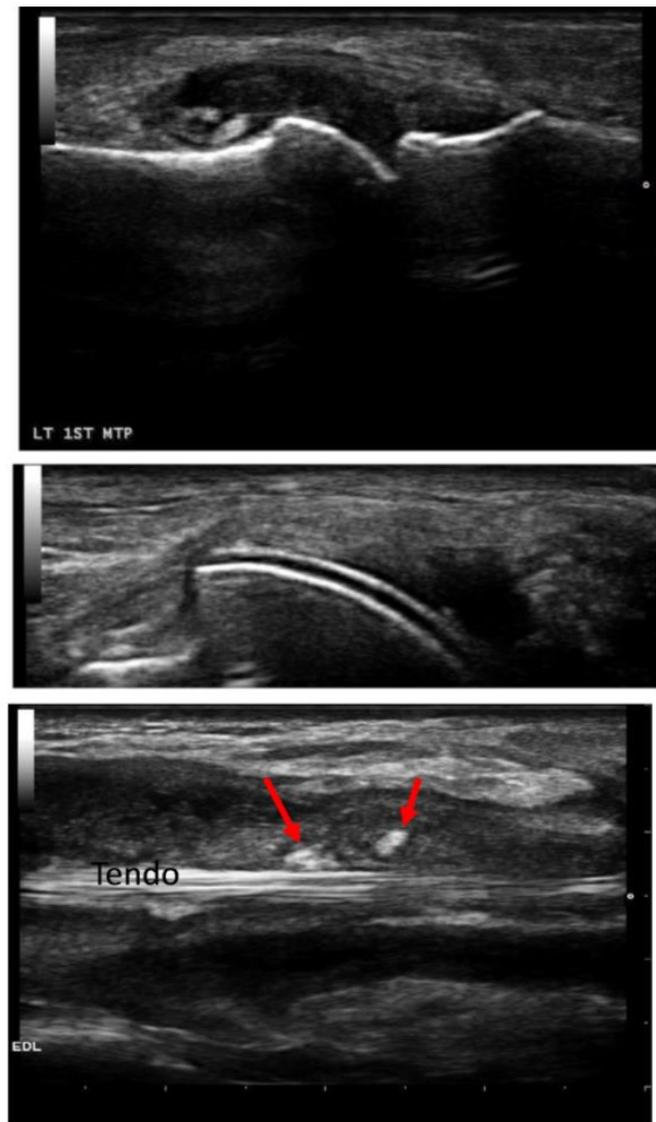


Figure 2-9 The three examples of ultrasonography shown are (a) Intra-articular tophus in MTP joint, (b) double contour sign in the hyaline cartilage and (c) synovial effusion/hypertrophy and crystal formation (indicated by 2 red arrows) in the tendon (Ragab et al, 2017).

DECT, a new imaging technique, allows the differentiation of deposits based on their different X-ray spectra, as shown in Figure 2.10 below. Here, attenuation of tissues depends on their atomic number, density and photon beam energy. It is better than all other available imaging techniques in its ability to identify urate deposition in imaged areas (Omoumi et al, 2015b). DECT provides a rapid, non-invasive method to visualize MSU crystals, change in soft tissue and early onset erosions. DECT has a high accuracy for MSU

crystal detection in joints, ligaments, tendons and soft tissues, with an added high specificity of sub-clinical gout identification (McQueen et al, 2011). However DECT cannot detect crystal deposit on cartilage surfaces, a feat which US can detect (Huppertz et al, 2014). Moreover, US is more sensitive and has a high positive predictive value for early disease diagnosis, when the crystal deposition level is low (Wang et al, 2018). Both methods are used to monitor the MSU deposit level and the response to urate lowering therapy. Plain radiographs cannot show erosions, bone and joint changes until late stages of gout (Taylor et al, 2015).

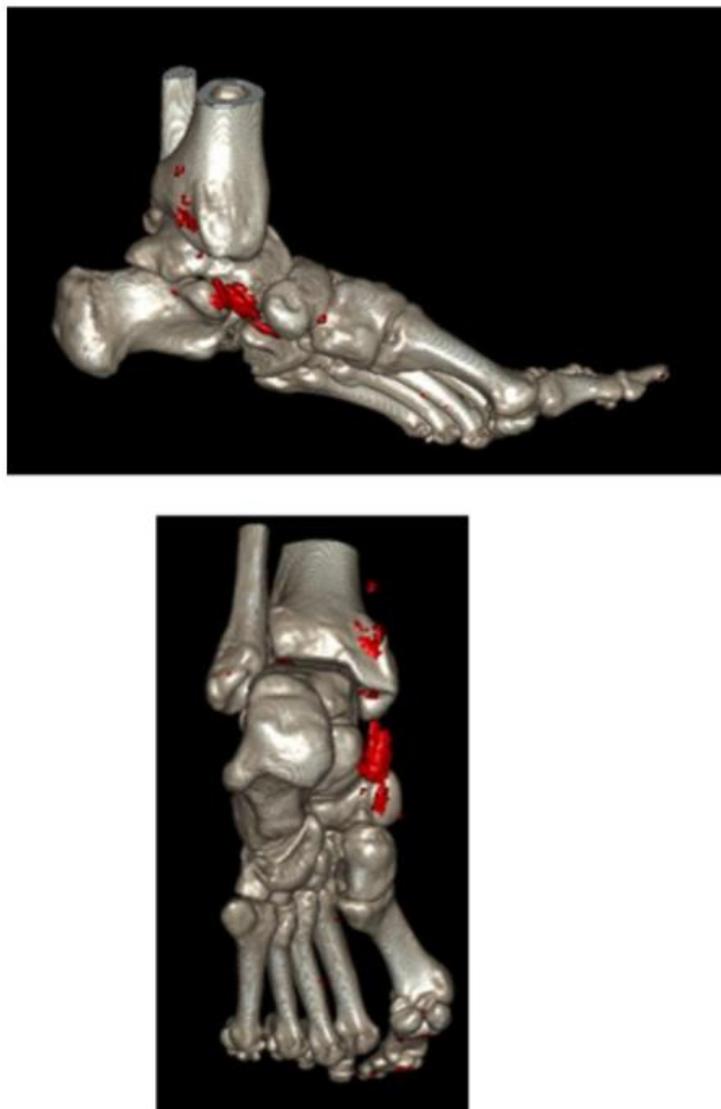


Figure 2-10 DECT showing two view points for MSU deposits (red deposits) in the tendon of a gout patient. (Ragab et al, 2017)

MRI and CT can mostly identify tophi presence and are used to assess soft tissue and bone involvement in gout and its comorbidities. Conventional CT (CCT) has good resolution and high contrast and is the best technique for identifying crystal arthropathies (Omoumi et al, 2016). However, CCT is not appropriate for use in acute gout diagnosis as it cannot detect synovitis, inflammation, tenosynovitis and osteitis. In chronic gout, CCT detects erosions better than Magnetic Resonance Imaging (MRI) and it detects tophi better than both US and MRI (Dalbeth et al, 2009; Gerster et al, 2002; McQueen et al, 2011). CCT can monitor disease burden and response seen to specific therapies, but with an added disadvantage of radiation exposure (Dalbeth et al, 2007; Omoumi et al, 2015a; Omoumi et al, 2015c). MRI shows numerous features of arthritis, such as synovial thickening, effusion, erosion, nonspecific inflammation and tophi presence. However, it has a limited use due to its high cost and limited availability (Chowalloor et al, 2014).

## **2.1.6 Risk factors for gout**

### **2.1.6.1 Hyperuricaemia**

The main hallmark of MSU formation is hyperuricaemia, which can in turn be heightened by several factors and ultimately increase the risk of getting gout. These factors are needed in conjunction with hyperuricaemia in order to induce MSU deposition in joints and these are called risk factors. Risk factors include hyperuricaemia itself, followed by genetics, age, gender, diet, alcohol consumption and lead exposure, which are further detailed below. Hyperuricaemia is the most important risk factor for gout. A clear dose response relationship was proven between the serum uric acid (SUA) levels and the incidence of gout. The Normative Aging study showed that hyperuricaemia alone rendered other risk factors null in a Cox regression model (Campion et al, 1987).

A study in Taiwan, showed a high 5-year cumulative incidence of 18.83% for new gout cases in a cohort of asymptomatic hyperuricemic patients. A 32-fold increase in gout risk was found in hyperuricemic patients, compared to people with normal SUA levels in

Germany. Another 5-year study in France showed an increase in the prevalence of gout from 1.3%, 3.2% to 17.6% in patients with SUA levels of <6.0 mg/dl, 6.0–7.9 mg/dl and >8.0 mg/dl respectively. This linear relationship between hyperuricaemia and risk of gout was also seen in the Framingham study (Kuo et al, 2015)

#### **2.1.6.2 Age, gender and socioeconomic factors**

Age and gender have always been the first identified risk factors to any type of disease. Epidemiological studies have confirmed male patients and older age as risk factors for both hyperuricaemia and gout (Arromdee et al, 2002; Mikuls & Saag, 2006). Recent studies found a two-fold increase in incidence of gout in men and a positive linear relationship between the incidence of gout and age in both genders (Soriano et al, 2011). Women are known to have a lower prevalence compared to men at all ages, but this increases exponentially after menopause, due to oestrogen's uricosuric action. In a Nurses' Health Study, the risk of gout was increased by menopause, but was decreased by using hormone replacement therapy. Other socioeconomic factors linked to a higher gout risk are urban locations and less-privileged and wealthy areas. Two European studies and one Taiwanese study also confirmed this lower gout risk in rural residents. A recent CPRD study challenged the notion that gout is usually linked to wealth and debauchery and proved higher levels of gout in Wales and Northeast England, two lesser-privileged areas (Kuo et al., 2015). Another factor linked to the gout risk is the type of occupations of patients, where non-manual jobs had a higher risk, compared to professionals (Kuo et al., 2015).

### 2.1.6.3 Genetics

Genetics also play a role in the development of gout, where a distinct prevalence variation of gout in different populations and countries point towards a racial and genetic difference. This is further demonstrated by the varying prevalence of gout in different populations, which can range from 2.6% to 47.2%. SUA levels also attest to this, where levels are similar in North American Caucasians and UK population, but has higher prevalence in African Americans compared to Caucasian Americans. Other populations with high gout prevalence are Pacific Islanders (including Maoris, Cook Islanders and Micronesians) and aboriginal Taiwanese. A Taiwanese nationwide study showed familial aggregation of gout and the importance of genetics, with a reported heritability of 35.1% and 17.0% in men and women. The interaction between genes and the environment was demonstrated in a study with a nine-fold increase in gout risk in migrants in New Zealand, compared to locals (Kuo et al., 2015).

Rare monogenic disorders were found to cause hyperuricaemia and premature gout, such as a deficiency of hypoxanthine-guanine phosphoribosyl transferase called Lesch-Nyan syndrome, over-activity of phosphoribosyl-1-pyrophosphate synthetase, and hereditary renal diseases such as familial hyperuricaemia nephropathy and medullary cystic kidney disease. Polymorphism of a gene responsible for urate renal clearance, by the urate transporter 1 (URAT-1), was identified as a cause of hereditary, common gout. Limited information and case studies are available on the familial aggregation of gout and needs further research (Choi et al., 2010). A 2013 Genome Wide Association Study (GWAS) on Europeans identified and replicated 28 loci linked to SUA levels. Only 7% out of the loci were known associations for the variance in SUA levels, emphasizing on how much is left unresolved in the heritability of hyperuricaemia by scientists. 'Missing heritability' is often seen in complex disorders and is based on those unidentified risk factors, rare penetrating risk variants and unknown non additive interactions between other genes and environmental factors (Kuo et al., 2015).

#### **2.1.6.4 Alcohol consumption**

Alcohol intake and gout has been linked by metabolic studies showing that alcohol has a hyperuricaemic effect and eventually contributes to the manifestation of the disease. This relationship can be seen clearly in men who drink an excessive amount of beer, a beverage known to have a higher purine content compared to other alcoholic beverages. 50 g of alcohol a day was enough to increase the gout risk by 2.5 times. Moreover, higher levels of alcohol consumption were linked with hypertension, diuretic use and an increased meat consumption, which in turn, further increase the risk of gout (Choi et al., 2004).

#### **2.1.6.5 Dietary factors**

Dietary factors' restriction and maintenance is one of the key factors to prevent further worsening in the life of gout patients. Gout patients are usually advised by doctors to limit their daily intake purine-rich foods such as seafood, high purine content vegetables, meat and proteins. However, the extent to which these dietary products must be limited or included in a gout patient's diet is not well defined. By unnecessarily limiting or eradicating some of those crucial dietary products, the patient's health may suffer in other ways unrelated to gout. It is vital to define the specific effect of all those dietary factors in gout patients. The confounding effect of protein-rich diets on gout is a prime example, where even in the presence of high purine levels, excretion of uric acid is promoted and thus ultimately decreases the serum uric acid level (Choi et al, 2004).

Choi et al. concluded a direct relationship between higher seafood and meat consumption and an inverse relationship with increased protein and purine-rich vegetables with gout. In terms of numbers, an additional daily serving of meat causes a 21% increase in risk of gout, whereas an additional weekly serving of seafood caused an increase of 7%. This increase in gout risk was seen to be higher in gout patients compared to healthy individuals, as gout patients already have a compromised renal urate clearance level where the uric acid level cannot be regulated properly. Dairy products, especially low-fat products, demonstrated a strong inverse relation with gout incidence. All these associations were independent of any

other dietary factors or other gout risk factors such as age, alcohol consumption, hypertension or chronic renal failure (Choi et al, 2004).

A higher intake of vegetable or animal **protein** is not linked to an increased incidence of gout. Choi et al. showed that vegetable protein can have a protective effect in gout, though this effect is smaller when compared to dairy proteins. Moreover, uric acid excretion can be raised in the presence of high-protein diets, thus contributing to a reduction in uric acid level. One study increased the intake of protein in patients and this caused a decrease in the number of recurring gout attacks. As such, intake of protein does not increase, but decrease gout incidence and cannot be directly linked to their purine levels (Choi et al, 2004).

## **2.1.7 Treatment of gout**

### **2.1.7.1 Introduction**

Several drugs are available for treatment of acute gout attacks, such as colchicine, non-steroidal anti-inflammatory drugs (NSAIDs), corticosteroids and intramuscular adrenal corticotrophin hormone (ACTH). The aim of long-term management is to reduce SUA levels below the saturation point of crystals, in order to prevent further MSU crystal accumulation and cause existing crystals to dissolve. The general principles of gout management are seen in Figure 2.11 below, based on the 2012 American College of Rheumatology (ACR) guidelines and the 2016 European League Against Rheumatism (EULAR) guidelines (Dalbeth et al., 2019).

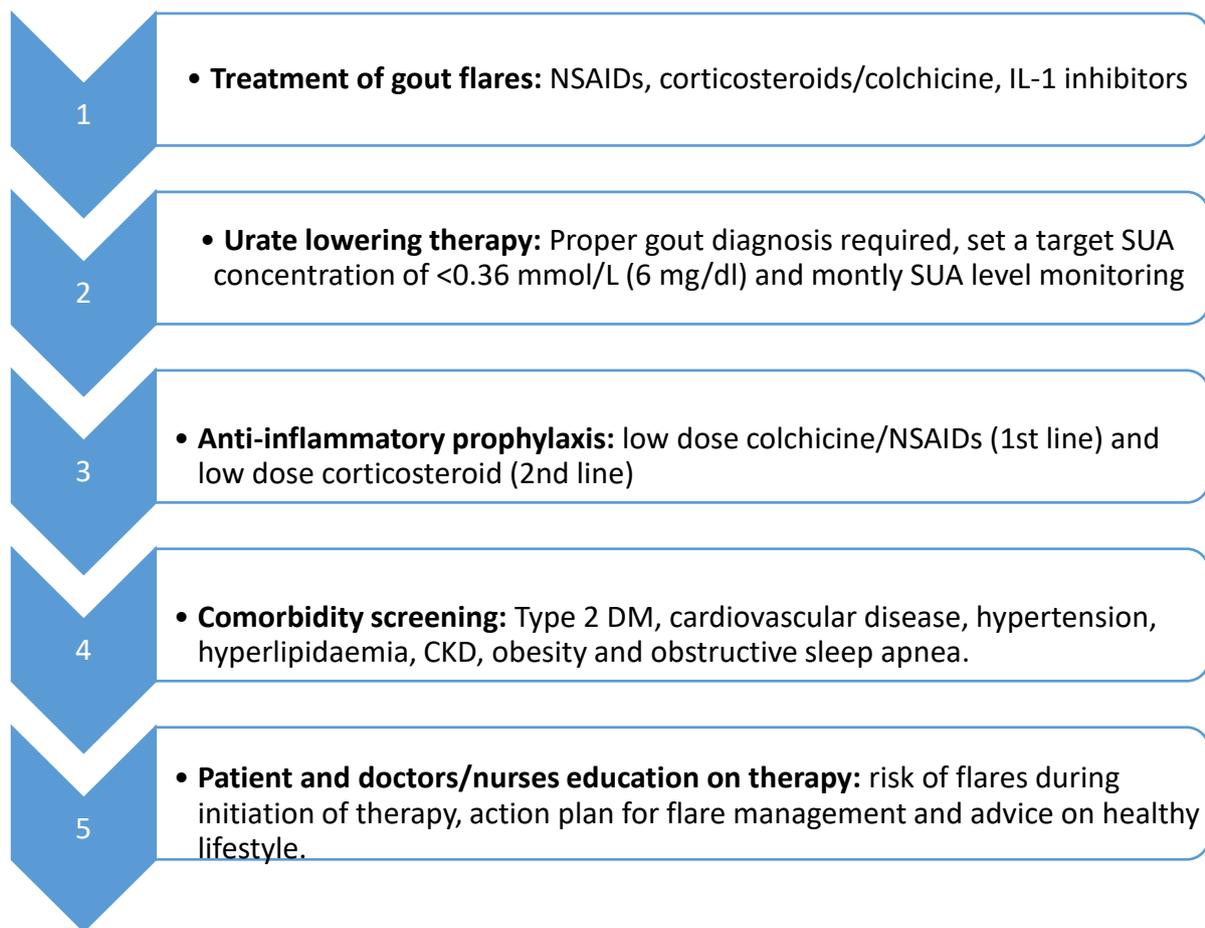


Figure 2-11 General principles of gout management with pharmacological and non-pharmacological guidelines based on EULAR recommendations (Zhang et al., 2006).

Urate lowering therapies (ULT) are divided into two categories, namely xanthine oxidase inhibitors and uricosuric drugs. Allopurinol and febuxostat are mainly used as xanthine oxidase inhibitors, while probenecid, benzbromarone and sulphinyprazole are three commonly used uricosuric agents. Initiation of ULT causes frequent acute attacks, due to the dissolving of MSU crystals in cartilage which is then redirected towards joint cavities. The latter can be decreased in 85% of cases by designing a prophylactic therapy with initial regular, low doses of colchicine or NSAIDs. However, with all the possible drug combinations and therapies available, less than 25% of gout patients become free of gout attacks (Dalbeth et al., 2019). This suggests that using drugs for gout treatment is not enough and that other factors play a big role in gout management. Factors to be limited during gout treatment are the intake of alcohol, beverages with fructose, seafood intake, etc, as previously mentioned. Other measures to take include, a careful monitoring of body

weight, adequate exercise, liberal water intake and proper treatment of any other comorbidities, such as hypertension, DM and dyslipidaemia. EULAR proposed a set of 12 recommendations for treatment of gout, which includes pharmacologic and non-pharmacologic means, as summarized in Figure 2.12 (Zhang et al., 2006). Patient education, lifestyle advice and management of comorbidities were the first recommendations set. Allopurinol was labelled as the best first-line urate-lowering treatment, followed by all other alternative uricosuric. Effective acute treatment consisted of colchicine/NSAID use and joint aspiration with corticosteroid injection. The most crucial aim was to maintain SUA levels below 360umol/L. Colchicine and NSAIDs were used as prophylaxis and losartan should replace diuretics' use in hypertension cases. (Zhang et al., 2006).

Figure 2.12 below shows a flowchart of the EULAR recommendations with all the possible case scenarios and decisions to be made by doctors. The first step consists of determining SUA level, followed by educating patients about gout, the lifestyle they should have and other comorbidities present. Once a proper medical history has been established, prophylactic treatment starts with allopurinol use or febuxostat in case of allopurinol allergy. In cases where those two drugs are not useful to achieve target SUA, combined therapy and pegloticase are used as a last option (Zhang et al., 2006; Ragab et al., 2017).

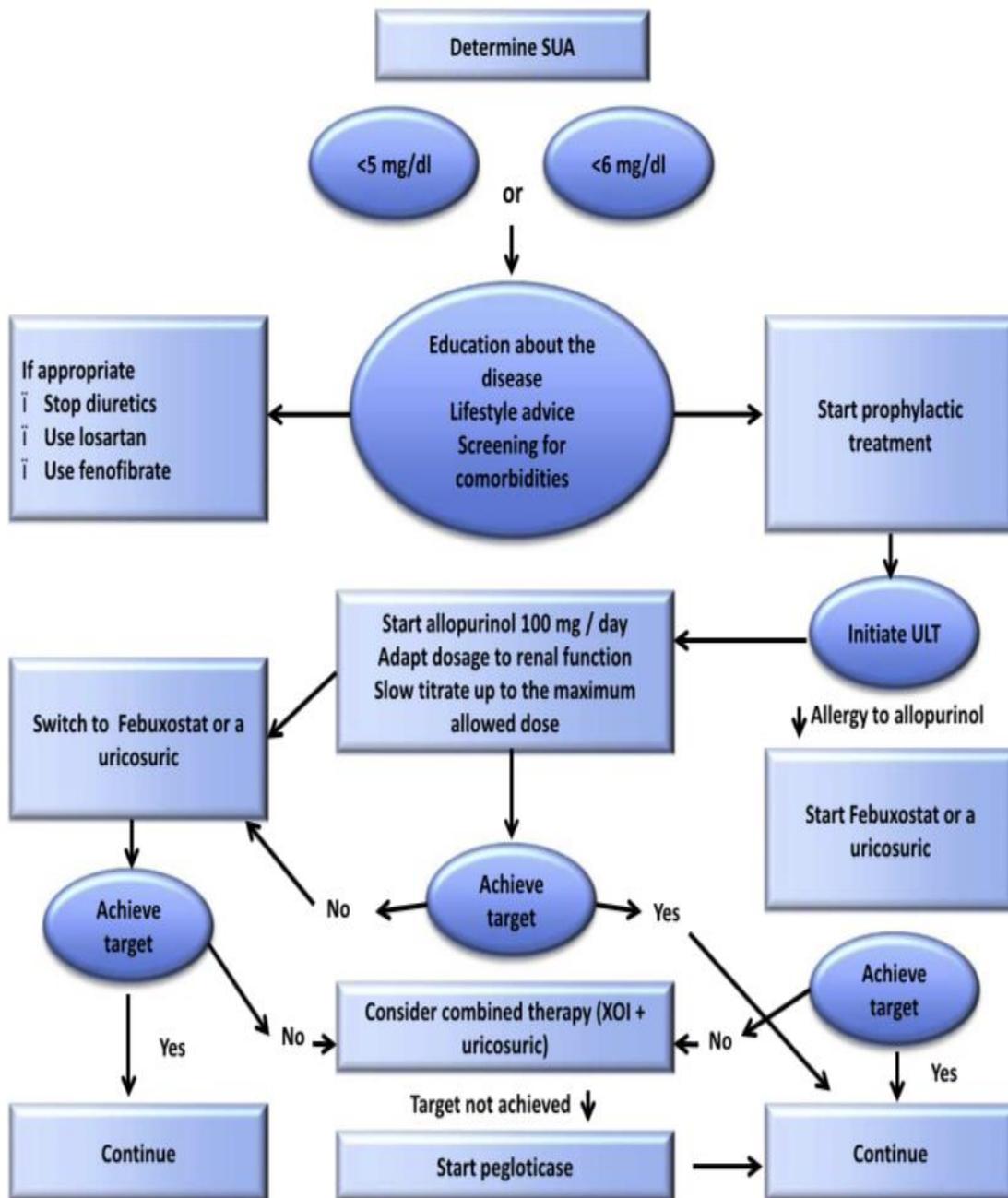


Figure 2-12 Simplified flowchart from the EULAR recommendations for management of hyperuricaemia in gout patients (Ragab et al., 2017).

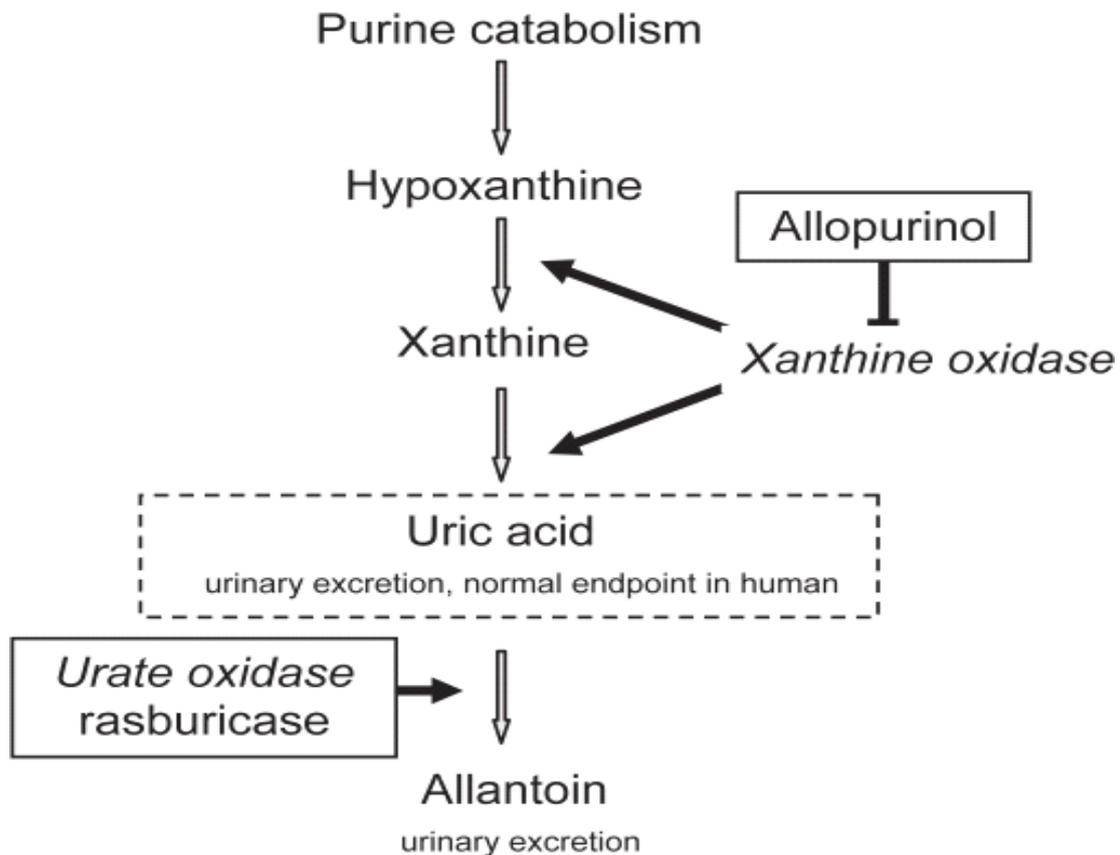
## 2.1.7.2 Urate lowering therapies (ULT)

### 2.1.7.2.1 Allopurinol

Allopurinol is the first-line, inexpensive urate-decreasing therapy commonly used in gout patients. 70 tonnes of allopurinol amounting to 240 million doses were used only in the 1980s (Lee et al., 2008). Allopurinol is used to treat gout, high levels of uric acid in the body caused by certain cancer medications and kidney stones. It is in a class of medications called xanthine oxidase inhibitors. It comes as a tablet to take by mouth, usually taken once or twice a day, preferably after a meal. The usual treatment plan starts with low dose allopurinol and it is gradually increased. The full benefit of allopurinol is usually felt after several months, but it may increase the number of gout attacks during the first few months. Allopurinol may be given with other medications like colchicine, to prevent gout attacks for the first few months. Mild side effects of allopurinol are; upset stomach, diarrhea and drowsiness. Some side effects are more serious, such as skin rashes, painful urination, blood in urine, irritation of the eyes, swelling of the lips or mouth, fever, sore throat, chills, loss of appetite, unexpected weight loss and itching (Lee et al., 2008).

Allopurinol inhibits xanthine oxidoreductase by its product oxypurinol, thus lowering plasma urate levels and increasing levels of substrates hypoxanthine and xanthine. Conversion of hypoxanthine to inosine, inosine monophosphate, adenosine and guanine monophosphates follows, which in turn leads to feedback inhibition of amidophosphoribosyl transferase and ultimately limits purine biosynthesis. The combined effect of allopurinol and oxypurinol is decreased urate production which ultimately decreases *de novo* purine synthesis. The first, 'pseudo irreversible' mechanism of xanthine oxidoreductase inhibition occurs when allopurinol acts as the substrate along with a stable reduced enzyme-oxypurinol complex. This complex slowly breaks down to release oxypurinol. Oxypurinol itself binds strongly to the reduced enzyme in the major and second type of inhibition. Oxypurinol has a longer half-life and higher plasma levels than allopurinol, contributing majorly to the latter's hypouricaemic effect (Lee et al., 2008). The full mechanism of allopurinol's action can be seen in Figure 2.13.

Excretion of oxypurinol occurs mainly through the kidney and this decreases in cases of renal failure and increases by uricosurics. Dose requirements are increased by a rise in body weight and diuretic use. Allopurinol is prescribed at a dose of 300mg/d or less in around 90-95% gout patients worldwide. However, due to an increase in uricemia and weight in patients nowadays, the recommended dose is not enough to reduce the SUA level to its target. The maximum dose for this drug is 800-900 mg/d in certain countries and in patients with normal renal function (Lee et al., 2008). Xanthine oxidase inhibitors are used for patients with joint damage, tophi or more than two gout cases per annum, after the acute gout phase is over. Patients with renal insufficiency are given allopurinol doses based on their creatinine clearance level (Lee et al., 2008). Allopurinol is the most commonly available drug in Malaysia, with superior efficacy compared to all other drugs (Ministry of Health, Malaysia, 2008). Allopurinol shows attractive advantages such as rapid lowering of serum urate levels and affordable cost over other drugs such as probenecid and febuxostat. However, 2-5% of patients taking allopurinol will show Severe Cutaneous Adverse Reactions (SCARs) which is the main factor limiting allopurinol's efficacious use in such cases. Patients should be warned that SCARs may manifest suddenly, in the first few weeks of treatment and that they should immediately stop taking the drug and visit their doctor as soon as possible (Lee et al., 2008).



Barred lines=inhibition. Arrows=activation or consequences

Figure 2-13 The mechanism of action of allopurinol is depicted here, where it inhibits xanthine oxidase action as seen by the back T symbol. Barred lines shows where inhibition is taking place and arrows shows activation or consequences. Adapted from American Society of Haematology, 2015.

#### 2.1.7.2.2 Febuxostat

Febuxostat, a xanthine oxidase inhibitor, is usually given orally, once a day, and proves to be more potent at a dose of 80-120 mg/d compared to 300 mg/d allopurinol. With its mixed renal and hepatic metabolism, it can be prescribed even in patients with moderate renal failure and creatinine clearance below 30 mL/min. This drug undergoes hepatic metabolism, where conjugation happens by uridine diphosphate-glucuronosyltransferase enzymes. Oxidation is carried out by CYP1A2, CYP2C8 and CYP2C9 to form active metabolites and

finally excretion happens via the kidneys (Lee et al., 2008). Febuxostat is usually used as an alternative drug in Malaysian hospitals, for patients who cannot tolerate allopurinol and probenecid. These two drugs have always been used as first-line treatment for gout treatment and febuxostat was only introduced in 2009 as an alternative (Ministry of Health, Malaysia, 2009). The American College of Rheumatology (ACR) recommends febuxostat as a first line ULD, while EULAR recommends it only for patients intolerant or refractory to allopurinol. Dose titration is again recommended here for ULD-induced flares, but without any proper evidence that tolerance to febuxostat increases subsequently (Zhang et al., 2006).

#### **2.1.7.2.3 Uricosurics**

This set of drugs lower the SUA level by increasing the uric acid excreted in urine, but has a disadvantage of increasing the risk of uric acid stone at the start of treatment. Uricosurics include probenecid, sulfipyrazone, benzbromarone, lesinurad and fenofibrate. These drugs are not administered as monotherapy in patients with previous history of UA stones or hyperuricuria.

#### **2.1.7.2.4 Colchicine**

Colchicine is used as an alternative drug for patients who cannot take NSAIDs and COX-2 inhibitors. This drug works by inhibiting numerous pro-inflammatory mechanisms and also elevates the level of anti-inflammatory mediators. 1.8 mg of colchicine is taken within 12 hours after flare onset and shows effectiveness as good as the initial higher doses recommended. However, this drug is not so efficient when administered long after the flare onset. Thus colchicine is only used within 12-24 hours of flare onset, according to the EULAR and ACR. Moreover, this drug has a narrow toxicity window, with several side effects such as gastrointestinal intolerance, neutropenia, multi-organ failure and eventual death. Colchicine is contra-indicated in patients with renal failure, renal insufficiency, hepatic failure, CKD stage 5 patients and is poorly tolerated in elderly patients (Ragaab et al., 2017).

#### **2.1.7.2.5 NSAIDs**

Potent NSAIDs are the first-line treatment of for acute gouty arthritis and the underlying inflammation and pain. This group of drugs can rapidly decrease pain and inflammation in acute gout cases, specially when the drugs are taken right after the attack starts and in proper, full doses. NSAIDs are used frequently at their maximum dose, together with proton inhibitors. Examples of NSAIDs used are diclofenac, indomethacin and ketoprofen, while aspirin is avoided as it causes urate retention (Ragaab et al., 2017).

#### **2.1.7.2.6 Steroids**

This set of drugs are best used when patients cannot use NSAIDs or colchicine, for example, in CKD patients. Steroids work by suppressing the immune system and decreasing inflammation in gout patients. Intra-articular steroid injections are very effective for the management of mono or pauci-articular flares, as recommended by the ACR and EULAR. However, steroids are contra-indicated in patients with hypertension and diabetes as it can worsen these conditions. Prednisone, considered as first-line therapy for flares, is effective when administered orally, at a dose of 30 mg/d for 7 days. Prednisone is metabolized in the liver to its active form, prednisolone, a glucocorticoid agonist corticosteroid (Ragaab et al., 2017).

#### **2.1.7.2.7 IL-1 blockers**

The ACR and EULAR recommended IL-1 blockers in patients with frequent flares and who are contraindicated for NSAIDs, colchicine and steroids. This set of drugs will regulate the immune system and inflammation in patients. MSU crystals can directly activate the NLRP3 inflammasome, a key regulator of the pro-inflammatory cytokine interleukin-1 $\beta$ , thus this set of drug focuses on IL-1 blockade. Anakinra, an IL-1 receptor antagonist, and canakinumab, a long-lasting antibody to IL-1 beta are two examples of drugs tested in trials, both with different mechanism of action. (Ragaab et al., 2017).

## 2.2 Severe Cutaneous Adverse Reactions (SCARs)

### 2.2.1 Introduction

Adverse drug reactions (ADRs) are the leading causes of morbidity and mortality worldwide. According to the World Health Organization (WHO), “an ADR is a response to a drug which is noxious and unintended, and which occurs at doses normally used or tested in man for the prophylaxis, diagnosis, or therapy of disease, or for the modification of physiological function”. ADRs causes millions of deaths yearly and has a medical cost of billions of dollars. Populations (Bharadwaj et al., 2012). ADRs was classified into 2 types, Type A ADRs and the idiosyncratic, Type B ADRs. Type A reactions account for more than 80% of all ADR cases and are caused by overdosage the drug’s pharmacological action. Type B consist of 20% of the unpredictable ADR reactions, which are either immune-mediated or non-immune related. Immune-mediated reactions were classified into 4 different subtypes as shown in Figure 2.16 below (Bharadwaj et al., 2012).

IgE-mediated and delayed hypersensitivities are the most common out of the 4 different subtypes. One-fifth of all ADRs are cutaneous ADRs, making upto 2-3% of hospitalized cases. SCARs form part of Type 4 delayed T-cell mediated reactions, as shown in Figure 2.16 below, and accounts for around 2% of all cutaneous reactions ADRs. Stevens-Johnson syndrome (SJS), toxic epidermal necrolysis (TEN) and drug rash with eosinophillia and systemic symptoms (DRESS) all form part of SCARs' clinical spectrum. Life-threatening SCARs develops in 0.1-0.4% of patients taking allopurinol, but results into a great disease burden worldwide (McDonagh et al., 2014). SCARs are known to have a high mortality rate (10-50%), high morbidity rate (60%), along with numerous complications and high global economic burden. Numerous risk factors are related to SCARs and can be summarized into two groups, namely drug user-related or drug-related. Drug user-related factors include age, sex, genetics, comorbidity and previous exposure, while drug-related factors only are the nature of the drug, cross-reactivity, exposure degree and route of administration. Out of all those factors, genetic factors are the most important, as shown in carbamazepine, abacavir and allopurinol-induced SCARs (McDonagh et al., 2014).

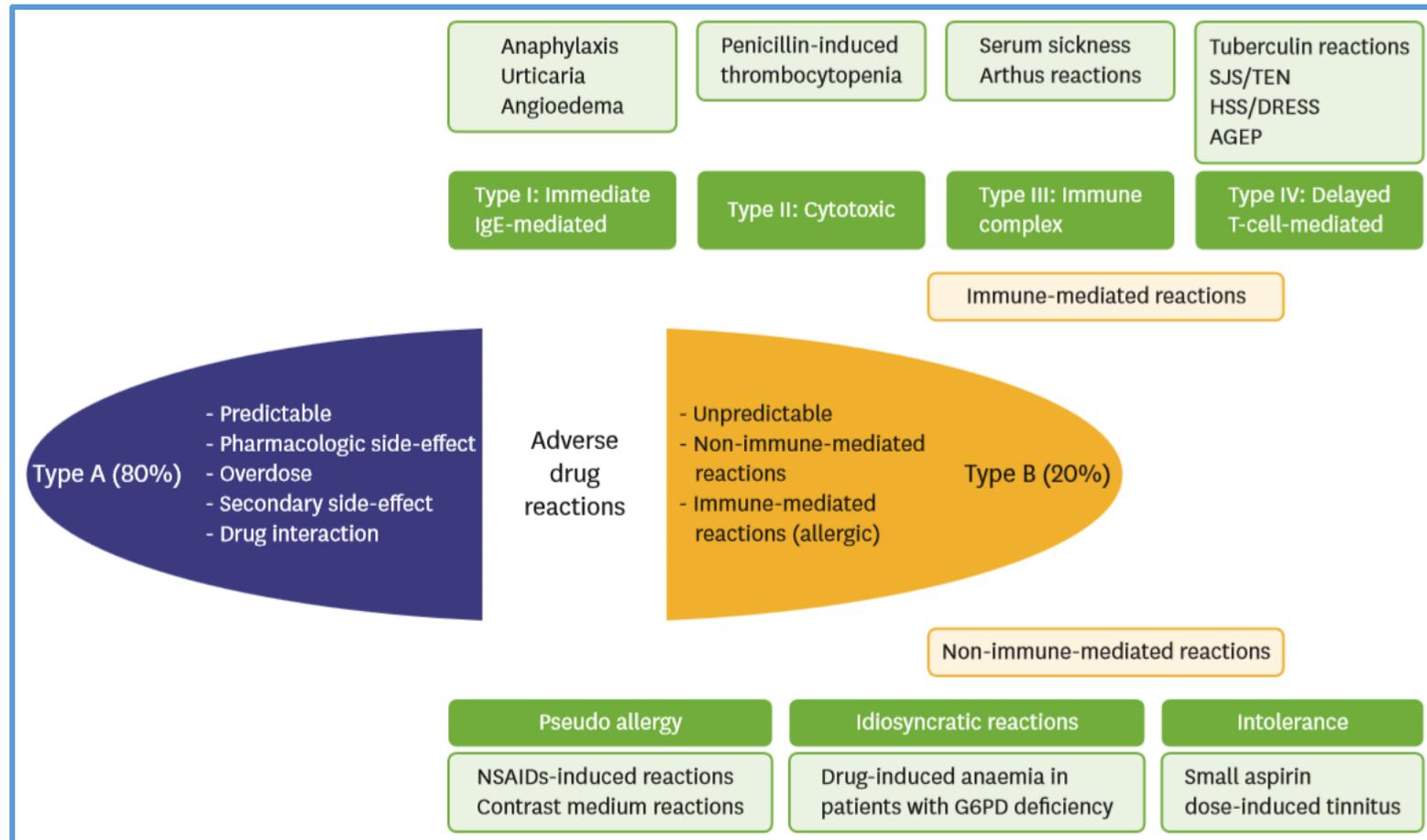


Figure 2-14 Classification of ADRs into Type A and Type B and further into immune-mediated and non-immune-mediated reactions. SCARs such as SJS/TEN, HSS/DRESS and AGEP are found in the Type 4: Delayed T-cell-mediated reaction box on the top right corner (Bharadwaj et al., 2012).

### **2.2.2 Classification of SCARs**

SCARs have some overlapping characteristics which makes it difficult to distinguish them apart properly. Key features that need to be taken into consideration are systemic and cutaneous involvement, latency periods, histological and laboratory characteristics and differential diagnosis generated. The different SCARs will be discussed and presented in the following subsections (Bharadwaj et al., 2012).

#### **2.2.2.1 Exanthematous drug eruption**

The most common type of drug hypersensitivity reaction is exanthematous drug eruption, also called morbilliform or maculopapular drug eruption. The latter are characterized by symmetrical erythematous macules and papules, which can be generalized and confluent and appear within a week on the initiation of drug treatment. In severe forms, the mucosae (oral, conjunctival, nasal, or anogenital) and skin appendages (hair and nails) may be involved. Other features of exanthematous drug eruptions are pruritus, mild eosinophilia and mild fever. All of these features are usually self-limiting, but without proper care, these may also progress to cause SJS/TEN and DRESS (Bharadwaj et al., 2012).

#### **2.2.2.2 Stevens Johnsons Syndrome (SJS) and Toxic Epidermal Necrolysis (TEN)**

SJS and TEN are severe systemic disorders whereby cellular apoptosis causes mucous membrane erosions and severe detachment of the epidermis, accompanied by constitutional symptoms. SJS and TEN are the 2 most fatal cutaneous ADRs in the hospital, accounting for 10-50% mortality rate. SJS and TEN have a high mortality rate of 5-10% and 30-40% respectively (Bharadwaj et al., 2012). Age, degree of skin involvement and serum urate concentration determines the prognosis of SJS and TEN. TEN is characterized by extensive, drug-related, skin damage of more than 30%, mortality rates of 30-40% and an incidence of 0.4-2.0 cases/million person-years. SJS is the result of drug action, infections

and a combination of both of those factors, with an incidence of 1-6 cases/million person-years. These figures might seem infrequent and small, but these conditions kill and disable healthy patients rapidly and is accompanied by an increasing economic and medical impact nowadays (Bharadwaj et al., 2012). More than hundreds of drugs have been linked to SJS and TEN, where the most dangerous are sulphonamides, allopurinol and anticonvulsant drugs. Large case studies have shown that allopurinol is a worrying frequent cause of SJS and TEN due to its long administration periods and its strong genetic association. A better knowledge on the causes of these fatal disease will pave the way for better medical therapy and decisions which will ultimately lead to decreased morbidity and mortality worldwide (Bharadwaj et al., 2012).

SJS and TEN are characterised by significant keratinocyte apoptosis in the dermis which detaches from the epidermis, followed by epidermal necrosis and mucocutaneous shedding. SJS and TEN are also accompanied by myocardial infarction, fever, gastrointestinal problems, hepatitis, respiratory and renal failure (Chung et al., 2008). The first symptoms of SJS/TEN appear after a few days and within 4 weeks of exposure to a specific drug. Early symptoms include fever (above 38°C), sore throat and ocular problems. This progresses to mucous membrane involvement and appearance of skin lesions, which gradually spreads everywhere. Lesions in the mucous membrane are known as blisters, erosions and erythema which target the eyes, oropharynx, nasopharynx, anus and genitalia (shown in Figure 2.17 below) (Chung et al., 2008).



Figure 2-15(a) Patient with dusky red confluent macular lesions, along with tense bullae and large areas of dermoepidermal detachment (shown by asterisk). Figure 1 (b) shows the typical distribution of skin and mucous membrane lesions in SJS/TEN. Figure 1 (c) shows overt mucous membrane involvement and moderate skin detachment with crusts in an SJS case (Chung et al., 2008).

SJS is a minor form of TEN, where detachment of body surface area (BSA) is less than 10%, while in TEN, BSA detachment is more than 30%. An overlapping phenotype of SJS and TEN is defined by BSA detachment in between 10-30%, as shown in Table 2.3 below. The acute phase of SJS/TEN involves failure of internal organs and results into hepatic injury, epithelial necrosis of the digestive/respiratory tract and kidney impairment. Internal organ failure is caused by loss of the skin barrier, infection due to leukopenia and disturbances of electrolytes and fluids linked to blisters (Chung et al., 2008). It is imperative to classify and distinguish all the specific clinical features for SJS, SJS/TEN overlap and TEN, to properly administer required treatment and care. The latter can be distinguished by their primary lesions, distribution on the body, mucosal involvement, systemic symptoms and their percentage of body surface area detached, as shown in Table 2.3 below.

Table 2-3 Detailed clinical features to distinguish SJS, SJS/TEN overlap and TEN. Plus (+) sign indicates strength of confluence, where + is lower than ++ (Chung et al., 2008).

Clinical entity	SJS	SJS/TEN overlap	TEN
<b>Primary lesions</b>	Dusky red lesions Flat atypical targets	Dusky red lesions Flat atypical targets	Poorly delineated erythematous plaques Epidermal detachment Dusky red lesions Flat atypical targets
<b>Distribution</b>	Isolated lesions Confluence (+) on the face and trunk	Isolated lesions Confluence (++) on the face and trunk	Isolated lesions (rare) Confluence (+++) on the face, trunk and elsewhere
<b>Mucosal involvement</b>	Yes	Yes	Yes
<b>Systemic symptoms</b>	Usually	Always	Always
<b>Detachment (percentage of body surface area)</b>	<10%	10%-30%	>30%

The SCORE of Toxic Epidermal Necrosis scale (SCORTEN scale) (Bastuji-Garin et al., 2000) is used universally to classify and determine the severity of SJS/TEN by using 7 independent risk factors, as shown in Table 2.4 below, and estimating the possible mortality rate for a patient, shown in Table 2.5 below. This method of classification should be applied in all the hospitals for a standardized and accurate SCAR diagnosis nationwide.

Table 2-4 Seven independent risk factors used to classify SJS/TEN as part of SCORTEN scale. The scale of 0 or 1 shown in the last two columns are attributed according to the risk factors they fall in (Bastuji-Garin et al., 2000).

Risk factor	0	1
Age	< 40 years	> 40 years
Associated malignancy	no	yes
Heart rate (beats/min)	<120	>120
Serum BUN (mg/dL)	<28	>28
Detached or compromised body surface	<10%	>10%
Serum bicarbonate (mEq/L)	>20	<20
Serum glucose (mg/dL)	<252	>252

Table 2-5 Second part of SCORTEN estimation by using Table 1's initial calculated scores and linking it to an estimate of the possible mortality rate in terms of percentage (Bastuji-Garin et al., 2000).

No of risk factors	Mortality rate
0-1	3.2%
2	12.1%
3	35.3%
4	58.3%
5 or more	>90%

Long term complications, even after recovery, further complicate matters and sometimes even require surgery. Adhesion of mucous membranes and narrowing of the digestive and respiratory tracts are the most common. Severe ocular problems are also commonly seen, such as severe dry eye, symblepharon, trichiasis, corneal conjunctival invasion and vision disturbance. The available treatments for SJS/TEN focuses mostly on withdrawing the suspected drugs first, then deals with symptomatic and supportive treatment. The crucial part lies in identification and withdrawal of the culprit drug and in cases where this is delayed, mortality and morbidity rates increase. Symptomatic and supportive care includes decreasing pain, skin care methods, restoring fluid balance, giving nutrients and avoiding severe complications (Nguyen et al., 2019).

Immunosuppression is currently being used as the main therapy for SJS-TEN, but often fails as secondary infections and complications comes about. Another debatable therapy is intravenous immunoglobulin, which was recently shown to decrease the symptoms and length of immune reaction. However, there is no treatment that can properly tackle these diseases to stop their significant death toll. The only way to fight SJS-TEN urgently is to prevent them from happening in the first place and to do this more research needs to be done on its genetic link to the HLA-B alleles. As it is said, 'prevention is better than cure'. HLA-B alleles are increasingly being seen in the research and medical world nowadays, where it is being used to prevent fatal immune-linked ADR (Nguyen et al., 2019).

### 2.2.2.3 DRESS/HSS

A severe idiosyncratic ADR is hypersensitivity syndrome (HSS) or DRESS, characterized by non-bullous skin lesions and involvement of internal organs. Incidence rates for DRESS are 1/1000 to 1/10000 drug users, along with a mortality rate upto 10%. Drugs frequently causing DRESS are allopurinol, sulphonamides, aromatic anticonvulsants, dapson and minocycline, with allopurinol being the most common. DRESS can be seen in the initial 8 weeks of allopurinol therapy and may also cause manifestation of eosinophilia, lymphadenopathy and mononucleosis-like atypical lymphocytosis. Early signs and symptoms of DRESS include fever (above 38°C) and rashes. Skin lesions can manifest as maculopapular/morbilliform rashes, exfoliative dermatitis and desquamation (shown in Figure 2.18) (Tsai & Yeh, 2010).

DRESS also includes manifestation of fever, hepatitis, facial oedema, nephritis and pneumonitis. Infiltration of eosinophils and lymphocytes in tissues ultimately affects internal organs, mainly the liver and the kidney. DRESS treatment first involves, removing the culprit drug and targeted treatment. Targetted therapy involves use of corticosteroids, intravenous immunoglobulin and antiviral agents in necessary cases. Moreover, long-term complications are also seen, namely Hashimoto's disease, Grave's disease, thyroiditis and type 1 diabetes (Tsai & Yeh, 2010).



Figure 2-16(A) Patient with DRESS showing scaly patch with papules on his forearm and Figure 2.16 (B) shows desquamation of his soles with petechiae visible. (Choudhary et al., 2013)

### 2.2.3 Immunological SCARs mechanisms

The immune mechanisms of SCARs consist of the activation of drug-specific cytotoxic T-cells, inflammatory cells and regulatory T-cells (T-Regs), along with the secretion of inflammatory cytokines. The first known effector mechanism for epidermal necrolysis in SJS/TEN is the appearance of drug-specific cytotoxic cells. The latter's effect is amplified by the huge production of death mediators, flawed negative regulation of drug-specific immune responses and modified anti-apoptotic pathways in target cells (Chung et al., 2008).

The pathogenesis of SJS/TEN suggest that they are mediated by the immune system as administering the same drug to the same person decreases the disease's incubation period and increases its intensity. It has been shown that drug presentation by the major histocompatibility class I (MHCI) causes clonal proliferation of the CD8+ cytotoxic T lymphocytes (CTLs), CD56+ NK and NKT cells, which thus causes an immune response. Skin lesions of SJS-TEN patients are shown to be permeated by CTLs and Natural Killer (NK) cells. However, the extent of keratinocyte death cannot be explained by the infiltration of inflammatory cells only in the skin lesions. There are numerous suggested pathways involved, i.e, granule-mediated exocytosis, for example granzyme B and perforin, and Fas-Fas ligand contact. Other studies show the involvement of a high amount of granulysin released from CTLs, NK and NKT cells, which causes severe proliferation of keratinocyte apoptosis (Chung et al., 2008).

Granulysin was shown to be the crucial cytotoxic molecule which causes major keratinocyte necrosis via the action of cytotoxic or NK-cell-mediated cytotoxicity in the absence of any direct cellular contact. Perforin-granzyme B pathway has a minor role, while Fas-Fas ligand interaction has no detectable effect. Severity of SJS/TEN increased with an increase in granulysin and interleukin 15 concentration, while interleukin 15 was highly correlated with mortality (Chung et al., 2008). Receptor-interacting protein kinase 3 (RIPK3) had a high expression level in TEN lesions, making the latter an important factor in apoptosis and necrosis of keratinocytes. DRESS syndrome is mediated by activated drug-specific T cells in skin and internal organs. Human herpes virus 6 (HHV6+) peripheral mononuclear cells are recruited to damaged skin, in order to allow virus transmission and replication of CD4+ T

cells. Patients with severe DRESS syndrome had a huge number of CD8+ T cells expressing granzyme B in their skin samples (Chung et al., 2008).

In certain cases, viruses are involved in DRESS syndrome, but this was not proved in in-vitro studies. A study showed that patients with DRESS syndrome had tumour-necrosis factor (TNF)  $\alpha$  and interferon  $\gamma$  secreted by circulating CD8+ T cells. HHV was recognised by more than half of the circulating CD8+ T cells, whereas CD8+ T cells in visceral or skin infiltrates recognised the Epstein-Barr virus (EBV) mainly. EBV replication could also be triggered by culprit drugs, via the action of EBV-transformed B lymphocytes. Researchers compared the type and density of inflammatory cells present in SJS, TEN, DRESS syndrome and non-SCAR cytokine levels. SJS and TEN patients had more pro-inflammatory cytokines (TNF $\alpha$ , interleukin 6 and interferon  $\gamma$ ) and anti-inflammatory cytokines (interleukin 10 and interleukin-1-receptor antagonist, compared to other cutaneous ADRs (Chung et al., 2008).

SJS and TEN again distinguished themselves for analysis of lymphocyte/white blood cell subsets and immunoglobulin profiles, suggesting the possible involvement of a viral infection coinciding with drug exposure. These evidences show a type of non-drug-specific immune activation, where T-cells are ready to react, similar to a viral infection. Another factor influencing SCARs' phenotype is T-regs during acute and chronic SCAR stages. T-reg frequency did not change during early and late stages of SJS/TEN, but non-T-reg frequency did increase upon resolution of SJS/TEN. Levels of functional T-regs increased during the acute stages of DRESS syndrome, while the opposite occurred for SJS/TEN. Resolution of DRESS syndrome caused T-regs to become functionally deficient, while functionality came back after SJS/TEN stages (Chung et al., 2008).

## 2.2.4 SCARs treatment

### 2.2.4.1 Principles of symptomatic treatment

The main SCAR-management strategies are symptomatic, shown in **Table 2.6** below, where the focus centres on preventing short-term mortality, morbidity and severe long-term sequelae (Chung et al., 2008). Table 2.6 shows the possible specific and symptomatic treatment, the most crucial step being of culprit-drug identification and withdrawal, leading to better prognosis in all SCAR cases. An increased morbidity risk is observed in drugs with longer half-lives. Acute SCAR stages require intensive care due to the possibility of multiorgan failure and major fluid loss caused by skin damage. Supportive care then focuses on restoration of haemodynamic equilibrium and prevention of life threatening complications. Erythroderma, or epidermal detachment, leads to numerous complications, such as fluid and protein loss, renal insufficiency, hypovolaemia, sepsis and thermal dysregulation. This requires immediate and daily fluid replacement, maintaining an environmental temperature of 28°C and hypercaloric and hyperproteic feeding through a nasogastric tube. Skin debris removal pain in SJS/TEN is limited by using opioid agonists, along with proper respiratory monitoring (Chung et al., 2008).

The outcomes and sequelae in Table 2.6 are numerous and severe, showing the true extent of damage done to SCARs patients, along with the constant management required even after resolution. SJS/TEN have more sequelae along with vigorous management up until 5 years after resolution, along with numerous specialist consultations over the years (Chung et al., 2008).

Table 2-6 Overview of SCARS treatment, outcome, main sequelae and management after resolution. (RCT= Randomised controlled trial) (Chung et al., 2008).

SCARs	Treatment		Outcome	Sequelae	Management after resolution
	Specific	Symptomatic			
<b>SJS/TEN</b>	Drug withdrawal, No RCT*- validated curative treatment	Supportive care, cutaneous/mucous membrane care, enteral feeding, fluid-loss treatment, analgesia, no systematic intubation, environmental temperature $\geq 28^{\circ}\text{C}$ , anxiolytics	Bacterial superinfection, visceral-specific involvement, lung failure	Dystrophic scars, hyperpigmentation, alopecia, nail loss, visual loss, synechiae, dry eye, symblepharon, dental agenesis, sialadenitis, tooth decay, genital synechiae, psychiatric disorders	Patch testing at month 6; follow-up at least at month 2, month 6, month 12, and every year for 5 years; specialist consultations: dermatology, ophthalmology, ear, nose, and throat examination, gynaecology, psychiatry, pulmonology
<b>DRESS</b>	Drug withdrawal, No RCT*- validated curative treatment, Topical/systemic corticosteroids	Symptomatic, antipyretics	Acute organ failure, virus reactivation, relapses	Autoimmune diseases, lupus, thyroiditis, diabetes, scleroderma	Patch testing at month 6; follow-up at month 2, 3, 4, 5, 6, and 12, then annually

#### **2.2.4.2 Dermatological care**

Daily wound care is crucial in SJS/TEN patients and involves diluted antiseptic sprays and antiseptic baths. Skin injuries must be avoided at all costs, for example, no adhesive must be used and care must be taken during transportation and manipulation of the patients. Aggressive skin debridement must be avoided as the necolytic epidermal sheets act as biological dressing themselves. The use of petroleum jelly is encouraged on all areas of detached skin, along with the occasional use of non-adhesive dressings on pressure points such as the back. Mucosa lubrications with emollient are used to decrease formation of mucosal adhesion and other functional sequelae during acute SCAR phases. Mucosal bleeding and erosions necessitates the use of mouthwash, topical analgesic, swabs, local administration of adrenaline and clotting agents. Ocular management for inflammatory debris removal includes daily saline rinses and topical anesthesia, followed by debris removal with cotton buds or smooth blunt instruments. Amniotic membrane transplantation has also been proposed to avoid eyelid scarring, conjunctival and corneal sequelae. DRESS syndrome's dermatological care mostly focuses on skin moisturisation and in some cases mucous-membrane management can be done similarly in aforementioned methods for SJS/TEN (Chung et al., 2008).

#### **2.2.5 Outcome and sequelae**

SCARs are all life threatening, with a confirmed risk of severe sequelae, specially SJS, TEN and DRESS. The main complication of acute SJS/TEN is visceral involvement, such as sepsis, renal failure and lesions in the intestine, eye and lungs. Respiratory insufficiency in SCARs patients is either caused by SCAR's influence on organs itself or by superinfection, caused by the inhalation of foreign substances. Patients are vulnerable to infections in the acute SCARs stages due to defective skin barrier function and translocation of gut bacteria (Chung et al., 2008).

The most common sequelae seen for SJS/TEN patients are dystrophic scars, hyperpigmentation, alopecia, nail loss, visual loss, synechiae, dry eye, symblepharon, dental

agenesia, sialadenitis, tooth decay, genital synechiae and psychiatric disorders. The highest cause of death in SJS/TEN is sepsis. Acute-stage SCARs have significant mortality range of at least 10% for SJS and 40% for TEN, along with a total hospital mortality rate of 22%. SJS/TEN patients admitted to hospitals are predicted to have a 3-day mortality by SCORTEN, along with a high individual death risk (Chung et al., 2008).

In the absence of laboratory data, a five-point auxiliary score was used to predict the severity of illness for SJS/TEN patients. The latter showed that after a year, mortality rates were still high, around 34% for SJS and 49% for TEN patients. Sequelae, which occur after remission in SJS/TEN patients, are numerous and includes cutaneous, buccal, dental, ocular, pulmonary and genital lesions, as well as psychiatric disorders (Chung et al., 2008). Routine screening is needed for sequelae, in order to reduce their impact on quality of life. Cutaneous and ocular sequelae are the most frequent ones and need to be treated properly. Chronic SCAR stages have up to 65% of patients with late eye complications and these can worsen, even after discharge from the hospital (Chung et al., 2008).

Managing ocular sequelae includes rehabilitation of visual function and treating corneal inflammation, dry eye, and refractory ocular disease. SJS/TEN patients also experience salivary acidity which causes tooth decay, chronic sialadenitis and dental atrophy. Male genital synechiae with circumcision are often seen, along with complications in women such as, strictures of vaginal mucosa, birth-canal stenosis which cause difficulties in vaginal delivery and normal sexual intercourse. DRESS syndrome is mainly linked to pulmonary/myocardial lesions and haemophagocytosis, with an overall acute-stage mortality of 5%-10%. Studies showed a cumulative incidence of long-term DRESS sequelae of 11.5%, with mostly autoimmune diseases, PTSD, anxiety and depression. DRESS remission can be followed by chronic virus activation, which in turn causes autoimmune diseases such as diabetes, scleroderma, thyroiditis and lupus erythematosus (Chung et al., 2008).

## **2.2.6 ADRs and SCARs reports in Malaysia**

High levels of SCARs caused by allopurinol has been reportedly shown in numerous countries such as Malaysia, Europe, Israel and many other Southeast Asian populations (Tsai & Yeh, 2010). A study on Asian populations showed that allopurinol, carbamazepine, phenytoin, lamotrigine and sulfamethoxazole are linked to SJS/TEN in more than 50% of cases. Allopurinol was singlehandedly found to be the second SJS/TEN causing drug in 19.65% cases in Asia, after carbamazepine. Moreover, allopurinol was responsible for 22.78% of all SJS/TEN cases in Europe populations (Tsai & Yeh, 2010). The two main sources of SCARs/ADRs reports in Malaysia are from Malaysian Adverse Drug Reactions Advisory Committee (MADRAC) and data collected through various studies. Data from MADRAC (section 2.3.6.1) shows the increase in number of ADR reports with the years, percentage of different SCARs recorded, the identification of allopurinol as a culprit drug and inclusion of warnings and safety measures to be implemented in hospitals. Section 2.3.6.2 presents SCARs data mined from different hospitals in different states of Malaysia, specifically Johor, Pahang and Penang. Few studies were available and they performed four to ten year reviews of SCAR cases and patterns in those states and compiled them for a clearer view of the situation.

### **2.2.6.1 Malaysian Adverse Drug Reactions Advisory Committee (MADRAC)**

The annual ADRs reports generated by the Malaysian Adverse Drug Reactions Advisory Committee (MADRAC) were analysed for several years. Reports from 2000 to 2019 were accessed and each year allopurinol was mentioned, along with its link to SCARs. In 2000, 792 reports were recorded and this increased to 5850 case reports in 2009. The total number of new ADR reports increased from 5,550 cases in 2010, to 12,306 cases in 2015, showing a significant increase in five years. Moreover, all the cases reported were benchmarked against the VigiGrade™ Completeness Score, used predominantly by the WHO Collaborating Center for International Drug Monitoring in Sweden. This score range shows the level of clinical accuracy in a case, with a poorly documented case at 0.07 and a

well-documented one at a score of 1.00. NRPA applied this score technique and the completeness core increased from 0.45 in 2010 to 0.72 in 2015, showing the continuous effort to improve quality of ADR case report. The main type of ADR reported was related to skin and subcutaneous tissue disorders, making up 30.0% of the total reports in 2015 (MADRAC bulletin 2016).

Allopurinol was mentioned as the cause of ADRs as early as the year 2000, where 493 reports were obtained, with more than 90% of them relating to skin and appendages disorder. 16 cases out of 493 were fatal and allopurinol was the main contributory factor. A review of all NPRA reports, showed that the commonest allopurinol-induced SCARs recorded from 2001 to 2009 were SJS (26.9%), maculopapular exanthem (23.2%) and TEN (5.7%), along with severe reactions in a total of 36.8% of the total number of patients. The main indication for allopurinol usage was for gouty arthritis (59%), asymptomatic hyperuricaemia (34%), non-specific arthritis/arthralgia (3%), renal calculi (2%) and chemotherapy treatment (2%). Fifteen patients died due to the SCARs, where 6 of them were given allopurinol for asymptomatic hyperuricaemia.

An alert in 2010 was raised as allopurinol was not being prescribed for the right indications and cases. Allopurinol is contraindicated in acute gout and asymptomatic hyperuricaemia as stated by the Ministry of Drug Formulary. However, it was noted that healthcare professionals were not adhering to these clear instructions and the prescribing information, with 30.2% of cases recorded with inappropriate indications in 2010 (MADRAC bulletin 2010). In 2010, the incidence of allopurinol-induced SCARs was 15 cases per hundred thousand population years, with the most common SCARs being SJS (27.3%), maculopapular exanthema (23.1%) and TEN (5.7%) (MADRAC bulletin 2012).

Deeper analysis of the all the data since 2000 showed that 21% of all the ADR reports were linked to allopurinol use for asymptomatic hyperuricaemia. In 2002, 60% of ADR reports on allopurinol, were linked to its inappropriate use. This problem was partly decreased by disseminating an advisory to all prescribers in 2004, with the correct allopurinol indications. A decline in the inappropriate use of allopurinol was seen in 2005 and 2006, however this increased again (25-30% increase) from 2007 to 2009. This triggered the use of a list risk minimisation actions in all hospitals, which lead to a decrease in number of cases to 5.6% in

2010 and 11.4% in 2011. The latter figures shows that inappropriate use of allopurinol still remains a cause for concern and contributes to SCARs occurrence (MADRAC bulletin 2012).

It was highlighted that allopurinol should not be used indiscriminately in any case with high SUA, that it doesn't have anti-inflammatory action and may intensify and prolong inflammation during acute gout phase. Moreover, a warning was issued about colchicine's severe drug interactions with P-glycoprotein and strong CYP3A4 inhibitors. Several deaths were recorded when these drugs were administered together. Patients with hepatic and renal impairment taking colchicine should not take P-glycoprotein or strong CYP3A4 inhibitors. Colchicine dose must be reduced or treatment stopped altogether in patients with normal hepatic and renal function in case the other two are needed urgently. Grapefruit and its juice must both be avoided while taking colchicine (MADRAC bulletin 2010). In 2018, the NPRA reported 25,127 ADR case reports, doubling the number of cases compared to 2015 as aforementioned. Moreover, a safety alert was reported about Beta-lactam antibiotics causing SCARs in January 2019 (MADRAC bulletin 2019).

### 2.2.6.2 Data mining from Malaysian hospitals

Data available from different states of Malaysia were limited and mined, to see the trend of allopurinol-induced SCARs and its mortality rates. A tertiary hospital in Johor, the Hospital Sultanah Aminah Johor Bahru, was analysed for the pattern of SCARs in 280 new patients obtained from 2001 to 2008. The most common SCARs seen were maculopapular eruption (39.5%), SJS (28.1%), DRESS (6.8%), TEN (5.7%) and others. Allopurinol was implicated in 13.9% of the total SCARs cases recorded and was amongst the top three main causative drugs of SJS/TEN at a percentage of 18.8%. DRESS was mainly caused by allopurinol, at 52.6% of the total DRESS cases. This study was compared to other papers worldwide and similar trends were seen, with allopurinol being amongst the top three high-risk drugs most often associated with SJS/TEN. Halevy et al. showed that 17.4% of the total SJS/TEN cases were caused by allopurinol in Johor (Ding et al., 2010).

Lin et al. (Lin et al., 2005) and Khoo et al. (Khoo et al., 1996) both showed similar high percentages of allopurinol-induced SJS/TEN cases, at 17.1% and 13.0% of the total number of cases, respectively, in Johor. In an earlier study, allopurinol-induced ADRs only consisted of 8% of the total cases, from 1992 to 1997. The National Pharmaceutical Control Bureau, Ministry of Health, Malaysia, received an increasing number of allopurinol-ADRs cases, from 22 cases in 2002 to 80 cases in 2008, with a total of 71.8% of severe ADRs from allopurinol (Ding et al., 2010).

A four year review of SCARs pattern in Pahang, in the Department of Dermatology of Hospital Tengku Ampuan Afzan, was performed from 2013 to 2016. A SCARs incident rate of 0.34% was noticed amongst 7,353 new patients, with the highest rate seen in the indigenous and Indian people (0.63% and 0.62% respectively). Allopurinol was identified as the most common culprit drug in all SJS and TEN cases, at 46.7% and 33.3% respectively. Moreover, this study discovered that allopurinol has been in the top three culprit drugs causing SCARs for around 20 years. The author also compared clinicoepidemiological SCAR studies in Malaysia to various Asian countries, showing that SJS was the most commonest SCAR reported and that Malaysia had a higher mortality rate compared to other countries (Ramalingam, 2018). Two other Malaysian studies were analysed by the author, the first one in Johor Bahru with 144 total SCAR cases, 26.4% SCAR cases linked to allopurinol, a SCAR

incidence of 0.34%, prevalence of 39.8% and a mortality rate of 6.9%. The second study in Klang had 33 SCAR patients, 33.3% of allopurinol-induced cases and a mortality rate of 6.1% (Choon et al., 2012).

A 10-year review of SCARs cases was carried out in Penang General Hospital from 2006 to 2015, resulting in a total of 189 mined SCARs cases. The rate of hospital admission due to SCARs was 0.3/1000, similar to other international studies. SJS was the commonest SCAR present (55.0%), followed by TEN (23.8%) and DRESS (12.7%) and they had respective mortality rates of 1.9%, 13.3% and 12.5%. Allopurinol was implicated in 18.9% of the total SCARs cases, in 20.8% of SJS/TEN cases and 46% of all DRESS cases. Allopurinol being the commonest culprit drug was also confirmed internationally in studies from Asia, Europe and Israel. The higher percentage of allopurinol-induced DRESS cases compared to SJS/TEN was also confirmed in other local studies such as Ding WY et al. (52.6%), Choon SE et al. (44%) and Tee SH et al. (40%) (Loo et al., 2018).

## 2.3 Pharmacogenomics

### 2.3.1 Introduction

Pythagoras first recognized pharmacogenomics when he connected the dangers of fava bean ingestion with haemolytic anemia and oxidative stress, around 510 BC. The latter was linked to G6PD deficiency and called favism later on in the 1950s. This step marked the official start of the pharmacogenomics field and the first official publication was made in 1961. Friedrich Vogel first coined the term pharmacogenetic in 1959 and the first proof of genetic involvement in drug metabolism was confirmed in late 1960s. The 1990s announced the spreading of the term pharmacogenomics worldwide, with the first FDA approved pharmacogenetic test for CYP2D6/CYP2C19 in 2005. The first step in understanding the pharmacogenomics field is understanding its genetic source, which is the Human Leukocyte Antigen (HLA) genes, implicated in drug induced SCARs (Shankarkumar, 2004). Section 2.4.2 below will give an introduction on the HLA genes, followed by its mechanism of action in SCARs in section 2.4.3. Once a proper foundation has been set to understand the HLA genes' role in SCARs, section 2.4.4 will explore the numerous pharmacogenomics success stories worldwide, along with its limited presence and study in Malaysia.

### 2.3.2 Human Leukocyte Antigen (HLA) genes

Genetic factors are found to be strongly linked and implicated in the aetiology of Type B ADRs after drug administration, specifically the Human Leucocyte Antigen (HLA) genes. The HLA loci are known to be highly polymorphic in human beings with the HLA-B locus having more than 1500 alleles which in turn have countless variants each. The HLA genes, also known as the Major Histocompatibility Complex (MHC) genes, codes for cell-surface receptors which capture and present peptides to T cells in an immune response. The MHC molecules can be classified into 2 types; the MHC Class I in nucleated cells and the MHC Class II found in antigen-presenting cells (APCs). MHCI and MHCII are involved in peptide presentation to CD8+ and CD4+ T cells respectively. HLA-A, HLA-B and HLA-C loci codes for MHCI while HLA-DR, HLA-DQ and HLA-DP loci codes for MHCII molecules. All these genes are linked in the MHC region on chromosome 6p21.3. HLA can be broken down into 3 classes; namely Class 1 (HLA-B, -B, -C), Class II (HLA-DR, -DQ, -DP) and Class III (Bf, C2, C4A). Naming of HLA genes follows a simple nomenclature as shown in Figure 2.17 below (The Sequencing Center, 2022).

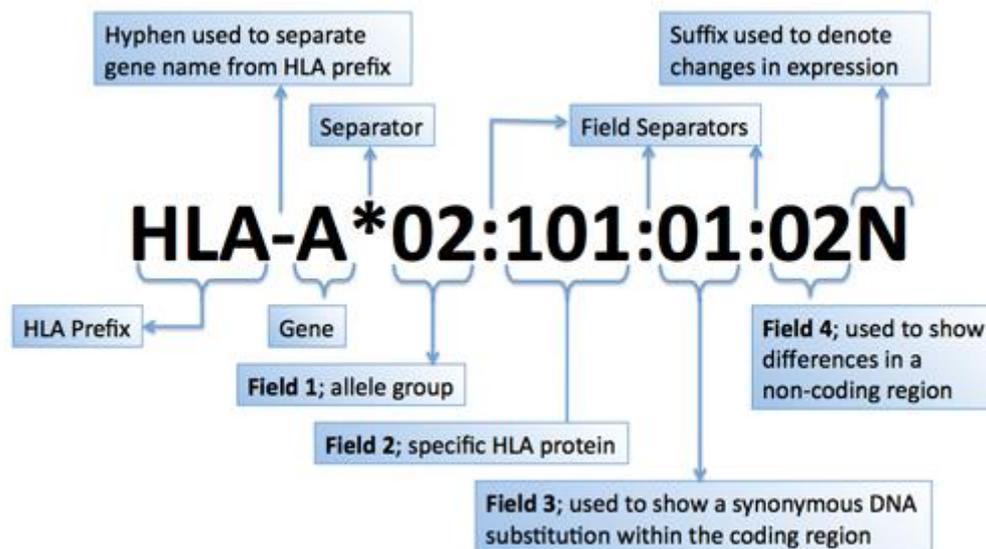


Figure 2-17 HLA Nomenclature, where the HLA prefix is followed by the specific gene (A,B or C), a separator, the allele group and more detailed information about a specific gene/allele (The Sequencing Center, 2022).

Linkage disequilibrium is known to occur between allelic combinations at linked loci, thus suggesting that this phenomenon may be the cause of the numerous ADRs-HLA associations. MHC1 molecules are found to be at the core of immune responses found in allopurinol ADRs and SCARs (Bharadwaj et al., 2012). MHC1 molecules are made up of a polymorphic  $\alpha$  chain with 3 domains ( $\alpha 1$ ,  $\alpha 2$  and  $\alpha 3$ ) and a conserved single domain called the  $\beta 2$ -microglobulin.  $\alpha 1$  and  $\alpha 2$  domains form the peptide-binding groove of MHC1 with a  $\beta$  sheet at the base of the peptide-binding cleft and the cleft walls formed by an  $\alpha$  helix. The binding cleft consists of peptide-binding pockets which will bind to amino acids' side chains and anchor the whole peptide to the MHC1 molecule. Some amino acid side chains will interact less with the MHC1 and more with the T cell receptor (TCR). In the presence of a neopeptide, the latter will be presented by MHC1, causing specific complementary TCRs on T cells to be activated, ultimately leading to an antigen-specific, CD8+ T cell response. Interactions will thus occur between TCRs of CD8+ T cells and the MHC1-peptide complex. CD8+ T cells proliferation follows causing effector cytotoxic functions (Bharadwaj et al., 2012).

The main function of HLA antigens is to recognise and fight against microorganisms. With the extensive polymorphism of the HLA loci, it ensures that the human body can deal with any type of attack from foreign organisms. Furthermore, polymorphisms are shown to be population specific. Linkage disequilibrium is also seen in HLA alleles which causes HLA Class I, II, III to be inherited together, thus diverging from Mendel's Law of independent segregation and creating a positive selective advantage. Cross-reactivity also occurs in HLA alleles where one antibody can react with numerous antigens and even share epitopes between antigens. This is due to a high similarity in the amino acid sequence, thus the molecular conformation, of HLA molecules (Shankarkumar, 2004).

Variations in HLA allotypes may be of 1-2 or 30 residues. Differences in allotypes are found mostly in the residues which will contact peptides within its peptide-binding groove, causing a change in the latter's structural and electrochemical layout. These changes in layout will cause HLA molecules to establish a distinct peptide specificity for each polymorphic allotype as well as distinct peptide-binding repertoires and motifs. Thus HLA allotypes with less similarities will have distinct binding sites for TCRs due to their different MHC exposed region and peptide repertoires. Micropolymorphisms of 1-2 amino acids will alter the

peptide presentation, dependency on the tapasin chaperone and ultimately impacts on T cell activation and recruitment. Immune responses by T cells thus depend on an individual's specific MHC allotype's capability in presenting new peptides from pathogens. Heterozygous HLA loci maintain HLA polymorphism by an elevated total number of peptide repertoire, which ultimately enhances the probability of peptide presentation from pathogens (Bharadwaj et al., 2012).

The variation in HLA frequencies and polymorphism in different populations arose due to specific selective pressure in different geographical areas. MHC restriction happens when peptides are bound to MHC molecules and the latter is recognized by T-cell receptors (TCR). Natural killer (NK) cells do not have TCRs and will be involved in cytotoxicity. NK cells are shown to destroy cells with a lower class I molecule expression, such as in some virus-infected cells or tumours. Furthermore, NK cell receptors are shown to bind to HLA class I molecules which in turn regulates NK cells' activity (Chung et al., 2008).

### 2.3.3 Pathogenesis of HLA induced SCARs

An increasing number of ADRs are being associated with HLA-B genes nowadays and this gave rise to different mechanisms of action in an effort to explain these complex interactions. These mechanistic models, as shown in Figure 2.18 below, explain the HLA genes' role, the recognition of drugs by T cells and the consequent boost of immune response caused by T cells (Nguyen et al., 2019). The first mechanism for MHC-dependent T cell action is the hapten-prohaptent concept (Figure 2.19 A & B). Haptens are chemically reactive small molecules, which bind covalently to bigger proteins/peptides and cause an immune response. Pro-haptens are on the contrary, not chemically reactive, but become so after being metabolized (Nguyen et al., 2019). This mechanism suggests that a drug, also called the hapten, will react with a self-protein, thus generating a haptentated product. This modified product will be processed to form a new MHC ligand on the MHC molecule which will in turn stimulate antigen-specific T cells on the cell surface. This can be seen clearly in Figure 2.18 A below (Nguyen et al., 2019).

One example of the hapten model is penicillin-induced hypersensitivity, where penicillin derivatives bind to serum albumin and go through intracellular processing, in order to produce chemically modified peptides and cause an immune response (Nguyen et al., 2019). In the second mechanism rapid T cell activation occurs which causes the immune response to be independent of antigen processing and cellular metabolism. This gave rise to the concept of pharmacological-interaction (p-i) with immune receptors. In the p-i concept, the drug directly binds to T-cell receptors or MHC proteins by non-covalent bonds on the cell's surface (as shown in the Figure 2.18 C below) eliciting a rapid immunogenic complex formation. This concept is known to be unstable compared to the hapten concept (Bharadwaj et al., 2012). The p-i concept can be shown in CBZ-induced SJS/TEN, where CBZ binds directly to the protein encoding the HLA-B\*15:02 allele, via a non-peptide processing pathway. This was confirmed in studies showing that fixation of antigen-presenting cells only can cause a specific immune response, without the involvement of the antigenic peptide-processing pathway. The latest physiopathological hypothesis discovered was the altered peptide repertoire model, shown in Figure 2.18 D. In this concept, the drug binds non-covalently in the MHC's binding pocket, altering the chemistry of the binding cleft and

the self-peptide repertoire and ultimately leads to cytotoxic T-cell activation. In the abacavir-hypersensitivity model, ABC changes the repertoire of self-peptides by triggering conformational alterations in endogenous peptides presented by the protein encoded by the HLA-B\*57:01 allele. This in turn leads to a polyclonal T-cell response and induces hypersensitivity reactions. The drug does not directly interact with the HLA repertoire here, but it interacts with peptides which change the binding cleft of the HLA molecule, and which is in turn recognized as foreign antigens by APC (Bharadwaj et al., 2012).

Once the immune system has been stimulated by the aforementioned mechanisms, cytotoxic T-cells are in turn activated, such as CD4+, CD8+ and regulatory T-cells, as shown in Figure 2.18 below. The different cytotoxic T-cells will produce cytotoxic proteins and thus cause SCARs development, be it SJS/TEN, DRESS or acute generalised exanthematous pustulosis (AGEP) (Nguyen et al., 2019).

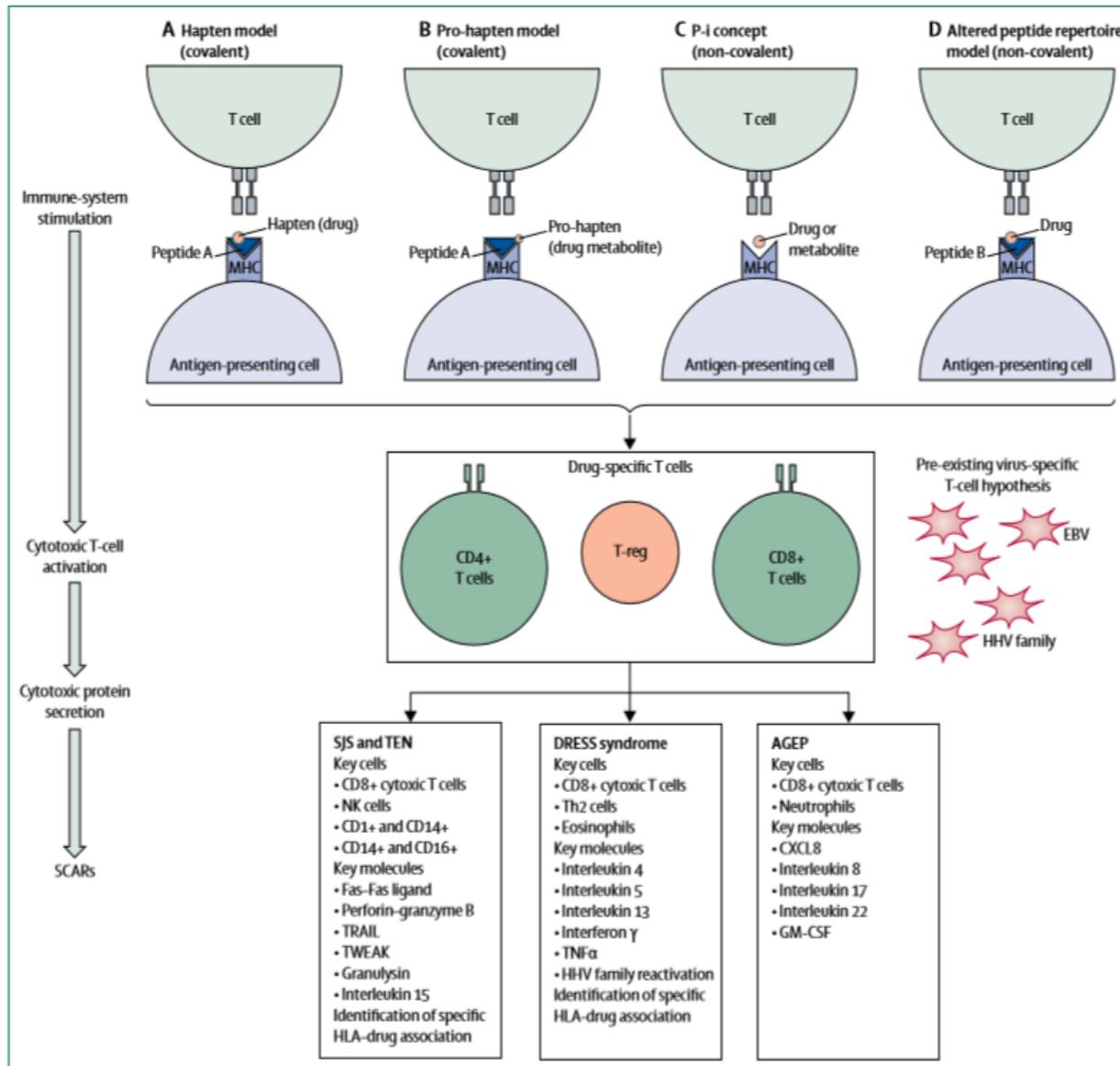


Figure 2-18 Schematic diagram of the different mechanisms of action for immune stimulation of the Major Histocompatibility Complex (MHC) involved in SCARs development. Figure 2.19 A & B depict the hapten and pro-hapten models while C and D show the p-i and the altered protein repertoire concept (Nguyen et al., 2019).

### 2.3.4 Pharmacogenomics success stories

The strongest ADR-HLA associations discovered were the role of HLA-B\*57:01 in Abacavir hypersensitivity in HIV cases, the role of HLA-B\*15:02 in SJS patients taking carbamazepine and most recently, the role of the HLA-B\*58:01 allele in allopurinol hypersensitivity syndrome (HSS), SJS and TEN (Bharadwaj et al., 2012). The odd ratios of these ADR-HLA associations are greater than 500, 1000 and 800 respectively. These strong HLA associations are clearly involved in disease progression, while the weaker associations suggests the presence of other contributing factors which ultimately makes the drug hypersensitivities' mechanisms more complicated to decipher (Bharadwaj et al., 2012). The first success in pharmacogenomics implementation started with abacavir (ABC) hypersensitivity syndrome and the HLA-B\*57:01 allele. In 1998, ABC was first discovered to cause abacavir hypersensitivity reaction (ABC HSR) in its premarketing stage, where 8% of treated patients experienced malaise, fever and gastrointestinal side effects. 70% of patients later experienced mild to moderate rashes over their body (Mallal et al., 2008). In 2002, two studies first showed the association between the HLA-B\*57:01 allele and ABC HSR in Caucasian populations. This association was then verified immunologically by using patch testing. 100% of white and black patients in a case-control study was shown to have positive patch test results, ABC HSR and the presence of the HLA-B\*57:01 allele (Mallal et al., 2008).

The Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines included HLA-B\*57:01 screening prior to ABC treatment in 2012. It took around 15 years for this to be accepted and implemented in clinics worldwide and treatment guidelines. A decline in ABC HSR incidence was proven, along with this strategy's cost-effectiveness (Karnes et al., 2020). However, it was a unique model with a positive predictive value of 55%, negative predictive value of 100% and narrow HLA restriction. The higher positive predictive value was the turning point which allowed easy implementation in clinical settings.

The second pharmacogenomics global success story shows how SJS/TEN emergence in epileptic patients were linked to carbamazepine's (CBZ) reaction with HLA-B\*15:02 allele. This link was first shown in Han Chinese, followed by Thais, Malaysians and Indians.

Screening for HLA-B\*15:02 is recommended by the US Food and Drug Administration for South East Asians, even with a significantly smaller positive predictive value of 3% to 7.7% and a 100% negative predictive value. However, the HLA-B\*31:01 allele was surprisingly recently found to be associated with CBZ SJS-TEN in Caucasian populations instead of the HLA-B\*15:02 allele. This shows the importance of screening different populations and ethnic groups and finding the HLA-B allele which is best associated with drugs and diseases. The HLA-B\*15:02 allele's involvement is also seen for the Malaysian population, mostly the Malay and Chinese subgroups, and further proves its beneficial use as a genetic marker for SJS-TEN prevention. This is further proven later, where almost all the papers studied for HLA-B allele frequencies showed high HLA-B\*15/15:02 frequencies. The slow but sure integration of HLA-B\*15:02 testing in Malaysia is greatly due to the allele's high frequency in all the ethnic groups. The Malaysian Society of Neurosciences recommends the screening of HLA-B\*15:02 amongst epilepsy patients. However, this still hasn't been made obligatory for all patients taking carbamazepine or included on the drug label as done in other countries (Rani et al., 2018; Then & Raymond, 2019).

### **2.3.5 Pharmacogenomics in Malaysia**

The SCAR cases and trend were in Malaysia analysed in section 2.3.6, with the main focus on MADRAC reports and a few studies in different states. Allopurinol has been mentioned since 2000, up until 2019 in MADRAC reports as the culprit drug in SCAR cases but the cause remained unknown. An alert was only raised in 2010's report to all hospitals nationwide to emphasize on the incorrect prescription of allopurinol. Pharmacogenomics was only mentioned in 2014 MADRAC report, along with a list of HLA variants and gene variants to be on the lookout for along with their culprit drugs and ADRs. Pharmacogenetic testing was also mentioned, with emphasis on warfarin, psychotropics and a testing kit developed for HLA-B\*15:02 in epileptic patients taking carbamazepine. HLA-B\*58:01 was linked to allopurinol-induced SCARs in the following years, but no kit or studies were done on this association (MADRAC bulletin 2014).

Rani et al. developed a High Resolution Melt (HRM) screening technique in UMBI to screen for the HLA-B\*15:02 allele in epileptic patients from Hospital UKM (HUKM). This method has been implemented and used successfully for a few years in preventing SCARs and its related death toll in HUKM (Rani et al., 2018). Pharmacogenetic screening was seen to be hospital specific in Malaysia, where laboratories affiliated to hospitals develop their own screening method.

## 2.3.6 HLA-B\*58:01's role in allopurinol-induced SCARs

### 2.3.6.1 Introduction

The genetic predisposition for allopurinol-induced hypersensitivity reactions has been extensively investigated. The HLA-B\*58:01 allele has been indicated as a genetic susceptibility factor for allopurinol-induced hypersensitivity reactions with studies of allopurinol-induced SCARs strongly associated with individuals tested positive for the HLA-B\*58:01 allele (Hung et al., 2005). To date, the genetic association between HLA-B\*58:01 and allopurinol-induced SCARs has been validated in a number of Asian countries and European countries (Wu et al., 2016).

The association is still a theory, which needs to be established in other countries, like Malaysia. It is postulated that around 10.1% of individuals in the Malaysian population carry this allele (Hung et al., 2005). This will lead to a significant prevalence of allopurinol-induced ADRs. Despite the lack of extensive studies for the allele frequency in Malaysia, a susceptibility-allele frequency of 10.4% is still relatively significant, given the severity of allopurinol-induced ADRs (Gonzalez-Galarza et al., 2015). Such observations may be a result of varying HLA-B\*58:01 frequency in different populations. The HLA-B\*58:01 allele has been reported to occur at the highest frequency among those of Han Chinese descent, but it has also been identified in other population ethnicities. The susceptibility-allele frequency harbors an impact on the practicality of HLA-B\*58:01 genotyping implementation in clinical settings, in such a way that the need for the implementation is directly proportional to the susceptibility-allele frequency in a population (Hung et al., 2005).

This is why our proposal for HLA-B\*58:01 genotyping prior to allopurinol prescription may effectively reduce the incidences of allopurinol-induced ADRs, concomitantly minimizing iatrogenic morbidity, mortality, and extra medical costs among allopurinol-taking patients (Hung et al., 2005).

### 2.3.6.2 HLA-B\*58:01 allele's mechanism of action in SCARs

The HLA-B\*58:01 allele is the most well-established and proven genetic link to allopurinol induced-SCAR. The mechanisms behind this genetic association are shown to follow the HLA-dependent T cell activation by the drug. Allopurinol or oxypurinol bind peptides to produce haptenated products which are processed and presented by HLA-B\*58:01 (hapten concept) as shown in Figure 2.19 below. Both drugs may also directly react with the HLA-B\*58:01-MHC-peptide complex and TCR on a cell's surface without any antigen processing (p-i concept). Proof of the p-i concept comes from the observation that oxypurinol has a higher affinity for the HLA-B\*58:01 molecules in docking experiments. Allopurinol/oxypurinol may also be incrustated in the HLA-B\*58:01 allele's peptide-binding groove, altering the peptide repertoire to be presented by the HLA-B\*58:01 alleles which ultimately results in alloreactivity (McDonagh et al., 2014).

CD8+ T cells activation will ensue, accompanied by cytotoxic effects on keratinocytes which ultimately results in the characteristic epidermal aggression and necrolysis in SJS/TEN. The full mechanism of HLA-B\*58:01 activation of T cells in the presence of allopurinol is shown in Figure 2.20 below (McDonagh et al., 2014).

The allopurinol hypersensitivity(AH) link to the HLA-B\*58:01 allele was first shown in 1989 by Chan and Tan (Jung et al., 2011). Patients risk SCARs development 80-97 times more in the presence of the HLA-B\*58:01 allele (Choo et al., 2014). The risk of AH is shown to be directly proportional to the HLA-B\*58:01 allele frequency in a population. Malaysia is seen to have a high frequency of 10.1%, showing a high risk of allopurinol-induced SCARs (Hung et al., 2005).

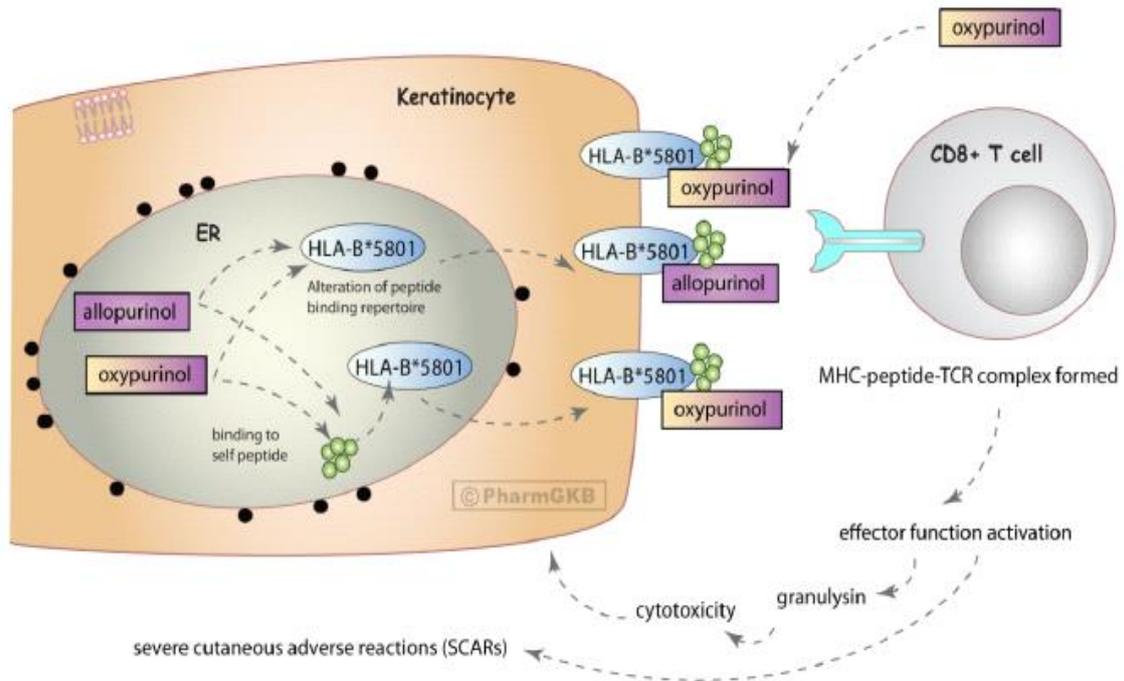


Figure 2-19 This figure shows the potential mechanisms which are involved in the allopurinol-induced SCARs. The haptent concept, p-i concept and the altered repertoire model are depicted. The interaction of allopurinol and oxypurinol with the HLA-B\*58:01 molecule can be clearly seen here (McDonagh et al., 2014).

### 2.3.6.3 Global profiling of the HLA-B\*58:01 allele's role in allopurinol-induced SCARs

#### 2.3.6.3.1 Website search for global HLA-B\*58:01 allelic frequency

A systematic search was performed on The Allele Frequency Net Database, <[www.allelefrequencies.net](http://www.allelefrequencies.net)>, which has 134,008 frequencies stored on HLA alleles from a total of 10,509,338 individuals (Gonzalez-Galarza et al., 2020). Data was extracted by using the keyword “HLA-B\*58:01” for all countries to set up an appropriate global background for this specific allele. The search yielded 281 hits from countries around the whole globe (Gonzalez-Galarza et al., 2020). The highest global allelic frequency for the top fourteen hits and countries were shown in Table 2.7 below.

Table 2-7 Highest global allelic frequency mined for top fourteen hits and countries (Gonzalez-Galarza et al., 2020)

Entry number	Population	Allele frequency	Sample size
16	Cameroon Baka Pygmy	0.1500	10
20	Cameroon Sawa	0.1150	13
30	China Guangdong Province Meizhou Han	0.1700	100
93	India Khandesh Region Pawra	0.1500	50
137	Kenya Nandi	0.1000	240
142	Malaysia Jelebu Temuan	0.1040	25
145	Malaysia Patani	0.1000	25
146	Malaysia Peninsular Chinese	0.1211	194
149	Malaysia Sarawak Bau Bidayuh	0.1400	25
179	Singapore Chinese	0.1040	149
209	Taiwan Hakka	0.1090	55
210	Taiwan Han Chinese	0.1060	504
214	Taiwan pop 2	0.1000	364
215	Taiwan pop 3	0.1010	212

From Table 2.7 above, seven hits (highlighted in yellow) are considered more reliable due to the higher sample size, with a minimum of 100 people. China takes the lead with 17%, followed by Malaysia Peninsular Chinese with 12.11%, Taiwan Han Chinese with 10.60%, Singaporean Chinese with 10.40% and others at 10.00%. Interestingly, Malaysia and Singapore both fall in the highest global HLA-B\*58:01 frequency, focusing mostly on the Chinese ethnic groups. From the whole table, it is noted that five hits correspond to Malaysia and Singapore out of fourteen, i.e, 35.7% of the total. This highlights the importance of further researching all the studies done previously in Malaysia which can contribute to a better and more accurate estimation of the allelic frequency in all its different ethnic groups

#### **2.3.6.3.2 Proof of the strength of association of the HLA-B\*58:01 allele globally**

It is crucial to study the strength of association between the HLA-B\*58:01 allele and allopurinol-induced SCARs in existing studies worldwide in order to grasp the extent to which it can help in minimizing SCARs in hospitals. The allelic frequency of HLA-B\*58:01 in various regions worldwide were searched and adapted from <[www.allelefreqencies.net](http://www.allelefreqencies.net)>, a database on allelic frequencies. A methodical search was carried out on these following databases: PubMed, PreMEDLINE, and NUSearch Primo library to identify studies that investigated the association between HLA-B\*58:01 and allopurinol-induced ADRs. The keywords employed in the search included “allopurinol” AND “HLA-B\*58:01” AND “hypersensitivity” OR “Stevens Johnson Syndrome” OR “Toxic Epidermal Necrolysis” or their acronyms. The results gathered were detailed in Table 2.8 shown on the next few pages.

Hung and colleagues first reported a convincing association between HLA-B\*58:01 and allopurinol-induced SCARs in the Han Chinese population in Taiwan, where 100% (51/51) of the subjects who developed SCARs following allopurinol consumption were found to carry this allele. In contrast, only 14.8% (20/135) of allopurinol-tolerant subjects (OR = 580.3, 95% CI = 34.4 – 9780.9) and 20.4% (19/93) of healthy subjects (OR = 393.5, 95% CI = 23.2 – 6665.26) were shown to carry this allele (Hung et al., 2005). Led by this pioneering

investigation, a handful of other studies validated the genetic association in question in the Han Chinese population. Cao et al. indicated that 100% (38/38) of susceptible subjects enrolled in their study were carriers of the HLA-B\*58:01, whereas only 11.1% (7/63) of the allopurinol-tolerant group (OR = 580.07, 95% CI = 32.18 – 10456.80) and 14.0% (80/572) from the general population (OR = 471.09, 95% CI = 28.66 – 7744.39) were found to carry this allele (Cao et al., 2012). An association of similar magnitude among Han Chinese was reported by two other studies: 100% (19/19) of subjects with SCARs had the allele vs. 13.3% (4/30) of allopurinol-tolerant controls (OR = 229.7, 95% CI = 11.7 – 4520.4) (14); 94.6% (87/92) of subjects suffering from SCARs had been shown to carry the allele while only 12.0% (9/75) of allopurinol-tolerant (OR = 127.6; 95% CI = 40.8 – 398.6) controls and 10.1% (10/99) of healthy subjects (OR = 154.9; 95% CI = 50.9 – 471.5) had been tested positive for the allele (Cao et al., 2012).

This association between HLA-B\*58:01 allele and allopurinol-induced ADRs was also reported in various regions of Asia. Studies conducted on native Thais or Thai-Chinese who developed SCARs post-allopurinol use demonstrated a solid genetic association between the allele and SCARs. Tassaneeyakul and colleagues revealed that all of the susceptible patients (27/27) in their studies were carriers of HLA-B\*58:01, while a mere 13.0% (7/54) of allopurinol-tolerant controls (OR = 348.3, 95% CI = 19.2 – 6336.9) and 8.1% (8/99) of healthy controls (OR = 592.1, 95% CI = 33.1 – 10587.7) were carriers of said allele (Tassaneeyakul et al., 2009). In an association study undertaken by Jantararoungtong et al, 100% (12/12) of subjects who developed allopurinol-induced SCARs had been shown to carry the allele, while 50.0% (1/2) of subjects who developed MPE had been seen to be carriers of the allele. In contrast, only a mere 10.7% (6/56) of allopurinol-tolerant patients (OR = 108.33, 95% CI = 11.96 – 980.82) were found to carry the HLA-B\*58:01 allele (Jantararoungtong et al., 2011). This finding was further supported by a multicentre case-control study carried out by Sukasem and coworkers, where the HLA-B\*58:01 allele was found in 100% (23/23) of subjects who developed SCARs following allopurinol consumption (Sukasem et al., 2016). *Sukasem et al.* also reported a HLA-B\*58:01 frequency of 85.7% (6/7) in patients who developed MPE (a form of mild adverse drug reaction related to allopurinol), validating an association not only between HLA-B\*58:01 and allopurinol-induced SCARs, also that between HLA-B\*58:01 and allopurinol-induced cADRs. In comparison, only 4.0% (4/100) of

allopurinol-tolerant patients (OR = 696.00, 95% CI = 74.81 – 6475.01) and 10.1% (111/1095) of healthy subjects (OR = 257.08, 95% CI = 34.68 – 1905.57) were reported to carry this allele (Sukasem et al., 2016).

This association was similarly observed in the Korean population. Kang et al. indicated that as many as 92.3% (24/26) of patients who suffered from allopurinol-induced SCARs had the HLA-B\*58:01 allele while merely 10.5% (6/57) of allopurinol-tolerant patients (OR = 97.8; 95% CI = 18.3 – 521.5) and 12.2% (59/485) of healthy controls (OR = 83.0, 95% CI = 19.0 – 361.3) were shown to be carriers of HLA-B\*58:01 (Kang et al., 2016). A study conducted by Jung et al. in 2011 further validated this association among the Korean population. 100% (9/9) of SCARs patients were shown to be HLA-B\*58:01 carriers, in contrast to susceptibility-allele frequencies of 9.5% (41/432) and 12.2% (59/485) in allopurinol tolerant (OR = 179.24, 95% CI = 10.19 – 3151.74) and healthy controls (OR = 136.29, 95% CI = 7.79 – 2380.85), respectively (Jung et al., 2011). Despite its notably lower frequency in the Japanese population, the HLA-B\*58:01 allele has still been found to be significantly associated with allopurinol-induced ADRs. Kaniwa et al. reported that 20% (4/20) of SCARs patient carried the allele whilst only 0.6% (6/986) of healthy controls (OR = 40.83, 95% CI = 10.5 – 158.9) carried HLA-B\*58:01 (Kaniwa et al., 2008). Tohkin et al. had shown that among 36 SJS/TEN patients, only 10 of them were found to carry this allele (27.8%) (Tohkin et al., 2013). In contrast, only 0.6% of the general population was tested positive for this allele (OR = 62.8, 95% CI = 21.2 – 185.8) Another Japanese study<sup>18</sup> reported that 2 out of 3 SJS patients were carriers of the HLA-B\*58:01 allele (OR = 65.6, 95% CI = 2.9 – 1497.0). However, the Japanese population has been described to face a lower risk for allopurinol-induced ADRs as compared to other Asian populations. HLA-B\*58:01 association with allopurinol-induced adverse drug reactions was also replicated – although at a more modest strength – in European populations. Cristallo et al indicated that 42.8% (19/31) of SCAR subjects were carriers of the HLA-B\*58:01 allele while only 5.2% (6/115) of the healthy controls (OR = 13.625, 95% CI = 2.774 – 69.448) recruited in the study carried the allele (Cristallo et al., 2011). In a Caucasian population, 61.0% (19/31) of SCARs patients had been shown to carry the allele (Wu et al., 2016). A study conducted by Lonjou et al. among a Caucasian population reported a HLA-B\*58:01 frequency of 61% (19/31) among those susceptible to

allopurinol-induced SCARs. Only 1.5% (28/1822) of the general population had been reported to carry this allele (OR = 61.0, 95% CI = 32 – 118) (Lonjou et al., 2008).

In the 14 studies aforementioned, the HLA-B\*58:01 allele has been described as a susceptibility factor for allopurinol-induced adverse drug reactions (both mild and severe forms). In other words, a significant association between HLA-B\*58:01 and allopurinol-induced ADRs has been observed in the populations aforementioned, with the exception of the predisposition to simple rash, as reported by Jung et al. in their study (Wu et al., 2016).

Table 2-8 Summary of studies reporting the association of HLA-B\*58:01 and allopurinol-induced ADRs (including cADRs and SCARs) worldwide. CI: Confidence interval; ADRs: Adverse drug reactions; SCARs: Severe cutaneous adverse drug reactions; cADRs: Cutaneous adverse drug reactions; SJS: Stevens-Johnson syndrome; TEN: Toxic epidermal necrolysis (Wu et al., 2016; Chang et al., 2020)

Study and year	Group of patients suffering from types of ADRs	Frequency of HLA-B*58:01 allele (%)	Odds ratio of susceptible group vs. the allopurinol-tolerant group (95% CI)	p-value	Odds ratio of susceptible group vs. the general population (95% CI)	p-value
<b>Han Chinese population</b>						
<b>Hung et al.</b>	SCARs	100.0 (51/51)	580.3 (34.4 – 9780.9)	$4.7 \times 10^{-24}$	393.5 (23.2 – 6665.26)	$8.1 \times 10^{-18}$
	Allopurinol tolerant	14.8 (20/135)	-	-	-	-
	General population	20.4 (19/93)	-	-	-	-
<b>Cao et al.</b>	cADRs	100.0 (38/38)	580.07 (32.18 – 10456.80)	$7.01 \times 10^{-18}$	471.09 (28.66 – 7744.39)	$3.15 \times 10^{-38}$
	SCARs	100.0 (16/16)	248.60 (13.48 – 4585.35)	$7.40 \times 10^{-12}$	201.89 (11.99 – 3398.49)	$1.82 \times 10^{-18}$
	SJS/TEN	100.0 (13/13)	203.40 (10.93 – 3785.04)	$8.24 \times 10^{-10}$	165.19 (9.72 – 2806.13)	$2.10 \times 10^{-15}$
	DRESS	100.0 (3/3)	52.73 (2.47 – 1124.13)	$2.00 \times 10^{-3}$	42.83 (2.19 – )	$2.00 \times 10^{-3}$

					836.89)	
	MPE	100.0 (22/22)	339.00 (18.58 – 6186.39)	$9.21 \times 10^{-14}$	275.31 (16.54 – 4583.53)	$3.74 \times 10^{-24}$
	Allopurinol tolerant	11.1 (7/63)	-	-	-	-
	General population	14.0 (80/572)	-	-	-	-
<b>Chiu et al.</b>	cADRs	95.0 (19/20)	123.5 <sup>1</sup> (12.8 – 1195.1) 151.3 <sup>2</sup> (10.8-2115.5)	$1 < 1 \times 10^{-4}$ $2 \times 2 \times 10^{-4}$	-	-
	SCARs	100.0 (19/19)	229.7 (11.7 – 4520.4)	-	-	-
	Allopurinol tolerant	13.3 (4/30)	-	-	-	-
<b>Cheng et al.</b>	SCARs	94.6 (87/92)	127.6 (40.8 – 398.6)	$< 1 \times 10^{-3}$	154.9 (50.9 – 471.5)	$< 1 \times 10^{-3}$
	Allopurinol tolerant	12.0 (9/75)	-	-	-	-
	General population	10.1 (10/99)	-	-	-	-

Table 2-8 (continued): Summary of studies reporting the association of HLA-B\*58:01 and allopurinol-induced ADRs (including cADRs and SCARs) worldwide. CI: Confidence interval; ADRs: Adverse drug reactions; SCARs: Severe cutaneous adverse drug reactions; cADRs: Cutaneous adverse drug reactions; SJS: Stevens-Johnson syndrome; TEN: Toxic epidermal necrolysis (Wu et al., 2016; Chang et al., 2020)

Study and year	Group of patients suffering from types of ADRs	Frequency of HLA-B*58:01 allele (%)	Odds ratio of susceptible group vs. the allopurinol-tolerant group (95% CI)	p-value	Odds ratio of susceptible group vs. the general population (95% CI)	p-value
<b>Japanese population</b>						
<b>Kaniwa et al.</b>	SJS/TEN	20.0 (4/20)	-	-	40.83 (10.5 – 158.9)	< 1x10 <sup>-4</sup>
	General population	0.6 (6/986)	-	-	-	-
<b>Tohkin et al.</b>	SJS/TEN	27.8 (10/36)	-	-	-	-
	General population	0.6	-	-	62.8 (21.2 – 185.8)	5.4x10 <sup>-12</sup>
<b>Niihara et al.</b>	SJS	66.7 (2/3)	65.6 (2.9 – 1497.0)	9.73x10 <sup>-4</sup>	-	-
	cADRs	50.0 (2/4)	26.0 (2.0 – 336.1)	-	-	-
	Allopurinol tolerant	-	-	-	-	-

Korean population						
<b>Jung et al.</b>	SCARs	100.0 (9/9)	179.24 (10.19 – 3151.74)	$< 1 \times 10^{-3}$	136.29 (7.79 – 2380.85)	$< 1 \times 10^{-3}$
	Rash	0.0 (0/7)	0.63 (0.005 – 5.32)	0.9	0.48 (0.03 – 8.67)	0.9
	Allopurinol tolerant	9.5 (41/432)	-	-	-	-
	General population	12.2 (59/485)	-	-	-	-
<b>Kang et al.</b>	SCARs	92.3 (24/26)	97.8 (18.3 – 521.5)	$2.45 \times 10^{-11}$	83.0 (19.0 – 361.3)	$2.47 \times 10^{-16}$
	SJS/TEN	80.0 (4/5)	34.0 (3.2 – 356.1)	$1.60 \times 10^{-2}$	28.9 (3.2 – 263.4)	$1.00 \times 10^{-2}$
	DIHS	95.2 (20/21)	161.5 (18.2 – 1430.9)	$1.45 \times 10^{-10}$	137.5 (18.1 – 1046.2)	$2.83 \times 10^{-14}$
	Allopurinol tolerant	10.5 (6/57)	-	-	-	-
	General population	12.2 (59/485)	-	-	-	-

Table 2-8 (continued): Summary of studies reporting the association of HLA-B\*58:01 and allopurinol-induced ADRs (including cADRs and SCARs) worldwide. CI: Confidence interval; ADRs: Adverse drug reactions; SCARs: Severe cutaneous adverse drug reactions; cADRs: Cutaneous adverse drug reactions; SJS: Stevens-Johnson syndrome; TEN: Toxic epidermal necrolysis (Wu et al., 2016; Chang et al., 2020)

Study and year	Group of patients suffering from types of ADRs	Frequency of HLA-B*58:01 allele (%)	Odds ratio of susceptible group vs. the allopurinol-tolerant group (95% CI)	p-value	Odds ratio of susceptible group vs. the general population (95% CI)	p-value
<b>Thai population</b>						
<b>Tassaneeyakul et al.</b>	SCARs	100.0 (27/27)	348.3 (19.2 – 6336.9)	$1.61 \times 10^{-13}$	592.1 (33.1 – 10587.7)	$1.00 \times 10^{-20}$
	Allopurinol tolerant	13.0 (7/54)	-	-	-	-
	General Population	8.1 (8/99)	-	-	-	-
<b>Jantararoungtong et al.</b>	SCARs	92.8 (13/14)	108.33 (11.96 – 980.82)	$< 10^{-6}$	-	-
	SJS/TEN	100.0 (9/9)	217.26 (12.41 – 925.35)	$< 10^{-6}$	-	-

	DRESS	100.0 (3/3)	80.0 (3.42 – 372.87)	$< 10^{-6}$	-	-
	MPE	50.0 (1/2)	-	-	-	-
	Allopurinol tolerant	10.7 (6/56)	-	-	-	-
<b>Sukasem et al.</b>	cADRs	96.7 (29/30)	696.00 (74.81 – 6475.01)	$< 10^{-3}$	257.08 (34.68 – 1905.57)	$< 10^{-3}$
	SJS/TEN	100.0 (13/13)	579.00 (29.50 – 11362.67)	$< 10^{-3}$	238.40 (14.08 – 4037.80)	$< 10^{-3}$
	DRESS	100.0 (10/10)	430.33 (22.64 – 8958.88)	$< 10^{-3}$	185.42 (10.79 – 3185.84)	$< 10^{-3}$
	MPE	85.7 (6/7)	144.00 (13.85 – 1497.03)	$< 10^{-3}$	53.19 (6.35 – 445.85)	$< 10^{-3}$
	Allopurinol tolerant	4.0 (4/100)	-	-	-	-
	General population	10.1 (111/1095)	-	-	-	-

Table 2-8 (continued): Summary of studies reporting the association of HLA-B\*58:01 and allopurinol-induced ADRs (including cADRs and SCARs) worldwide. CI: Confidence interval; ADRs: Adverse drug reactions; SCARs: Severe cutaneous adverse drug reactions; cADRs: Cutaneous adverse drug reactions; SJS: Stevens-Johnson syndrome; TEN: Toxic epidermal necrolysis (Wu et al., 2016; Chang et al., 2020)

Study and year	Group of patients suffering from types of ADRs	Frequency of HLA-B*58:01 allele (%)	Odds ratio of susceptible group vs. the allopurinol-tolerant group (95% CI)	p-value	Odds ratio of susceptible group vs. the general population (95% CI)	p-value
<b>European population</b>						
<b>Lonjou et al.</b>	SJS/TEN	61.0 (19/31)	-	-	61.0 (32 – 118)	< 10 <sup>-8</sup>
	General Population	1.5 (28/1822)	-	-	-	-
<b>Cristallo et al.</b>	SCAR	42.8 (3/7)	-	-	13.625 (2.774 – 69.448)	0.248
	General population	5.2 (6/115)	-	-	-	-

#### 2.3.6.4 Data mining for the presence of HLA-B\*58:01 allele in Malaysia

A systematic search was performed on The Allele Frequency Net Database, <[www.allelefrequencies.net](http://www.allelefrequencies.net)>, which has 134,008 frequencies stored on HLA alleles from a total of 10,509,338 individuals (Gonzalez-Galarza et al., 2020).

A classical allele frequency search was done for the HLA-B\*58:01 allele in Malaysia. Data was also mined for all other countries to set up an appropriate global background for the allele. Using the keywords “HLA-B\*58”, “Malaysia”, a focused search was first done for these two countries of interest. “HLA-B\*58” was used instead of HLA-B\*58:01 in order to widen the number of hits, frequencies and ethnic groups to be examined, as there are a limited amount of studies done in Malaysia. The search for Malaysia specifically yielded 18 different frequency hits, with 9 hits containing the HLA-B\*58:01, as shown in Table 2.9 below. The sample size varied from a minimum of 25 to 1445 and the HLA-B\*58:01 allele frequency varied from 14% to 1.9% across all the 16 different Malaysian ethnic groups.

Table 2-9 Simplified search results from the database for Malaysia HLA-B\*58 and HLA-B\*58:01 hits (Gonzalez-Galarza et al., 2020)

Hit Number	Allele	Ethnic group	Percentage of individuals with the allele (%)	Allele frequency (%)	Sample size
1	B*58	Negeri Sembilan Minangkabau	-	3.00	34
2	B*58	Perak and Johor Banjar Bugis Jawa	-	3.20	94
3	B*58	Perak Rawa	-	7.00	23
4	B*58	Malaysia population 2	-	4.90	62
5	B*58	Malaysia population 3	-	5.80	1445
6	B*58	Sabah Kadazan	-	6.10	57
7	B*58	Sarawak Bidayuh	-	2.00	50
8	B*58	Sarawak Iban	-	5.90	51
9	B*58:01	Champa	10.3	5.20	29
10	B*58:01	Jelevu Temuan	-	10.40	25
11	B*58:01	Kelantan	7.1	3.60	28
12	B*58:01	Mandailing	3.7	1.90	27
13	B*58:01	Patani	20.0	10.00	25
14	B*58:01	Peninsular Chinese	22.2	12.11	194
15	B*58:01	Peninsular Malay	11.3	5.84	951
16	B*58:01	Perak Grik Jehai	-	6.00	25
17	B*58:01	Sarawak Bau Bidayuh	-	14.00	25
18	B*58:19	Peninsular Malay	0.1	0.053	951

The demographics of Malaysia, with its numerous ethnic groups, adds another level of complexity in understanding the role of HLA-B alleles and defining rock solid frequencies. The three main ethnic groups are the Malay/Bumiputera making up 68.8% of the population, the Chinese comprising of 23.2% and Indians with 7.0% (Negeri, 2015). The different ethnic groups in Malaysia can be classified as ethnolinguistic groups, with a total of around 27 large groups, as shown in Table 2.10 below (The Joshua Project, *joshuaproject.net*, retrieved 2015).

Table 2-10 Detailed list of all the ethnolinguistic groups of Malaysia (The Joshua Project, *joshuaproject.net*, retrieved 2015)

<b>Ethnolinguistic group</b>	<b>Population in numbers</b>
Malay, standard Malaysian	6,916,000
Malay, Kedah	3,095,000
Malay, East coast	2,448,000
Han Chinese, Hokkien	1,903,000
Tamil	1,796,000
Han Chinese, Hakka	1,729,000
Han Chinese, Cantonese	1,396,000
Han Chinese, Teochew	1,004,000
Han Chinese, Mandarin	986,000
Minangkabau	901,000
Indonesian	796,000
Iban	686,000
Javanese	640,000
Filipino, Tagalog	456,000
Han Chinese, Hainanese	396,000
Han Chinese, Northern Min	384,000
Brunei Malay, Kedayan	350,000
Malay, East Malaysia	280,000
Han Chinese, Eastern Min	256,000
Straits Chinese	244,000

Nepalese	224,000
Tausug	209,000
Dusun, Central	197,000
Malayali, Malayalam	170,000
Bugis, Buginese	139,000
Lun Bawang/Lundayeh	31,600
Kelabit people	5,000

Peninsular Malaysia holds around 79% of the whole Malaysian population when compared to East Malaysia. The HLA-B\*58:01 frequency of 12.11% seen in Table 3 for Peninsular Chinese (hit number 14) follows the trend seen for Chinese population from China, shown in Table 2.8 from the global profiling section. Peninsular Malays come in second with a HLA-B\*58:01 frequency of 5.84% (hit number 15). No frequency was listed in Table 3 for Indian Malaysians, but this should follow the general Indian population, which has a lower HLA-B\*58:01 frequency of around 1.00% as shown in The Allele Frequency Net Database.

Apart from HLA-B\*58:01, a handful of HLA variants have been discovered to be linked with the propensity for allopurinol-induced cADRs, namely the HLA-A\*33:03 and HLA-C\*03:02 alleles, shown in **Table 5**. In spite of that, these associations between the alleles aforementioned in allopurinol-induced SCARs are highly possible to be brought about by linkage disequilibrium (LD) with HLA-B\*58:01 (Wu et al., 2016; Chang et al., 2020)

The link between these three different alleles can be clearly seen with their similar percentages in allelic frequency in Table 2.11. These three alleles are said to be in linkage disequilibrium, hence showing the importance of screening for all three of them to give a better chance at improving screening efficiency. In a future world where HLA allele screening will be done routinely, screening for these three alleles before allopurinol treatment will only help in improving treatment for patients.

Methods employed by the seven papers from Table 2.11 are as follows; sequence-specific Oligonucleotide probes (SSOP), sequence specific primer (SSP), sequence based typing (SBT). These methods are considered outdated nowadays, especially with advances in genetic technologies like Sanger sequencing and Next Generation Sequencing. Future studies can

fully exploit NGS in order to provide higher resolution data and more accurate allelic frequencies.

The data gathered were mostly from anthropology studies, blood donations, bone marrow registries and controls from disease studies. There should be future specific studies done on gout patients taking allopurinol, with a significant sample size to truly assess the role of the HLA-B\*58:01 allele in all the different ethnic groups in Malaysia. Moreover, patients experiencing different levels of ADRs and SCARs need to be identified, along with the culprit allele in the majority of cases. The focus should first start on the three major ethnic groups in Peninsular Malaysia, followed by all other smaller ethnic groups. In this way, a more comprehensive study and analysis will be done on the role of the HLA-B\*58:01 allele in the Malaysian population (Gonzalez-Galarza et al., 2020).

Another possible method to help identify at risk population may be to trace the ancestors of all the different ethnic groups and classify them in high risk to low risk groups in hospitals. One example would be; Malaysian Indians will automatically be classified as a low risk group as their gene pool will follow the same trend as their ancestors who have a low HLA-B\*58:01 frequency. However, this method should be used only in initial groupings of patients as mixed population will result in different allelic frequencies. Doctors need to further check the patient's medical record to define their ethnic groups properly. This method should be used to classify and prioritize patients' treatment and wellbeing and in no case should lead to discrimination amongst the different ethnic groups. With the global approach on pharmacogenomics and the increasing association and discovery of HLA allele's role in ADRs and diseases, this type of discrimination will be erased. Patients will be screened for all their HLA alleles and be given medicines based on this and their susceptibility to ADRs and SCARs for a wide range of diseases.

Pharmacogenomics shouldn't be perceived only as the need to do genetic screening before drug prescription. It should be considered as an integrated plethora of factors, which will each contribute in reducing SCARs/ADRs for any diseases. Understanding and integrating all these small factors' role in diseases and patient treatment will lead to the ultimate personalized medicine era.

Table 2-11 Full table for all studies mined for the Malaysian population. (SSOP-sequence-specific oligonucleotide probes, SSP-sequence specific primer, SBT-sequence based typing) (adapted from Gonzalez-Galarza et al., 2020).

Study number	Alleles	Percentage of individuals with allele (%)	Allele Frequency (%)	Sample Size (n=)	Ethnicity	Ethnic origin	Method used	Source
1	HLA-B*58:01	22.2	12.11	194	Malaysia Peninsular Chinese	-	SSOP	Controls for Disease Study
	HLA-A*33:03	4.1	2.84					
	HLA-C*03:02	24.7	12.6290					
3	HLA-A*33:03	7.8	4.1535	951	Malaysia Peninsular Malay	Asian	SSOP	Control for diseases study
	HLA-B*58:01	11.3	5.8360					
	HLA-C*03:02	12.0	6.1514					
4	HLA-A*33:03	-	16.70	25	Malaysia Jelebu Temuan	Oriental	SSP	Anthropology study
	HLA-B*58:01	-	10.40					
	HLA-C*03:02	-	-					
5	HLA-A*33:03	36.0	18.0000	25	Patani Malaysia	Austronesian	SBT	Anthropology study
	HLA-B*58:01	20.0	10.0000					

6	HLA-A*33	-	3.50	57	Malaysia Sabah Kadazan	Austronesian	SSP	Anthropology study
	HLA-B*58	-	6.10					
	HLA-C*03:02	-	-					
7	HLA-A*33:03	7.4	3.7000	27	Mandailing Malaysia	Austronesian	SBT	Anthropology study
	HLA-B*58:01	3.7	1.9000					
	HLA-C*03:02	-	-					
8	HLA-A*33:03	28.6	14.3000	28	Kelantan Malaysia	Austronesian	SBT	Anthropology study
	HLA-B*58:01	7.1	3.6000					
	HLA-C*03:02	-	-					
9	HLA-A*33:03	24.1	12.1000	29	Champa Malaysia	Austronesian	SBT	Anthropology study
	HLA-B*58:01	10.3	5.2000					
	HLA-C*03:02	-	-					
10	HLA-A*33	-	7.50	94	Malaysia Perak and Johor Banjar Burgis Jawa	Oriental	SSP	Anthropology study
	HLA-B*58	-	3.20					
	HLA-C*03	-	17.60					
11	HLA-A*33	-	13.00	23	Malaysia Perak rawa	oriental	SSP	Anthropology study
	HLA-B*58:01	-	7.00					
	HLA-C*03:02	-	30.00					
12	HLA-A*33	-	12.90	1445	Malaysia Pop 3	austronesian	SSP	Bone marrow

	HLA-B*58	-	5.80					registry
	HLA-C*03	-	-					
13	HLA-A*33	-	4.90	51	Malaysia Sarawak Iban	austronesian	SSP	Anthropology study
	HLA-B*58	-	5.90					
	HLA-C*03	-	-					
14	HLA-A*33	-	4.00	50	Malaysia Sarawak Bidayuh	austronesian	SSP	anthropology
	HLA-B*58:01	-	2.00					
	HLA-C*03:02	-	-					
15	HLA-A*33:03	-	12.00	25	Malaysia Sarawak Bau Bidayuh	austronesian	SSP	anthropology
	HLA-B*58:01	-	14.00					
	HLA-C*03:02	-	-					
16	HLA-A*33:03	-	8.00	25	Malaysia Perak Grik Jehai	oriental	SSP	anthropology
	HLA-B*58:01	-	6.00					
	HLA-C*03:02	-	-					

## 2.4 Existing methods for HLA allele screening

### 2.4.1 General HLA allele screening methods

The first method in history to identify HLA alleles was devised in 1964, where lymphocytes are used to perform serological testing also known as lymphocytotoxicity. This technique uses the high expression of HLA antigens on lymphocytes to type for HLA Class I and II antigens. However, due to its disadvantageous cross-reactivity and limited number of antibodies, this method gave way to newer DNA centred techniques. Serological typing is labour intensive, with a high complexity and low resolving power where HLA-B alleles can have the same serological specificity (Shankarkumar, 2004).

A type of Mixed lymphocyte culture (MLC) was also developed, where the culturing of lymphocytes from two people led to a wider range of foreign HLA antigens being identified. Molecular genetic methods saw the light in 1989 with Restriction Fragment Length Polymorphism (RFLP) which uses enzymes to cut DNA sequences at specific DNA locations. This will generate different fragment sizes for different alleles cut by the same enzyme. RFLP has been long abandoned as it is slow, inefficient and needs a large amount of DNA. The Polymerase Chain Reaction (PCR) allows a specific part of any DNA sequence to be targeted, amplified and analysed. Sequence specific priming (SSP), a PCR-based method, allows amplification of DNA with specific primers which are complementary to distinct HLA alleles or antigens. Another PCR-based method named Sequence Specific Oligonucleotide Typing (SSO) amplifies a whole gene which is latter mixed with radioactive labelled probes specific to particular HLA antigens. SSOP and SSP require a high amount of reagents in numerous reactions and these methods can be hindered by the high level of sequence homology in class I genes. Detection of new variants by SSOP is restricted to alleles formed by one combination event, whilst SSP can only detect variants in cis conformation compared to primers used. A loop-mediated isothermal amplification (LAMP) experiment can be used to identify HLA alleles, but the latter lacks in specificity compared to a standard or multiplex PCR. The gold standard for HLA typing is DNA sequencing, but it is too time consuming, especially when immediate results are needed in hospitals for subsequent prescriptions. (Choo et al., 2007).

## 2.4.2 HLA-B\*58:01 specific screening methods

A handful of specific HLA-B\*58:01 screening methods and kits have been produced globally in the last decade. Real-time PCR has emerged as the new promising method in order to rapidly and quickly screen for the presence of HLA-B alleles in patients. Commercially available PCR-based methods, such as PCR-SSP, PCR-SSOP and PCR-SBT, are conventional tools for multiple HLA alleles typing as shown in Table 2.12 below. Laborious operation, large capital costs for equipment, long turn-round time and easy contamination in post-PCR handling hamper the wide application of these aforementioned methods in clinical settings (Choo et al., 2007).

Table 2-12 Existing HLA-B\*58:01 screening methods with their specific machines/kit used, cost, run time and type of analysis required.

HLA-B*58:01 Screening Method	Machine/kit used	Cost per sample (USD \$)	Run time and results output	Analysis required	Reference
Single-tube duplex real-time PCR assay	ViiATM 7 real-time PCR instrument	\$ 1.5	~2-3 hrs	ViiATM 7 machine software analysis	Kang et al. 2016
PG5801 DNA detection kit	Real-Time PCR machine	\$ 15	~3-4 hrs	Ct analysis	Ko et al. 2015
Loop-mediated isothermal amplification (LAMP)	Thermal cyclers	\$ 6.4	~2 hrs	Visualization of results	Kwok & Kwong, 2013
SSO	Thermal cyclers	\$ 32	~5 hrs 10 mins	Visualization of results	Kwok & Kwong, 2013

<b>SSP</b>	Thermal cycler	\$ 38.5	~2-3 hrs	Visualization of results	Kwok & Kwong, 2013
<b>SYBR Green Real Time PCR</b>	Thermal cycler	\$ 3.8	~1 hr 5 mins	Rotor gene v6 software Qiagen	(Nguyen et al., 2017)
<b>TaqMan assay</b>	ABI 7500 real-time PCR system (Applied Biosystems, USA).	\$ 2.0	~2 hr 35 min	Measuring fluorescence	Zhang et al. , 2004
<b>Dna BINDING Dye-based Duplex Allele-specific Melting Curve Analysis</b>	CFX96 Touch Real time PCR detection system (Biorad)	\$ 1.25	~2hrs mins	Visualization/qu alitative results	Huh et al. 2018

Kang et al. performed genotyping of HLA-B\*58:01 in four different, healthy Chinese populations (n=349) and was 100% validated by sequence-based typing (SBT). The single-tube duplex real-time PCR assay developed combined sequence-specific primers and a TaqMan probe, in order to increase the specificity of differentiating between the HLA-B\*58:01 allele and other homologous allele. The latter has a cost of \$1.5, a run time of 1 hour for 96 samples and analysis was done on its system itself. Another good point from this method is the use of three DNA standards, with two HLA-B\*58:01 positive (homozygote and heterozygote) and one HLA-B\*58:01 DNA negative sample. Furthermore, the target and reference gene was amplified together in one tube, compared to Zhang et al's method described below. This study is an upgrade from Zhang et al's method, in terms of number of tubes and primers. However, despite all the precautions, other HLA-B alleles such as 57:05, 58:05, 58:13, 58:24, 58:31 and 58:40 were still amplified (Kang et al., 2016).

Ko et al. screened for HLA-B\*58:01 for 6 years in Taiwanese people of Han Chinese descent (n=2910), who had not taken allopurinol yet and completely prevented SCARs from happening. The PG5801 DNA detection kit was developed by Pharmigene in Taiwan and is based on a real-time PCR method. It has a cost of around USD15 per test, an assay run time of 2 hours and results in this study were verified by reverse line blot with an HLA sequence oligonucleotide. However, this method used only the cycle threshold values (Ct) analysis to interpret the final results and thus can be deemed inadequate. Relying only on Ct values is quite risky for the screening of HLA-B alleles, which have highly similar DNA sequences and high polymorphism (Ko et al., 2015).

Kwok and Kwong developed a loop-mediated isothermal amplification (LAMP) assay to detect the HLA-B\*58:01 allele, by targeting exons 2 and 3. Twenty gout samples, with known genotype (HLA-B\*58:01 positive) were used for the LAMP assay and results were compared to a clinical genotyping method. However, this method has a higher cost, along with complicated primer design and result visualization by experts. LAMP assay has a lower cost and run time compared to the conventional SSO and SSP methods as shown in Table 2.12 above (Kwok & Kwong, 2013).

Nguyen et al. used a SYBR green real-time PCR method to screen for the HLA-B\*58:01 allele in 119 samples. This method's sensitivity, specificity, positive predictive value and negative predictive value were all 100% and results were compared to Luminex SSO/SBT/SSP. The region between exon 2 and 3 was targeted and a housekeeping gene actin (ACTB) was used as internal control. Detection of HLA-B\*58:01 was based on double peaks for positive samples (91.26°C and 80.97°C) and a single peak (81.42°C) for negative samples. This method was used by the Korean authors Huh et al. as shown below (Huh et al., 2018). Authors proved that it was more accurate to detect the HLA-B\*58:01 allele by using the melt curve, rather than Ct values or agarose gel visualization. Moreover, this qualitative method had a low cost of USD \$3.8, lower DNA detection limit of 0.8 ng/μl and a run time of around an hour as shown in Table 2.12 above (Nguyen et al., 2017).

Zhang et al. screened for the HLA-B\*58:01 allele by combining specific primers and TaqMan probes in a TaqMan assay. Primers had a mismatch introduced for better allelic discrimination, along with two probes targeting exon 2 and intron 2 regions. Samples included allopurinol-induced SCARs (48), allopurinol-tolerant (133) and healthy individuals

(280). However, this reaction used two different tubes and primers cannot distinguish HLA-B\*58:01 from HLA-B\*57:05 and 22 other HLA-B\*58 alleles (B\*58:04, B\*58:05, B\*58:09, B\*58:11, B\*58:12, B\*58:13, B\*58:15, B\*58:17, B\*58:19, B\*58:21, B\*58:22, B\*58:23, B\*58:24, B\*58:28, B\*58:31N, B\*58:33, B\*58:34, B\*58:35, B\*58:36, B\*58:37, B\*58:39N, and B\*58:40). This may cause non-specific and false positive results, even though these alleles are considered to be rare in Asian populations (Zhang et al. , 2004).

Ji et al. developed a DNA binding dye-based duplex allele-specific melting curve analysis method, which costs around USD 1.25 and has an assay run time of around 1 hour 15 minutes. Primers used had a mismatch which allowed them to identify positive samples (generates 2 peaks) as well as negative samples (one peak). The BRAF gene, a proto-oncogene for the B-Raf protein governing cell growth, was used as internal control. 150 total samples were used and results were compared against SBT results. However, authors used previously genotyped samples from umbilical cord blood samples, showing that there was no link to allopurinol use, SCARs presence and gout. Primers used could not discriminate the HLA-B\*58:01 allele from the HLA-B\*58:02, HLA-B\*58:31N and HLA-B\*58:76, three alleles that are in linkage disequilibrium with the HLA-B\*58:01 allele (Huh et al., 2018).

Real-time PCR has therefore emerged as an important tool for HLA-B\*58:01 genotyping because of its rapidity, sensitivity and flexibility of use (Kang et al. 2016). However, due to the high polymorphism and linkage disequilibrium, it is very important to design a method which checks all generated data, not only Ct values or number of peaks seen at specific temperatures. A more specific way of making sure the real-time PCR method works, would be to use all of the aforementioned results and check any pattern seen. The most important part is to have specific primers, identifying only the HLA-B\*58:01 allele by using SNPs, so as to avoid the presence of any other closely related HLA-B alleles.

The few HLA laboratory diagnostic tests available in Malaysia are very costly and time-consuming for clinical settings. The laboratory turn-around time (LTAT) of HLA Typing (Class I and II) for disease association test in the Institute for Medical Research (IMR) is 10 working days (Chong et al., 2018). The LTAT is calculated from a day after a sample received until the report is sent out to hospital (in working days). According to the Laboratory Service Guide from Department of Diagnostic Laboratory Services Universiti Kebangsaan Malaysia Medical

Centre, the LTAT is equal to 15 working days (Chong et al., 2018). Meanwhile, the LTAT of Pathology Service in the Hospital Ampang is 30 working days. The guidebook of Immunology Lab Service from the Hospital Universiti Sains Malaysia states that the LTAT for the test is 120 days (Chong et al., 2018). This causes a serious problem for patients who require immediate treatment and drug prescription in Malaysia.

The efficacy of pharmacogenetic testing highly depends on numerous factors such as sensitivity, specificity, accuracy, reproducibility and simplicity of testing. All these factors must be present in any pharmacogenetic screening method in order to fully exploit the latter's advantages and feasibility. The current laboratory tests available in Malaysia do not meet these required standards. Hence, the need arises for a screening method which can combine the use of real-time PCR, melt curve reading, Ct value use and peaks analysis as aforementioned.

## 2.5 Cost-effectiveness analysis of HLA-B\*58:01 screening

### 2.5.1 Introduction

An important factor to consider before developing a screening method, is its probable cost-effectiveness in a specific country. This will thus be a guide for the acceptance and implementation of HLA-B\*58:01 screening in a country. Cost-effective analysis (CEA) of HLA-B\*58:01 screening was done in numerous countries and still remains debateable to this day. All the countries with an established HLA-B\*58:01-allopurinol association and with screening possibilities showed that screening before drug prescription was cost-effective. HLA-B\*58:01 screening has been recommended by the FDA, American College of Rheumatology and other countries.

Specific HLA-B\*58:01 CEA studies were analysed from Korea, Taiwan, Thailand, UK, USA, Singapore and Malaysia in this section. Positive results were obtained in Korea, Taiwan, Thailand and for two ethnic groups in the USA. Negative results were reported in the UK, Singapore and Malaysia, but variation of a few factors could always change the results to be positive. However, common points for all the aforementioned studies are the lack of randomized controlled trials (RCTs) and large scale studies to confirm the HLA-B\*58:01 frequency in countries. More extensive studies are needed before CEA can be used as a final guide for the implementation of HLA-B\*58:01 screening. There was a similar debate for the implementation of HLA-B\*15:02 and HLA-B\*57:01 alleles initially, which led to their initial implementation for vulnerable groups over the years. Cost-effectiveness findings are sensitive to a number of factors, including medical care cost of gout management, incidence of allopurinol-induced SJS/TEN, probability of death with SJS/ TEN, HLA-B\*58:01 frequency and cost of genetic testing (Saokaew et al., 2014). A slight fluctuation of one factor can turn a positive result into a negative one and vice versa. The two most important factors are the HLA-B\*58:01 frequency and cost of genetic testing. A higher HLA-B\*58:01 frequency in a specific ethnic group always favoured overall testing in any country. Cost of testing plays a big part in CEA and can lead to genotyping rejection or use in a country. In cases where the

government subsidized the genotyping cost or where the cost was low, it was automatically used for all gout patients.

### **2.5.2 Positive CEA studies**

Park et al. published a twelve month study on the CEA of HLA-B\*58:01 genotyping on gout patients with chronic renal insufficiency (CRI) in Korea. Korea has a HLA-B\*58:01 frequency of 12.2% with a SCARs incidence of 18% in positive HLA-B\*58:01 CRI patients and a mortality rate of 27% in the base case scenario (Park et al., 2015). The CEA showed robust results where genotyping guided treatment was both cheaper and decreased the frequency of SCARs and deaths. The normal allopurinol therapy had a cost of \$1193, while HLA-B\*58:01 genotyping cost was \$1055 in this CEA. The probability of SCARs and death rate were 2.19% and 0.59% in conventional treatment and 0% for both in genotyping. The only way to negate the genotyping based treatment is if the prevalence of HLA-B\*58:01 decreased to 3.18% (Park et al., 2015). Five difference sensitivity analyses were performed where different factors were decreased, starting with HLA-B\*58:01 prevalence, SCARs incidence, death rate, cost of hospital admission and death rate and a combination of all above scenarios. All of latter were cost-effective, except the fifth scenario. Therefore, this study shows that for CRI and other at-risk patients, HLA-B\*58:01 genotyping can be cost effective and feasible, especially with cheaper testing methods (Park et al., 2015).

Ke et al. used a decision-analytical model to compare direct medical costs and effectiveness, including lifetime saved and quality-adjusted life years (QALY) gained in treating new patients. A 1-year time frame and third-party payer perspective were modelled for both the entire cohort (base-case) and for the subgroup of patients with chronic kidney disease (CKD). Four treatment options were used, starting with genotyping and allopurinol use for HLA-B\*58:01 non-carriers, universal benzbromarone, febuxostat and allopurinol use. The Taiwan Food and Drug Administration has suggested HLA-B\*58:01 screening for patients before prescribing allopurinol since 2009 (Ke et al., 2017). The incremental cost-effectiveness ratio (ICER) of genetic screening prior to therapy was estimated as US \$7508/QALY gained in the base-case cohort and was US \$7390/QALY for CKD patients. The ICER/QALY of genetic screening prior to gout therapy remained lower than the World Health Organization WTP threshold of US\$22,635–67,905. Based on HLA-B\*58:01 population frequency (18%) and incidence of allopurinol-related SCAR (2.2/1000 persons), authors

showed that 461 patients needed to be tested to prevent 1 SCAR case. According to the aforementioned incidence, the ICER of HLA-B\*58:01 screening was lower than for the base-case cohort compared with the same alternative treatment strategies. Hence, HLA-B\*58:01 screening gave good value for money in preventing allopurinol-induced SCAR in Taiwan (Ke et al., 2017).

Saokaew et al. used a decision analytical and Markov model to estimate life time costs and outcomes (as QALYs gained) in a societal perspective in Thailand. Input data were obtained from the literature and a retrospective database analysis. A base-case analysis was performed for patients at age 30. A series of sensitivity analyses including scenario, one-way, and probabilistic sensitivity analyses were performed. The overall odds ratio of the association of the HLA-B\*58:01 gene and allopurinol-induced SJS/TEN in Thai was based on a meta-analysis study as 348. Moreover, this study was conducted in accordance with the Pharmacoeconomic guideline in Thailand (Saokaew et al., 2014). Probenecid was used as the alternative to allopurinol as febuxostat was not available in Thailand. Benzbromarone is another alternative drug with mild/moderate renal complication and possible hepatotoxicity.

In the base-case analysis, HLA-B\*5801 genotyping before allopurinol administration decreased the incidence of SJS/TEN to 1.57 cases per 1,000 exposures, and decreased the death incidence to 0.18 cases per 1,000 exposures. Based on a hypothetical cohort of 1,000 patients, the incremental total cost was USD 29,804 and incremental QALY was 5.89 with an ICER of USD 5,062 per QALY gained. HLA-B\*58:01 genotyping before allopurinol use is thus cost-effective based on a standard threshold of USD 5,161 per QALY gained in Thailand (year 2013). The uncertainty in medical care cost of gout management, incidences of allopurinol-induced SJS/TEN, discount rate, and probability of death with SJS/TEN in Thai population had the largest influence on the ICER. An increase in ICER/QALY gained was seen when benzbromarone (cost/tab: USD 0.224) was changed to probenecid (cost/tab: USD 0.056) and with an increase in age (from 25 to 50 years old). In the probabilistic sensitivity analysis, genetic testing increased both costs and QALY in all iterations. At the threshold of USD 5,161 per QALY gained, 49.4% of the iterations were cost-effective. The genetic testing for HLA-B\*58:01 before allopurinol administration is considered a highly potential cost-effective intervention in Thailand. Future studies are still required in Thailand with cheaper and all-in-one testing kits for HLA-B\*58:01, HLA-B\*15:02 and others alleles. The kits for multiple allele

testing will definitely have better cost-effectiveness value due to its broader range of benefits (Saokaew et al., 2014).

### 2.5.3 Negative CEA studies

Plumpton et al. combined a systematic review and meta-analysis to make a decision analytic and Markov model to estimate lifetime costs and outcomes associated with genotyping prior to allopurinol initiation (febuxostat for positive patients) (Plumpton et al., 2017). The number of patients needed to test to prevent one case of adverse drug reaction was 11286. Cost and QALY gains were small, £103 and 0.0023, respectively, resulting in an ICER of £44954 per QALY gained. The probability of testing being cost-effective at a threshold of £30000 per QALY was 0.25. Reduced costs of testing or febuxostat resulted in an ICER below £30000 per QALY gained. Testing remained not cost-effective with an ICER of £38478 per QALY gained for patients with CRI (Plumpton et al., 2017).

Structural sensitivity analysis showed that the sensitivity of the ICER was mostly influenced by febuxostat's efficacy and genotyping cost. The probabilities of genotyping being cost effective at ceiling ratios of £20000 and £30000 per QALY are 0.05 and 0.25, respectively. Scenario analyses showed that genotyping will become cost-effective at a reduced cost of ≤ £21 per patient and availability of a cheaper generic febuxostat. HLA-B\*58:01 testing is unlikely to be cost-effective in the UK; however testing is expected to become cost-effective with reductions in the cost of genotyping, and with the future availability of cheaper, generic febuxostat (Plumpton et al., 2017).

Jutkowitz et al. evaluated the cost-effectiveness of HLA-B\*58:01 testing according to the different ethnicities in the USA; the Caucasians, Hispanics, African Americans and Asians. Costs, QALYs and ICERs were estimated over a lifetime. HLA-B\*58:01 negative patients were given allopurinol and HLA-B\*58:01 positive patients febuxostat. Prevalence of HLA-B\*58:01 was lower in Caucasians and Hispanics (0.7%) and higher in Asians (7.4%) and African Americans (3.8%) (Jutkowitz et al., 2017). Universal HLA-B\*58:01 testing was more costly and more effective for all ethnicities, compared to no testing, along with varying ICERs across ethnicities. Based on their HLA-B\*58:01 prevalence, CEA was positive for African Americans (ICER \$ 83,450) and Asians (ICER \$ 64,190) but negative for Caucasians and Hispanics (ICER \$ 183,720), based on a WTP threshold of \$109,000/QALY. All results were robust in one-way sensitivity analyses with varied parameters, except in the case of halving the SJS/TEN risk and varying the cost of febuxostat. The same conclusions were obtained when AHS risk, febuxostat cost, cost of SJS/TEN and long term complications were varied. Two-way sensitivity analysis showed that a slight change in HLA-B\*58:01 prevalence and testing cost (base-case value \$ 129) impacted the ICER greatly. An increase in HLA-B\*58:01 prevalence caused a decrease in the ICER of testing. In the base-case, universal testing was cost-effective, with a HLA-B\*58:01 prevalence more than 1.6% and a WTP of \$109,000/QALY. Testing cost needs to be less than \$52 to be cost-effective in Caucasians and Hispanics (Jutkowitz et al., 2017).

The probability sensitivity analysis showed that a WTP of \$109,000/QALY testing was cost-effective in 9%, 77% and 88% of simulations for Caucasians/Hispanics, African Americans and Asians respectively. HLA-B\*58:01 frequency varies within the same race by geographic regions, such as a 3.8% frequency in African Americans, but a greater frequency of 7% to 10% in Black Kenyans. However, the final decision to screen for HLA-B\*58:01 applies in both populations. Thus, HLA-B\*58:01 testing was cost-effective only for Asians and African American in the USA.

#### 2.5.4 CEA in Malaysia

Chong et al. performed a thorough cost-effectiveness analysis on the use of HLA-B\*58:01 screening before use of allopurinol in the Malaysian population. The final conclusion stated that HLA-B\*58:01 screening before allopurinol prescription will not be cost-effective. This was emphasized mainly through a probability of 99.9% of the current practice being cost-effective at a threshold of USD 8695 per QALY compared to HLA-B\*58:01 screening (0.1%) and probenecid (0%). Both HLA-B\*58:01 screening and probenecid prescribing were dominated by current practice, with 0.252 QALYs loss/patient at an additional cost of USD 322 and 1.928 QALYs loss/patient at an additional cost of USD 2203, respectively. This is because of the low incidence of allopurinol-induced SJS/ TEN in Malaysia and the lower efficacy of probenecid compared with allopurinol in gout control (Chong et al., 2018).

This CEA adopted a societal perspective with a lifetime horizon, along with a decision tree model and Markov models for cost and outcome estimation, on a hypothetical patient cohort of 10,000. Three treatment strategies were used, universal allopurinol use, HLA-B\*58:01 screening before allopurinol initiation and probenecid use without screening. The model was populated with data from literature review, meta-analysis, and published government documents. Cost values were adjusted for the year 2016 (discounted at 3% per annum) and a series of sensitivity analysis including probabilistic sensitivity analysis were done. However, numerous factors and conditions used in this CEA contradicted the CEA from Singapore and other aforementioned countries. Due to the lack of Malaysian HLA-B\*58:01 studies, this CEA might greatly underestimate the need of screening. The conflicting result in this study might be a result of the limited HLA-B\*58:01 studies done in Malaysia and the limited amount of affordable screening methods. The allelic frequency of the HLA-B\*58:01 allele in Malaysia has also not been properly established due to little or no statistically significant sample size studies. This study only looked at SJS/TEN and left out DRESS, FDE, AHS and the whole range of side effects associated with allopurinol hypersensitivity. All other factors, such as renal function, lifestyle and dietary intake, were assumed to be normal. This does not portray the real-life scenario of Malaysian gout patients who have suboptimal lifestyles, dietary habits, low compliance to drugs and high comorbidity (Chong et al., 2018).

This study uses the positive predictive value of 1.52%, from a HLA-B\*58:01 test done in Taiwan only on Han Chinese patients, as a reference. This is not representative of the Malaysian population where the multi-ethnic population is made up of Malays (68.8%), Chinese (23.2%), Indians (7.0%) and other aboriginal groups still not properly quantified yet (Negeri, 2015).

The main alternative drug used for this cost-effective analysis study was probenecid, which is less effective compared to Febuxostat. Probenecid has several disadvantages, such as kidney stones formation, hepatotoxicity, increased gout attack frequency, it is not used during sudden/severe gout attacks and may worsen a patient's condition. All these reasons show that probenecid is not the best primary drug to be used as a comparison for a Malaysian population (Chong et al., 2018). The specific type of HLA-B\*58:01 genetic screening and its cost were also not mentioned in this CEA. There are numerous affordable screening methods which have been developed nowadays and the choice of screening method can definitely impact on the final cost-effectiveness analysis. Aforementioned CEA papers specifically emphasized that price of HLA-B\*58:01 testing constitutes a great part of the CEA in any country.

Moreover, the full cost of various SJS/TEN sequelae on patients were not calculated and added. Sequelae can be quite severe, with lifetime treatments involved. This study only mentioned a 40.6% chance of developing ocular sequelae. The full range of sequelae from SJS/TEN patients include more than 10 different types of them, as aforementioned. Comorbidities were also not mentioned in this study, along with other risk factors affecting gout patients and subsequent treatment. An SJS/TEN case already costs ten times more than the first year of gout treatment, showing how it impacts patients' treatment expenditure. Avoiding SCARs altogether by screening for the HLA-B\*58:01 allele will only cost a fraction of that amount and save patients from all those severe sequelae. Base case analyses results show that the incidence of SJS/TEN was reduced from 20 cases to 2 cases per 10,000 patients. This number becomes more significant when considering a greater number of gout patients. 180 lives will thus be saved for 100,000 gout patients. Sensitivity analysis shows that febuxostat use worsens QALY loss per patient and is more than six times more costly in terms of additional cost, compared to HLA-B\*58:01 screening before allopurinol. Hence this shows that HLA-B\*58:01 screening is better than febuxostat use,

which is the next best drug available to properly control uric acid level in gout patients nowadays (Chong et al., 2018). This analysis showed that HLA-B\*58:01 genetic testing before allopurinol initiation is unlikely to be a cost-effective intervention in Malaysia. However, this decision might still change when other factors/conditions are varied and with more future HLA-B\*58:01 studies in Malaysia (Chong et al., 2018).

## 2.6 New approach: The High Resolution Melt (HRM) method

### 2.6.1 Introduction

The High resolution melt (HRM) method was introduced in 2002 and its use has been proven in genotyping and single nucleotide polymorphism (SNP) scanning, to name a few methods. This method is shown to overcome all the shortcomings of previous HLA typing and screening methods by being rapid, simple, cheap and accurate. Studies also shows its use in diagnosis of mutated genes in human diseases and in assessment of DNA methylation and microsatellite presence (Reed, 2007).

With the recent advances in technology offering better double-stranded DNA (dsDNA) binding dyes and advanced real-time PCR machines and analysis, HRM is now easier and more accessible to people. HRM can be used to differentiate between different DNA sequences based on factors such as their GC composition, length and DNA complementarity. HRM is used to identify genetic variations in DNA sequences by using the PCR melting curve method. It starts with an optimised PCR amplification of a precise DNA region with specific primers designed. The reaction mix contains the dsDNA-binding dye which fluoresces highly in the dsDNA-bound state and will have a 1000-fold decreased fluorescence level in the unbound state, i.e. single stranded DNA (ssDNA) state. The high resolution melting step then occurs, where the PCR machine's optical system efficiently captures a large amount of fluorescence data per small changes in temperature. As the PCR steps are running, the dsDNA strands denature into ssDNA strands (Figure 2.21), emits decreased fluorescence which is captured and generates a melting curve profile specific to the amplicon. Each DNA sequence is unique, and will thus generate a unique melt curve (Reed, 2007).

During HRM, temperature is increased in short increments of 0.008-0.2°C allowing detailed melting curve generation and analysis where a single base change (G to A) is detected easily in otherwise identical DNA sequences. HRM thus monitors the denaturation of dsDNA samples in real-time by using the intercalating dyes. The optical system in the HRM machine watches this whole process of DNA denaturation to single strands and generates a detailed melting profile to match the sample's detailed DNA sequence. The melt curve profiles will

change according to the length, GC percentage and the heterozygosity. Melting curve profiles will generate valuable data required for genotyping, screening of mutations and for methylation. HRM screening uses the detailed melting profile of a positive control as reference and screens numerous samples alongside in one run. Any sample which completely matches the melting profile for the positive control shows presence of the desired allele. Similarly, HRM can detect a small mutation with a difference of a fraction of a degree due to its high resolution and accurately translate this mutation observed into a change in melt curve shape or a shift of the whole melt profile. Moreover, homozygous variants can be identified by a temperature shift of the melt curve and heterozygotes will show a change in the melt curve's shape itself compared to the wild-type. Thus the HRM method can detect base-pair mismatching caused by destabilised heteroduplex annealing between wild-type and variant strands (Reed, 2007).

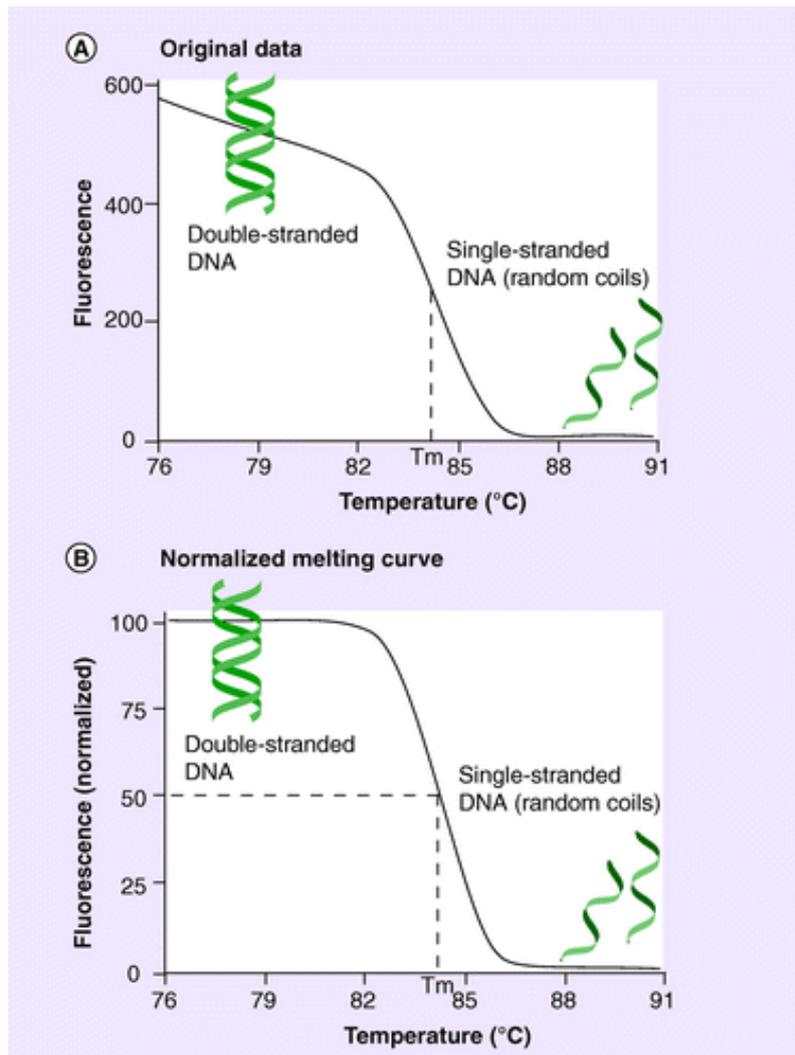


Figure 2-20 The characteristic HRM melting profile principle where an increase in temperature causes the denaturation of dsDNA to ssDNA. Fluorescence is shown to be inversely proportional to temperature rise as per the HRM principle. Figure A shows the original fluorescence data, while figure B shows the derived normalized melt curve after background fluorescence subtraction (Reed, 2007).

HRM profiles can be analysed in 2 ways. Firstly, melt curves generated with the same shape but different melting points ( $T_m$ ) each allow the differentiation of variant homozygous samples compared to a wild-type DNA sequence or a positive control. The difference in  $T_m$  is seen due to the DNA sequence variation in different samples compared to the reference DNA sequence used. The greatest change in fluorescence level will be seen at the  $T_m$  of the product. The  $T_m$  is the point in a melt curve where 50% of DNA is in the dsDNA form and the

other 50% in the ssDNA form (Reed, 2007). Secondly, distinct melt curves from the homozygote melt profile generated are caused by base-pairing discrepancies in the PCR reactions. Analysis of the HRM results does not need any post-PCR experiments such as gel electrophoresis. The software will rapidly generate easy amplification, normalized and difference melting profiles after the run (Reed, 2007).

This method uses a closed-tube system which decreases possibilities of errors and contamination. Moreover, the reaction cost is cheaper compared to other screening methods for genetic variations. The HRM method is very sensitive as changes in single nucleotides are detected and portrayed in the melting curves generated. Thus screening for a specific allele can be done by simply comparing DNA melt curves of an unknown sample to a wild-type of positive control reference sample. Similar melting curves would indicate the same DNA sequence in two samples while different melting curves would indicate the contrary. With an optimised HRM method, numerous samples can be screened for the presence of the HLA-B\*58:01 allele in a maximum of 1 hour and 30 minutes. The HRM method thus generates results faster than any screening/typing methods in history. DNA source can be obtained from blood, buccal cells or saliva. This method has been proven to have a 90-100% sensitivity with easy detections of low-level mutations, nucleotide substitution, insertion and deletions. The HRM method has been widely used for donor matching in organ transplantation, cancer research and genetic disease identification. The HRM technique is thus shown to be very cost-effective due to its simplicity, minimum reagents required and its rapid run and results generation (Reed, 2007).

### 2.6.2 HLA screening by the HRM method

In a pioneering study in 2004, Zhou et al. established HRM typing for HLA-A alleles by targeting its exon 2 and 3 and suggested that this method can be used as a rapid, inexpensive screen before transplantation. Because of the strong linkage between various HLA loci, HLA haplotypes are inherited as a block over 98% of the time. This study combined nested PCR to improve sensitivity and specificity, followed by HRM melting curve analysis in less than an hour. The percentage of heteroduplexes and homoduplexes after PCR and their melting temperatures ( $T_m$ ) determine the melting curves' shape. The  $T_m$  of each heteroduplex depends on the number of nucleotide mismatches and what the specific mismatches are. Six genotypes were screened (A\*01:01, A\*0201, A\*0301, A\*2402, A\*2602, and A\*3301) in all combinations, along with homoduplex and heteroduplex  $T_m$  estimation. The genotype inheritance was also established, along with HLA-A identity among siblings (Zhou et al., 2004).

HLA-B27 is strongly associated with the disease ankylosing spondylitis and is of diagnostic value because 90% of patients with ankylosing spondylitis have the B27 antigen. Two commonly used HLA-B27 flow cytometric assays are commercially available. Seipp et al. designed an allele-specific PCR melting assay for HLA-B27 and compared it with two available antigen assays on 371 clinical samples. The accuracy of the assays was measured by receiver operating characteristic analysis using the PCR method and sequencing as the reference standard. Using DNA sequencing as the gold standard, the sensitivity and specificity of PCR were 99.6% and 100%, those of the best single antigen assay were 98.2 and 97.6, and those of a reflex combination of both antigen assays were 98.8 and 97.6. Thus, the allele-specific PCR melting assay for HLA-B27 genotyping is easy to perform and has better sensitivity and specificity than antigen assays (Seipp et al., 2005).

Imperiali et al. designed and developed an easy, fast, and inexpensive HRM assay to detect HLA-B\*51, the strongest known genetic risk factor for Behçet disease. Results were mostly qualitative, with the presence of double peaks in the melting curve for positive samples and internal control presence, while negative samples showed only one peak for the internal control. This HRM assay genotyped 61 samples and confirmed the results by conventional PCR and gel electrophoresis. The HRM evaluated had 100% of specificity, sensibility, and

repeatability, and 0% of false positive and false negative rates. Therefore, this HRM analysis is easily applicable to the rapid detection of HLA-B\*51, exhibits a high speed, and requires a very low budget (Imperiali et al., 2015). Cui et al. developed a HRM method for the genotyping of at risk HLA-DQA1 and PLA2R1 alleles, which are linked to idiopathic membranous nephropathy (IMN). The SNPs rs2187668 and rs46643308, from the two alleles, were specifically targeted by primers and genotyped in 480 healthy volunteers. This study used the principle where the amplified PCR products differing in one single nucleotide for each SNP will result in different melt curves. Different genotypes were identified by a shifted melting temperature ( $T_m$ ) or by differences in melting curve shape. Reference samples with different genotypes were used to establish the normalization regions by first selecting the pre-melt and post-melt regions. Unknown samples were genotyped by comparison to the references. The two SNPs were easily identified by the altered shape of the melting curves or the homozygote (AA or GG) was identified by a change of  $0.4^\circ\text{C}$  in melting temperature ( $T_m$ ). The success of this HRM method was fully dependent on the highly specific primers targeting only the SNPs. Moreover, this HRM method was able to detect two SNPs in one amplicon simultaneously, thus showing the ability to screen fragments containing more than one close variant. Three different SNP genotypes were distinguished by HRM, along with mean sensitivity of 98.8% and mean error rate of 1.9% (Cui et al., 2013).

Lundgren et al. used HRM analysis to discriminate between MHC class I and killer cell lectin-like receptor allele variants, critical to natural killer (NK) cell-mediated viral control. HRM is thus capable to interrogate and quantify gene- and allele-specific variations due to differential regulation of gene expression. HRM analysis discriminated between highly polymorphic and polygenic murine NK receptor and MHC class I allele variants, and determined gene dosage, even in highly related gene families. HRM genotyping was successfully done for congenic, transgenic, and other gene modifications in genetically selected mice. As an identification strategy, HRM is a powerful method that can easily be harnessed for the purpose of following genetic manipulations in targeted mouse strains (Lundgren et al., 2012).

Chen et al. used HRM to genotype SNPs of VKORC (1173T/C, rs9934438) and CYP2C9 (1075A/C, rs1057910), which are major contributory factors on the sensitivity of warfarin in Chinese. 255 samples were screened via HRM and results matched DNA sequencing results 100% for rs1057910 and 99.2% for rs9934438. Different melt curve shapes and  $T_m$  were detected for presence of homozygous or heterozygous samples. Unexpected mutations, due to closely located variants, were also identified by completely different melting curve shape. The HRM assay in this study had an estimated turnaround time of 3-4 hours (DNA extraction and HRM run) and cost around \$10 per sample (Chen et al., 2015).

A Malaysian study conducted by Rani et al. used the HRM method to screen for HLA-B\*15:02 epileptic patients with SCARs presentation. This HRM screening method showed 100% sensitivity and specificity in screening for the HLA-B\*15:02 allele. Moreover, being a quantitative method, HRM obliterated the disadvantages of qualitative methods, such as multiplex PCR and real-time PCR. The cost of this HRM method was around USD 16.40 per test and takes around 1 hour and 30 minutes for the run time. Rapid turnaround time (TAT) is crucial for a pharmacogenomics screening method as the patient must wait for the results before getting their drugs prescribed by doctors. The HRM method thus shows the quickest TAT to be used for pharmacogenomics testing purposes, along with its other aforementioned advantages (Rani et al., 2018).

This HRM method has been successfully implemented in HUKM hospital in Malaysia for routine screening of the HLA-B\*15:02 allele for around 2 years (Rani et al., 2018).. Doctors have adopted this method before prescribing, and this led to a significant decrease in the number of SCARs. Thus, the HRM method shows a promising future for the screening of HLA-B\*58:01 allele in Malaysian hospitals.

### 2.6.3 Multiple HLA allele theory

Another phenomenon being slowly discovered is the presence of a set of different HLA genetic markers, previously linked to one disease only, being linked to more than one disease. We named this the multiple HLA allele theory, whereby the set of genetic markers previously identified (e.g; HLA-B\*15:02, HLA-B\*58:01, HLA-B\*57:01) are found linked to SCARs in different degrees for all the linked diseases, such as epilepsy, gout and AIDS. Moreover, the discovery of numerous HLA alleles linked to different degree of SCARs, and not all of them, further shows that even more variation is present and that one HLA allele will not be enough to act as a fool-proof genetic marker. For the past decades, scientists have focused on singling out one HLA allele for one disease for screening purposes, but the presence of haplotypes was not ignored. The presence of haplotypes and further studies showed that more than one HLA allele is linked to numerous diseases and novel HLA alleles are continuously being discovered in parallel. Hence, scientist might have to change their direction from single HLA allele detection to multiplex allele detection which will encompass a wider set of diseases and ADRs (Chang et al., 2020).

One such example is the identification of numerous HLA alleles involved in carbamazepine (CBZ), lamotrigine (LTG), phenytoin (PHT), and oxycarbazepine (OXC) -induced SCARs in epileptic patients. This association was more easily identified for epilepsy patients first, maybe due to this wider range of drugs causing SCARs in patients, hence opening up more possibilities. This starts with HLA-B\*15:02 and HLA-B\*15:21 found to be both associated to epileptic CBZ induced-SCARs in the Javanese and Sudanese population. Thai patients were also found to have HLA-B\*15:21 linked with CBZ-induced SJS/TEN. A large European GWAS study and a Japanese study showed the linkage of HLA-B\*31:01 allele to numerous CBZ-SCARs, namely SJS/TEN, HSS and MPE. HLA-B\*31:01 was found linked to CBZ-induced DRESS, while HLA-B\*15:11 was associated with CBZ-induced SJS in a Korean study. Another study suggested that HLA-B\*15:11 was associated with CBZ- induced SJS/TEN in Japanese patients too (Then & Raymond, 2019).

Interestingly, a Thai study found HLA-B\*15:02 and HLA-B\*58:01 to be linked with CBZ-induced maculopapular eruption (MPE) patients, as well as an association of HLA-B\*58:01 with CBZ-induced DRESS. This will be investigated in this study by NGS, to see if these two

alleles are present concomitantly in Malaysian gout and epilepsy patients. Tassaneeyakul et al. reported that 6 HLA alleles including HLA-A\*33:03, HLA-B\*38:02, HLA-B\*51:01, HLA-B\*56:02, HLA-B\*58:01, and HLA-C\*14:02 were significantly associated with PHT-related SJS/TEN, whereas only HLA-B\*51:01 was significantly associated with PHT-related DRESS. PHT is another SCARs causing drug given to epileptic patients which has been associated to the HLA-B\*15:02 allele. Another Thai study found that HLA-B\*13:01, HLA-B\*56:02/04, CYP2C19\*3 and co-medication with omeprazole were strong risk factors for PHT-induced DRESS. A Malaysian study also reported another novel HLA allele, HLA-B\*15:13 that was significantly associated with PHT-induced SJS/TEN. A Spanish study showed that concomitant HLA-A\*02:01/Cw15:02 alleles were linked to PHT-SJS/TEN (Then & Raymond, 2019; Chang et al., 2020).

A study showed that concurrent testing of CYP2C\*3/HLA-B\*13:01/ HLA-B\*15:02/HLA-B\*51:01 increased the sensitivity of screening from 16.41% up to 71.88% with an OR of 8.88 (95% CI of 5.63-14.01;  $P=2.12 \times 10^{-23}$ ) in the Taiwanese Han Chinese cohort. Oxcarbazepine (OXC), another anti-epileptic drug, was linked to HLA-B\*15:02 in SCARs and the newly discovered HLA-B\*13:02 in OXC-induced MPE. Other authors also reported the lack of association of HLA-B\*15:02 with OXC-induced MPE in patients of northern China, but it was associated with HLA-B\*38:02. The genotyping of patients with OXC-induced SJS/TEN also revealed its new association with HLA-B\*15:18. A Korean study in turn linked OXC-induced MPE with HLA-B\*40:02 (OR 14.64, 95% CI 1.73- 123.90;  $P=0.003$ ) and DRB1\*04:03 (OR 0.18, 95% CI 0.04-0.82;  $P=0.016$ ). Interestingly, this study also found that HLA-B\*15:01 was a protective allele (OR of 0.18, 95% CI 0.04-0.82;  $P=0.016$ ) (Then & Raymond, 2019; Chang et al., 2020).

Lamotrigine (LTG), another SCAR inducing anti-epileptic drug, was linked to HLA-B\*15:02, HLA-B\*33:03, HLA-B\*35:08 and HLA-B\*44:03. Another study showed the presence of HLA-B\*44:03 in patients with LTG-induced SJS/TEN. They did a follow-up study with larger sample size and confirmed that HLA-B\*44:02 may have an association with LTG-induced SJS/TEN with an OR of 12.75 (95% CI 1.03-157.14;  $P=0.053$ ). Another study showed the presence of HLA-A\*24:02, HLA-Cw\*01:02 and HLA-Cw\*07:02 in the LTG-MPE cohort, along with HLA-A\*33:03 reported as a protective allele against LTG-induced MPE. Fricke-Galindo et al. reported the HLA-A\*02:01/HLA-B\*35:01/HLA-C\*04:01:01 haplotype was significantly

associated with LTG-induced MPE in Mexican Mestizo patients TEN (Then & Raymond, 2019; Chang et al., 2020).

All the aforementioned studies showed how a bigger number of HLA alleles can be linked and used as a genetic marker for one disease. Hence, this gives rise to another theory, namely the multiple HLA allele theory, where a set of common HLA alleles are found linked to several SCARs-disease association, rather than a single association. NGS will thus be a tool to help in investigating this in this study. This theory will thus change the face of pharmacogenetic testing from single HLA allele detection to multiplex HLA allele detection for a wider set of SCARs and diseases globally. The ethnicity link might also be downplayed here as a wider set of HLA alleles will be screened for numerous diseases, acting as a HLA panel to detect and prevent ADRs early, similarly to allergy detection to food or other substances TEN (Then & Raymond, 2019; Chang et al., 2020).

# CHAPTER 3: METHODOLOGY

### 3.1 Introduction to Methodology section

This chapter will fully elaborate on all the methods used to achieve the objectives set for this research. Methods usually started with simple procedures, which then needed vigorous optimisation in order to reach the desired results. This chapter will provide the detailed final optimized protocols for each method used, so as to allow reproduction of this thesis' results by anyone. The operational flowchart below (Section 3.1.1) depicts all the different methods used in their sequential order, thus showing a logical flow and build-up to reach the desired aims. To start this project, an important question to be asked was the number of samples needed to prove and test the theory. Thus, the sample size was first calculated using proven methods, followed by the actual sample collection and DNA extraction. Once the starting material needed for experiments was in hand, i.e DNA samples, the primers needed to be designed to target the desired whereas of the allele being investigated. The starting point of the HRM method was the standard PCR reaction, which was optimized for the primers designed. Once a robust PCR method was designed and optimized, the last crucial element needed to set up the HRM method was established. A positive control, containing the HLA-B\*58:01 allele, needed to be used as reference in every single HRM run. Thus, to obtain an infinite amount of the limited positive control, the latter was cloned in *E. coli* vectors and stored for repeated use. The HRM method was then setup and optimized based on the previous two methods, in order to develop the HLA-B\*58:01 allele's screening method. The NGS method was used to validate the HRM method's results and also to delve deeper into the HLA-B\*58:01 allele's significance in this research. This concluded the methods used to prove the aims and objectives set at the start of this research.

### 3.1.1 Operational flowchart

The diagrammatic representation in Figure 3.1 below showed the summarized, step-by-step approach used to solve the questions set at the start of this research.

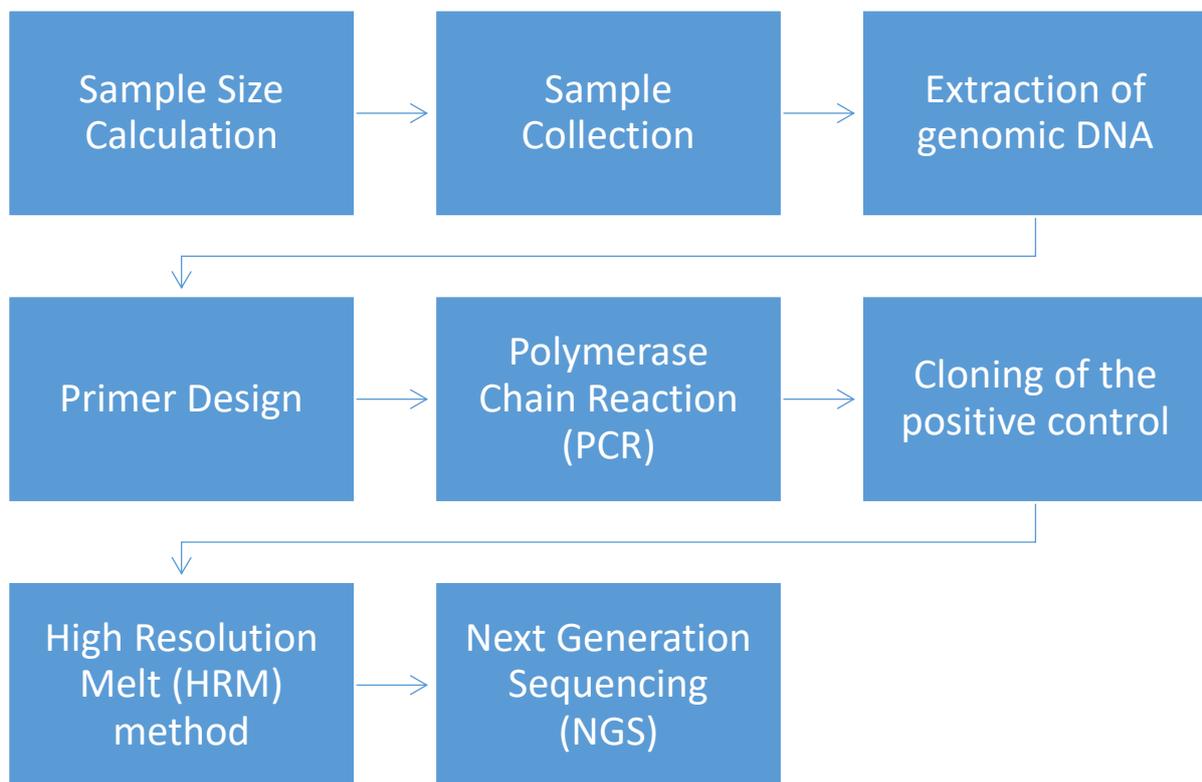


Figure 3-1 Complete operational flowchart, showing all the different methods used in sequential order (DIY).

## 3.2 Sample size calculation

### 3.2.1 Introduction

Sample size calculation has always been a crucial issue faced by biomedical researchers in the early stage of their study design. Numerous published studies in international journals reported incorrect and smaller sample size, thus leading to studies with less power in their respective fields. Different study designs required different methods of sample size calculations as one formula could not be used for all of them. Cross sectional studies were done to estimate a population parameter, such as the prevalence of a certain disease in an area. These studies could either be for qualitative or quantitative variables, whereby the sample size formula will differ for both. Studies investigating HLA alleles were examined and the best sample size calculation formula extracted from them, as shown in Section 3.2.2 below.

### 3.2.2 Sample size calculation

In this study, the High Resolution Melt (HRM) method was developed to screen for the HLA-B\*58:01 allele in the Malaysian population. The HLA-B\*58:01 allele frequency in Malaysia used at the start of this study was 10.49% (González-Galarza et al., 2015). This study investigated on the proportion of Malaysians who have the HLA-B\*58:01 allele and thus emphasized on the proportion part, which was a qualitative variable. The sample size needed for this study was thus calculated from the formula by Kish L (1965) as shown on the next page. This study will thus need around 145 subjects to have sufficient statistical power. 145 gout patient samples will be needed, along with 145 control samples, which will have matching age, gender and ethnicity.

Calculation for required study sample size by Kish L (1965) (3)

**Formula used:  $n = (Z_{1-\alpha})^2(P(1-P)/D^2)$**

Where;

$(Z_{1-\alpha})^2$ = the standard normal variate (at 5% type 1 error ( $P<0.05$ )) which is 1.96.  
Majority of studies use P values which were significant below 0.05. Value obtained from the normal distribution table, standard value for 95% CI

P= expected proportion of allopurinol-induced SJS/TEN patients with the HLA-B\*58:01 allele in Malaysian population based on previous studies (10.49%)

D= Absolute error or precision set at 5%

**Final calculation:**

$$Z_{1-\alpha} = Z_{0.95} = 1.96$$

$$P = 10.49\%$$

$$D = 5\%$$

$$n = (1.96)^2 \times [0.1049(1-0.1049)/0.025] = 144.3 = 145 \text{ (round off)}$$

### 3.3 Sample collection

Collection of blood samples from patients in hospitals required careful planning from the ethics applications, consent form designs, finding doctors for collaboration to the actual collection of samples. Many people were involved in this first stage and we had to ensure that all the procedures were done systematically. The flowchart in Figure 3.2 below showed the steps involved in blood samples collection and were further explained in the subsequent sections.

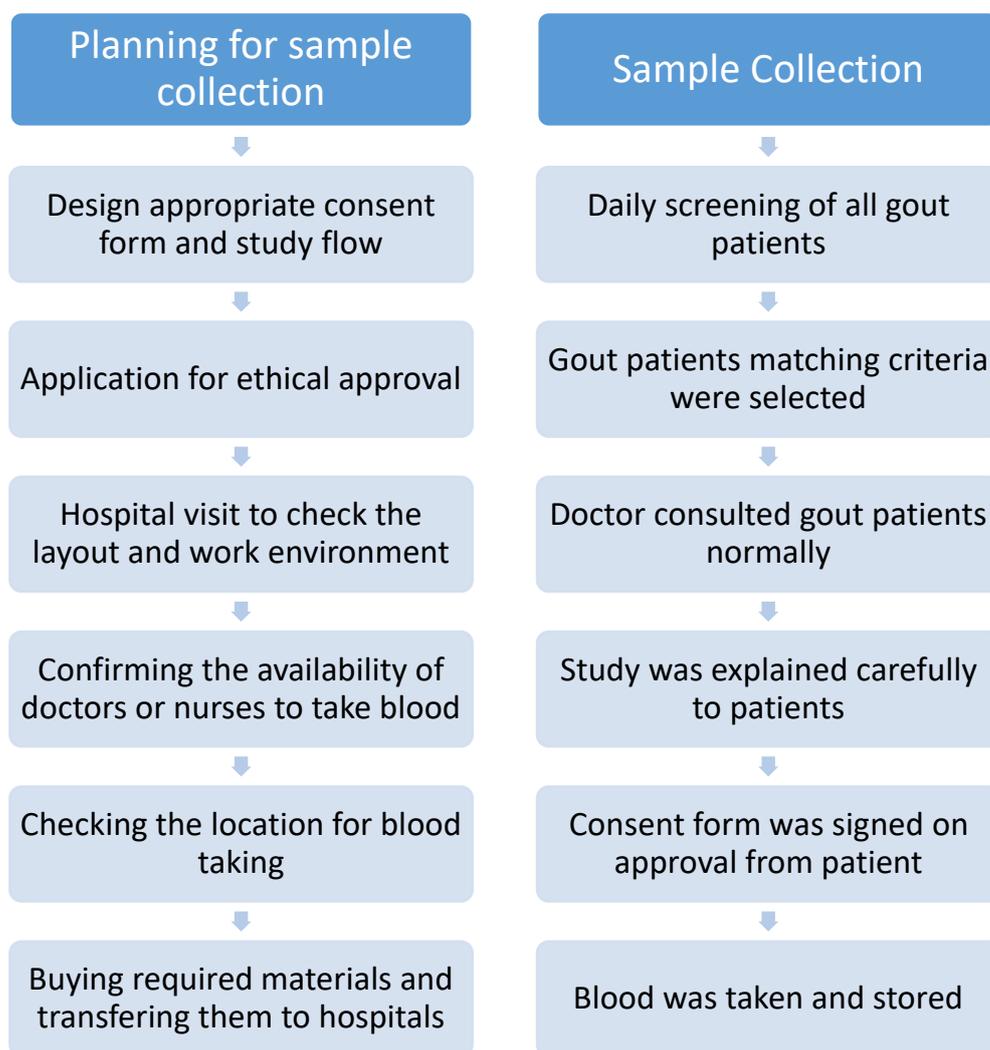


Figure 3-2 Flowchart showing the planning process before sample collection and the flow of sample collection done in hospitals (DIY).

### 3.3.1 Consent form design and ethics approval

The consent form (shown in Appendix A) was designed according to the guidelines on the National Medical Research Register (NMRR) of Malaysia and was checked and approved by all the doctors/principal investigators involved. The consent form was divided into three parts, namely, the patient's information sheet, the consent form and a questionnaire. The consent form was written in English and Bahasa Melayu (BM), to cater for language differences, especially for the elderly patients. The form started with an information sheet, with sufficient details about the study, explained in layman terms for the general audience to understand the blood taking procedure, followed by the benefits/risks of the study, confidentiality clause and the contacts of the principal investigators for any questions and enquiries. The principal investigator present during blood-taking was also required to verbally explain the whole study, in the volunteer's preferred language, as well as answering any questions. The second part consisted of signing the consent form, where volunteers were required to sign and jot down all their personal details. There was a separate consent form for children not of legal age, where their parents or legal guardians needed to sign and approve of their participation in the study. The last part of the consent form consisted of a questionnaire, which must be completed by the volunteer in order to check if they have completely understood all the details and what their participation in this study entailed. Volunteers were also asked if they wanted to be informed of their HLA-B\*58:01 screening results once the project was completed. Two copies of the consent form would thus be filled up and signed by both the volunteer and the principal investigator on site. One copy would be given to the volunteers and the other would be safely stored by the principal investigator. Full confidentiality and privacy would be kept throughout the study by several mechanisms. First, only the principal investigators had access to the participants' personal data written on the consent forms. Second, these consent forms were stored by only one principle investigator and was always safely locked. Third, all participants were assigned a sample code which will be used throughout the study and their identities or names would never be revealed.

### 3.3.2 Volunteer recruitment and eligibility

Blood samples were collected, on a voluntary basis, from gout patients receiving treatment in the rheumatology clinic in Pusat Perubatan Universiti Kebangsaan Malaysia (PPUKM) and Hospital Putrajaya (HP). The ethics for blood collection and handling were approved by the PPUKM Ethical Committee, the Medical research and Ethics Committee (MREC), Ministry of Health, Malaysia and the Nottingham ethical committee. Inclusion criteria for selected volunteers were Malaysian patients of any race, age, gender, the confirmed presence of gout by doctors, allopurinol use for at least three times and a treatment period of at least one to two weeks for ADRs manifestation. Exclusion criteria for volunteers were an absence of a clear gout diagnosis, only hyperuricemia mentioned by doctors and new patients who used allopurinol only once. Daily screening of gout patients were done in the two hospitals and eligible patients were selected by taking into consideration the aforementioned inclusion and exclusion criteria. 145 gout patients were recruited in total from the two hospitals, with 80 volunteers from HUKM and 65 volunteers from HP. The recruitment period started in September 2015 and ended in September 2018. Recruitment from two different hospitals helped in increasing the initial slow collection speed of blood samples.

### 3.3.3 Blood sample collection

#### Materials required for blood collection

- BD Vacutainer blood collection tubes (10ml)
- 70% Isopropyl alcohol pads
- Hypodermic 10ml syringes and needles
- Winged butterfly needles
- Gloves
- Adhesive plaster

There were numerous ways to collect blood samples from patients, such as arterial sampling, venipuncture sampling and fingerstick sampling. Venipuncture was the most common way of collecting blood from adults in a hospital setting. The latter involved collection from a superficial vein in the upper limbs, usually the median cubital vein which was close to the skin's surface and had only a few nerves located close to it. In this way, pain and discomfort was minimized for the patients. This procedure must be performed only by trained phlebotomists, i.e. doctors or nurses.

10ml of peripheral blood samples were collected from the patients by using BD Vacutainer blood collection tubes with EDTA as anticoagulant (Becton Dickinson, USA). EDTA, an anticoagulant, chelated calcium to prevent clotting and preserved the red blood cells' morphology. The doctors or nurses on site did the blood taking safely, with sterile hypodermic needles and syringes. The blood taking was either the inner elbow or wrist, depending on patients' veins visibility. In cases where blood collection was difficult due to smaller, rolling or spasming veins and delicate state of the patient, especially elderly patients, a minimum of 3ml of blood was enough for all experiments. The blood samples collected were then immediately stored in 4°C and the DNA was extracted on the same day. Patient confidentiality was strictly maintained and all the patients were given code names according to the hospital they were in. 80 gout patient samples from PPUKM were given codes HLAG1 to HLAG80, while 65 gout patients from HP were listed as HLAP1-HLAP65.

Once 145 gout patient samples were obtained, matching controls were retrieved from the UKM Molecular Biology Institute's (UMBI) biobank. The normal samples were retrieved through UMBI's biobank and came from patients who do not have gout and who have not used allopurinol before. The gender, age and ethnicity were matched as close as possible with the respective gout patient samples collected. Age was more difficult to match for each samples, hence a maximum difference of five years, before or after the patient's actual age, was allowed. Those 145 normal samples were all extracted from blood, by using the same method the gout patient samples were extracted. They were labelled as either N01 or NS01 according to their sequence in UMBI's biobank.

#### **3.3.4 Patient medical record examination**

The full medical records of the gout patients were extracted from PPUKM and HP in order to check for any signs of adverse reactions to allopurinol. Presence of all types of hypersensitivity reactions or ADRs to allopurinol were carefully recorded. Their other details were also noted down, such as their full list of drugs and diseases, in order to perform appropriate comparisons and analysis. PPUKM's medical records were stored on paper, in individual patient files and access needed to be requested for a maximum of twenty files at once. A request form needed to be filled with the patients' code and name and signed by the doctor in charge in the rheumatology clinic. This form must then be submitted to the medical records office and the files will be handed out, only for reading purposes, in their office. Hospital Putrajaya's medical records could be accessed on their electronic database, but only in the presence of the doctor in charge as the files were password protected. All these details were compiled in the results section and further analysed in the discussion section.

## 3.4 Extraction of genomic DNA

### 3.4.1 Preparation of reagents for DNA extraction:

#### 1) 1 X Red blood cells (RBC) lysis buffer (pH 8.0)

- a. 10.0 ml of 1 M Tris pH 8.0 (MW: 121.14)
- b. 3.3 ml of 3 M Natrium chloride (NaCl) (MW: 58.44)
- c. 5.0 ml of 1 M Magnesium chloride (MgCl<sub>2</sub>) (MW: 95.21)

Made to 1 litre with distilled water, in a clean and sterile Schott bottle, and autoclaved before use. All chemicals were originated from NacalaiTesque, Inc., Japan.

#### 2) Nuclei lysis buffer (pH 8.0)

- a. 10.0 ml of 2 M Tris pH 8.0 (MW: 121.14)
- b. 0.8 ml of 0.5 M EDTA
- c. 6.6 ml of 3 M Natrium chloride (NaCl) (MW: 58.44)

Added 200ml of distilled water along with ingredients and autoclaved before use. All chemicals were originated from NacalaiTesque, Inc., Japan.

#### 3) 20% Sodium Dodecyl Sufate (SDS)

- a. 100.0 g of SDS (MW: 288.38)

Added 250ml of distilled water along with ingredients and allowed to dissolve before making it up to 500 ml with distilled water. All chemicals were originated from NacalaiTesque, Inc., Japan.

#### 4) 6 M Saturated Salt solution

- a. 34.0 g of Natrium Chloride (NaCl) (MW: 58.44)

Added 100ml of distilled water along with ingredients and autoclaved before use. All chemicals were originated from NacalaiTesque, Inc., Japan.

5) 10x Tris EDTA (TE) Buffer (pH 8.0)

- a. 15.759 g of Tris-Cl (MW: 157.594)
- b. 2.92 g of EDTA (MW: 292.24)

Added 800 ml of distilled water into a clean 1 Litre Schott bottle along with the ingredients mentioned. Autoclaved before use. All chemicals were originated from NacalaiTesque, Inc., Japan.

6) Proteinase K

Bought directly from Qiagen Biotechnology Malaysia Sdn. Bhd.

### 3.4.2 DNA extraction protocol

10ml of peripheral blood samples were collected from the 145 gout patient samples using BD Vacutainer blood collecting tubes (Becton Dickinson, USA). DNA extraction was then performed by the salting out method as described below. This method can be divided into two parts; with removal of red blood cells in Figure 3.3 first, followed by DNA extraction, as shown in Figure 3.4.

10ml of whole blood was transferred to a labelled 50ml falcon tube, filled with 40ml cold, 1x Red Blood Cells (RBC) lysis buffer and inverted a few times. Residual blood in the BD Vacutainer tubes were washed with some RBC lysis buffer and added to the falcon tube to avoid wastage of blood collected. The falcon tube was then centrifuged (Eppendorf, Centrifuge 5810 R) at 2,500rpm for 10 minutes in a refrigerated centrifuge. The supernatant was then discarded slowly without disturbing the pellet with around 5ml of liquid left in the tube. All blood discards were poured down a waste container containing Clorox. 45ml of 1x RBC lysis buffer was added to the pellet, followed by resuspension by using a Pasteur pipette. Centrifugation was carried out again, followed by another round of the abovementioned steps until the pellet became fully white. This pellet can be sealed in the falcon tube and stored at -20°C for a maximum of a week, for later extraction, in cases where large number of samples had to be extracted. The preferred method here was to extract the sample right away. This first step called RBC lysis was depicted in Figure 3.3.

The DNA extraction step (Figure 3.4) started with addition of 2ml of nuclei lysis buffer to the pellet obtained, followed by resuspension. 100µL of Qiagen Proteinase K and 100µL of 20% SDS were immediately added next. This mixture was resuspended, the tube sealed with parafilm (Bemis Company, Inc., USA) and incubated at 60°C for 1hour in a hybrid oven (Binder, Model ED56). In case of incomplete blood digestion, another 100µL of Qiagen Proteinase K was added and incubation was repeated again. After incubation, the mixture was evenly split into two to four labelled 2ml sterile screw cap tubes (Heathrow Scientific, USA). The number of screw cap tubes varied according to the amount of white blood cells pellet obtained in the aforementioned steps. 333µL of 6M saturated salt solution was added

into each tube. The tubes were properly vortexed for around 1 minute until a uniform white colour was seen. The latter were then centrifuged at 15,000rpm for 30 minutes at 4°C. The resulting supernatant was decanted out and transferred to a labelled, 15ml falcon tube containing 5ml of cold, 100% ethanol. The falcon tube was gently inverted 10 times and DNA was precipitated as 'cotton wool' spools. The white DNA clump was carefully sucked out using a Pasteur pipette and transferred to a sterile and labelled 1.5ml microcentrifuge tube. This must be done quickly and swiftly as the DNA clump may be stuck to Pasteur pipette's tip. In that case, it must be carefully separated from the tip using a 10µL pipette tip.

Washing of the DNA pellet was done with 1ml of 70% ethanol, followed by centrifugation at 12,000rpm for 2 minutes at 4°C. Maximum amount of 70% ethanol was removed using 10µL-50µL pipette tips and the tube was spun again at 12,000rpm for 1 minute. The tube was then spun in a speed vacuum (Eppendorf, Concentrator plus basic device, Rotor F-45-48-11) for 5 minutes to dry out the DNA pellet. 100µL-200µL of sterile 1x Tris EDTA (TE) buffer was added to the DNA pellet to dissolve it. The microcentrifuge tube was sealed with parafilm and incubated in a water bath at 55°C overnight for reconstitution.

After DNA extraction, the concentration and purity of the isolated DNA was determined by using the Epoch microplate spectrophotometer (BioTek Instruments, USA) and the software Gen5. Absorbance was measured at 260nm and 280nm and purity was estimated by the Optical density (OD) ratio, 260/280nm. The first 2 wells of the Take 3 Micro-volume plate were filled with 1.5µL of distilled water each, followed by 1.5µL of the unknown sample in the next 2 wells. All samples were read in duplicates, with an average calculated, to increase the accuracy of the spectrophotometer reading. A pure DNA sample will have a purity/OD ratio of 1.8-2.0 and the concentration will vary depending on different individuals. The extracted DNA samples were immediately stored in 50µL aliquots at -20°C.

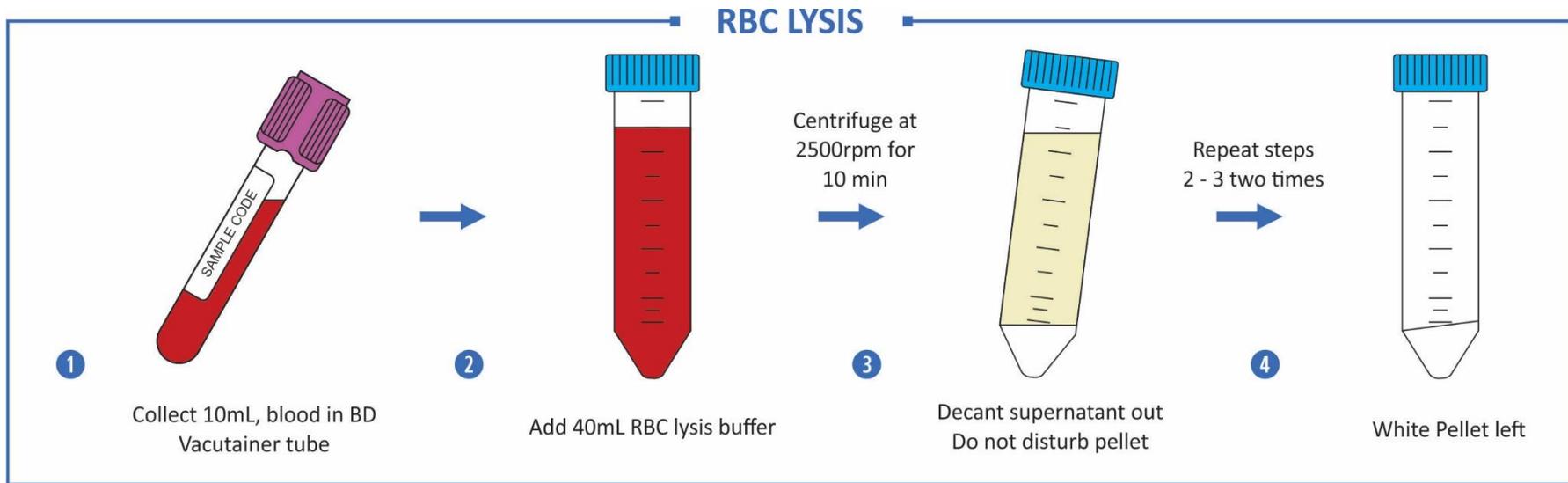


Figure 3-3 RBC lysis was carried out first, where the steps 1 to 3 were repeated 3 times until a white pellet was left. (DIY)

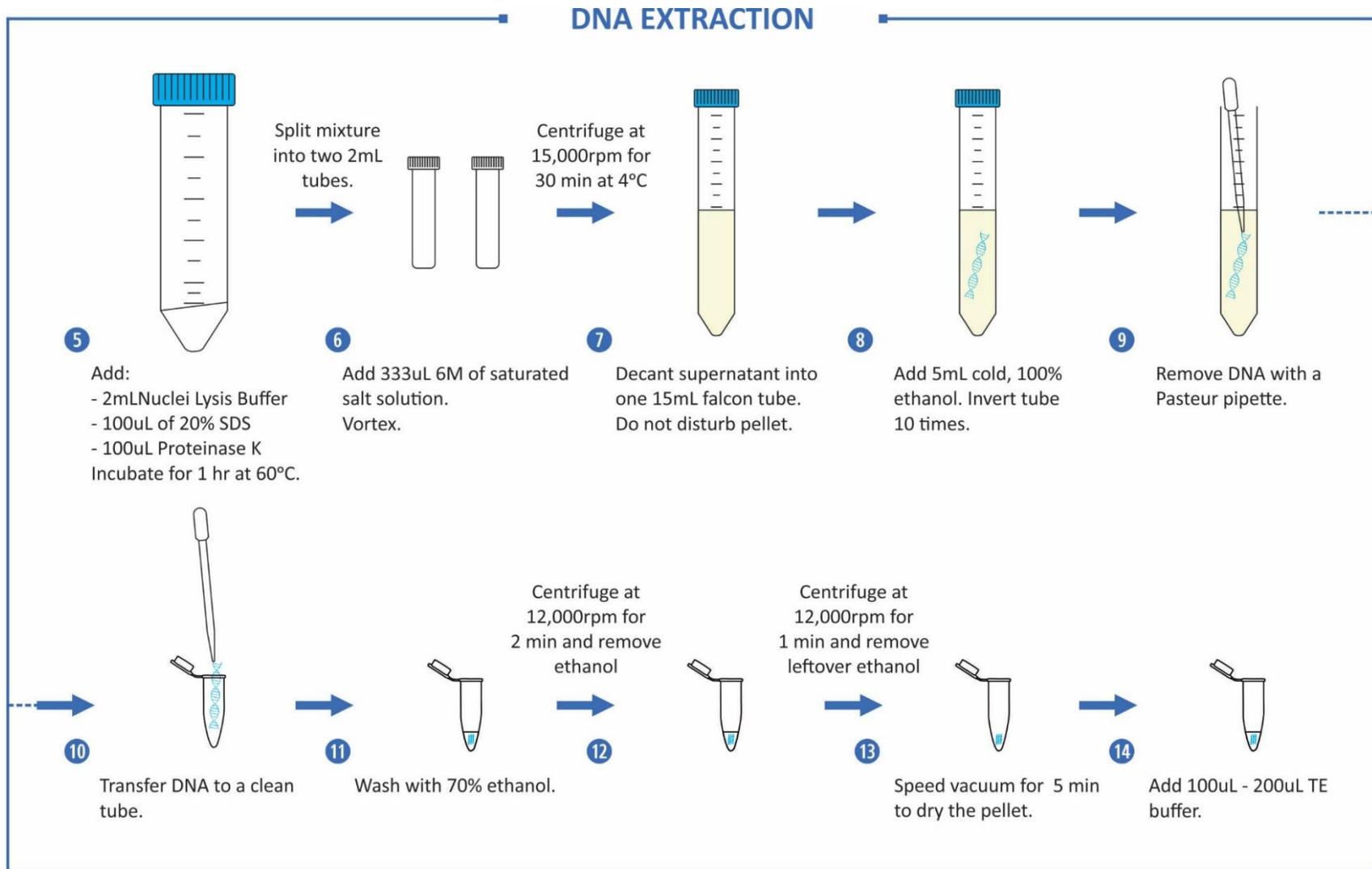


Figure 3-4 DNA extraction step which was done after RBC lysis, started with a white pellet and ended with the extracted DNA, dissolved in TE buffer. (DIY)

### 3.5 Primer Design

The complete sequence of the HLA-B\*58:01 allele was obtained from the IMGT/HLA database of EMBL-EBI. 3 sets of primers were designed by using the Primer designing tool from the National Center for Biotechnology Information (NCBI), the Primer3Plus website and further analysed using the OligoAnalyzer 3.1 from Integrated DNA Technologies (IDT). Primer specificity was of primordial importance for this whole project. A specific set of criteria need to be followed for primer design, such as a GC content of 40-60%, melting temperature of 50-60°C, a low self-annealing value and minimum hetero-dimer formation. Primer blasts were also carried out on NCBI to confirm specificity and avoid unwanted amplification.

Numerous primer sets were designed to find the best region with the highest specificity and consistent results. Designing primers for the whole gene left us with unspecific results due to the high similarity between HLA genes. These primers were further modified to be longer, around 50bp long, in order to optimize binding to a large product. Halving the full gene and primers designed accordingly resulted in good products but with a high level of complexity due to secondary structures, which could not be detected by Sanger sequencing. Overlapping primers were also designed and optimized to try to maximise product length and binding specificity, without avail.

The traditional method of targeting specific exons was thus used, in order to get highly specific products with a good enough length for the study's purposes. The Internal Control (IC) primers targeted the  $\beta$ -globin gene and were obtained from previous studies done on HLA-B alleles. This set of primer acts as housekeeping gene in molecular diagnosis to verify that the PCR conditions are optimum and to control amplification.

The two primer sets, named P2 and P3, targeted the exon 2 and exon 3 of the HLA-B\*58:01 gene, respectively. Exon 2 and 3 were targeted as they form the binding pocket of the HLA-B\*58:01 allele where allopurinol will bind and therefore cause hypersensitivity. The details

of the primers designed were shown in Table 3.1 below. Primers were sent for production at First Base Laboratories Malaysia Sdn. Bhd and were received after a few days. On delivery, primers were instantly diluted with RNAase free water (Qiagen, Germany) to make a stock of 100 $\mu$ M. This stock was divided into two stock tubes, one to be used immediately and one to be stored, for use when the first one was over or in cases of contamination where a fresh stock was needed. All primers were diluted to a working solution of 10 $\mu$ M by 10x dilution with RNAase free water. A few separate RNAase free water tubes were kept for primer dilutions to avoid contamination during experiments.

Table 3-1 Full primer details for the internal control and 2 pairs of primers used to amplify the HLA-B\*58:01 allele, including the primer sequence, length, Tm and GC %.

Primer names	Target region	Primer length (bp)	Primer sequence (5' to 3')	Amplicon size (bp)	Tm (°C)	GC (%)
<b>P2</b>	Exon 2	20	F-GACACCCAGTTCGTGAGGTT	149	57.2	55.0
		20	R-CGCAGGTTCTCTCGGTAAGT		56.6	55.0
<b>P3</b>	Exon 3	20	F-GGTCTCACATCATCCAGAGG	259	57.1	55.0
		20	R-TCCTCCCGTTCTCCAGGTA		59.6	55.0
<b>IC</b>	β-globin gene	20	F-GTGTACACATATTGACCAAA	423	51.1	35.0
		20	R-AGCACACAGACCAGCACGTT		52.2	55.0

## 3.6 Polymerase Chain Reaction (PCR)

### 3.6.1 Standard PCR protocol

Standard PCR reactions were carried out for all 3 primers designed with the Hot Star Taq Master Mix from Qiagen. A standard, optimised 25 $\mu$ L reaction was made of 12.5 $\mu$ L of Hot Star Taq Master Mix, 0.2 $\mu$ M of both reverse and forward primers (First BASE Laboratories Sdn Bhd, Malaysia), followed by 6.5 $\mu$ L RNAase-free water (Qiagen, Germany). and 4ng/ $\mu$ L template DNA. The PCR run was carried out on the Eppendorf Mastercycler Nexus Gradient machine (Eppendorf, Germany). The thermal cycler was programmed for 15 minutes at 95°C for initial denaturation followed by 35 cycles of 1 minute at 94°C for denaturation, 1 minute at 50°C for annealing, 1 minute at 72°C for extension and 10 minutes at 72°C for the final extension. The PCR products were analysed on a 1.0% agarose gel made with 1xTris Acetate-EDTA (TBE) buffer (Axon Scientific Sdn Bhd, Malaysia), agarose powder (Axon Scientific Sdn Bhd, Malaysia) and SYBR safe stain (Invitrogen, USA). Gel electrophoresis was (Bio-Rad Laboratories Inc., California) run at 80V for 60 minutes. The HLAG1 sample from PPUKM samples was used to optimise the initial PCR reactions to obtain the best results. Optimisation was done systematically by changing the annealing temperature, primer concentration and DNA concentration in the reaction mixture, as shown in Figure 3.5 on the next page. Optimised PCR products were verified by sending them for sequencing to First Base Laboratories Malaysia Sdn. Bhd.

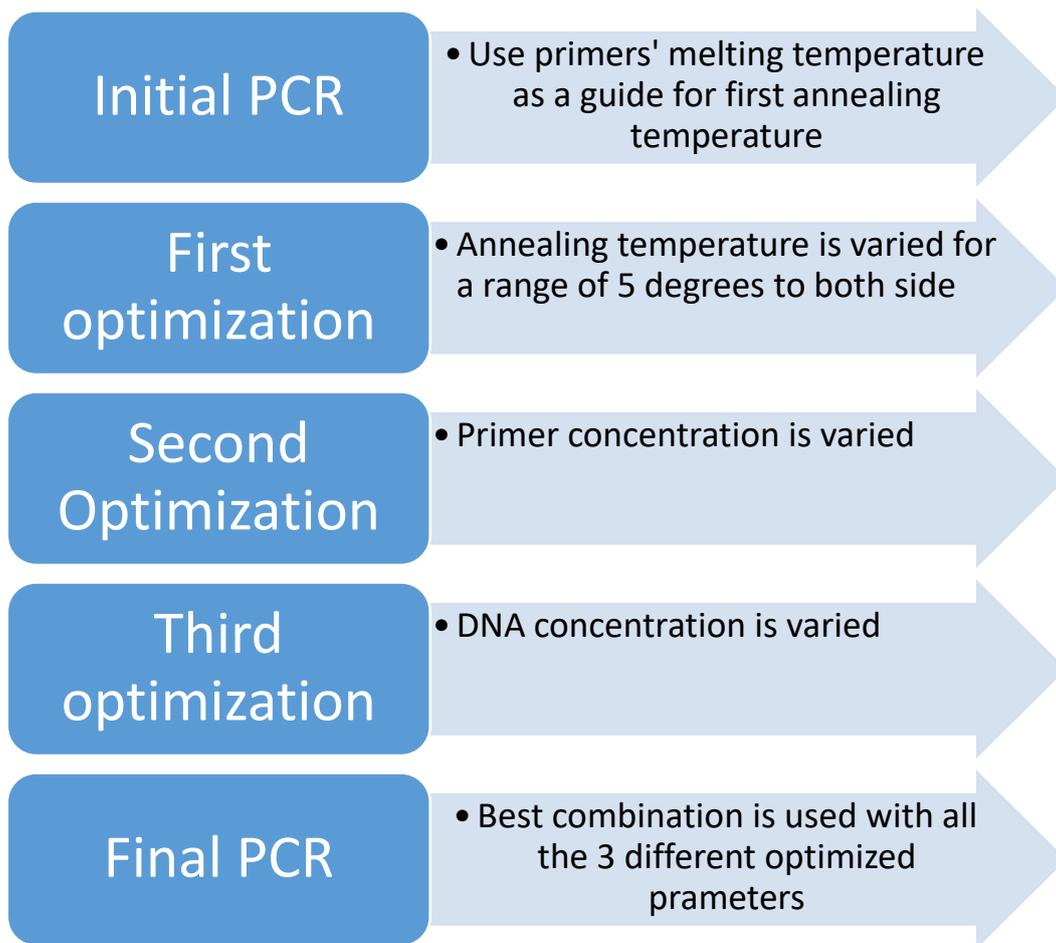


Figure 3-5 Flow of step-wise PCR optimisation done for all primers designed.

### 3.6.2 Multiplex PCR with IC

In order to check optimum condition and amplification by PCR, an Internal Control (IC) was used together with primers P2 and P3 separately. The first set of multiplex reaction with P2 and IC primers was named I2 and the second set with P3 and IC was named I3 multiplex. However, multiplex PCR was very complex and unstable, thus a lot of optimisation was needed to ensure proper targeting of each exon and the internal control together. Addition of two sets of primers in one reaction necessitates full optimisation of the annealing temperature, DNA and primer concentrations again. DMSO was further added as additive to the multiplex PCR reactions to increase stability of the two primers. The same reagents were used as previously mentioned in the standard PCR protocol. Final optimised conditions for I2 and I3 were, 47°C and 50°C for the annealing temperature respectively, followed by 6 ng/μL of DNA, a P2 and P3 concentration of 0.1 μM, an IC concentration of 0.3 μM and 0.5 μL of 100% DMSO as additive. The thermal cycler was programmed for 15 minutes at 95°C for initial denaturation followed by 35 cycles of 1 minute at 94°C for denaturation, 1 minute at 47/50°C for annealing, 1 minute at 72°C for extension and 10 minutes at 72°C for the final extension. The PCR products were analysed by gel electrophoresis on a 1.0% agarose gel run at 80V for 60 minutes with a 100bp ladder, as previously described for standard PCR.

### 3.6.3 PCR Clean-up

The QIAquick PCR Purification kit was used to purify the PCR products before being sent to Sanger sequencing (Qiagen, Germany). PCR reactions of 50 $\mu$ L were carried out to maximise amount of DNA product obtained. 2 $\mu$ L were used to perform a standard gel electrophoresis after PCR to check that specific amplification has taken place. 5 volumes of buffer BP, i.e. 240  $\mu$ L, was added to 1 volume (remaining 48  $\mu$ L) of the PCR product and mixed. One QIAquick column provided was placed into a 2 ml collection tube and labelled accordingly. The mixture from the first step was added to the QIAquick column carefully in the centre of the filter paper. Centrifugation (Sigma-Aldrich 1-14, USA) was carried out at 13,000 rpm (standard speed used throughout) for 1 minute, in order to bind DNA to the column. The flow-through was discarded and the QIAquick column was placed back in the same tube. The washing step ensued, by adding 750  $\mu$ L of Buffer PE to the column, followed by centrifugation for 1 minute, discarding the flow-through and placing the column back in the same tube. Centrifugation was carried out again for 2 minutes, in order to remove all the residual wash buffer in the column. The QIAquick column was then transferred to a clean, sterile and labelled 1.5 ml microcentrifuge tube for the elution step. 50  $\mu$ L of the Buffer EB was added to the centre of the column's membrane and was allowed to stand for 5 minutes. For increased DNA concentration, the elution volume was decreased up to 15  $\mu$ L to 25 $\mu$ L and allowed to stand for 5 minutes again. A final centrifugation step was carried out for 1 minute and the tube, with eluted DNA, was transferred to ice immediately. The DNA eluted was then analysed via a spectrophotometer (see next page) to obtain its purity and concentration. Samples which passed the purity and concentration standards set were stored in -20°C for those to be used shortly or in -80°C for prolonged storage.

### 3.6.4 Determination of DNA concentration and purity

Spectrophotometric analysis is commonly used to quantitate DNA by using a wide range of commercially available spectrophotometers. This instrument was based on the principle of ultraviolet (UV) light absorption in specific patterns. DNA samples were exposed to UV light at a wavelength of 260 nanometres (nm) and the amount of light passing through the sample was measured via a photo-detector. The amount of UV light absorbed was known to be proportional to the sample's nucleic acid concentration. A sample of higher concentration will thus emit less light, resulting in a lower detection by the photo-detector and a higher Optical Density (OD) detected. The Beer-Lambert law was used to calculate the concentration of the sample from the amount of light absorbed. Another method used by spectrophotometers is quantity measurement of nucleic acids with the 'A260 unit'; where one A260 unit was equal to the amount of nucleic acid in 1 mL and with an OD of 1. Therefore, one A260 unit was equal to around 50µg/mL of dsDNA (double stranded). These calculations were automatically done on the specific software provided with the spectrophotometer and the DNA concentration was displayed in ng/µL. Optimum concentration range used for the extracted samples were 1000 ng/µL to 3000ng/µL, where enough DNA was present for multiple experiments, as well as extra sample storage. The sample purity was determined by the 260 nm: 280 nm ratio. Pure DNA samples will have a ratio range of 1.8-2.00 as seen throughout literature. Purity and concentration were compiled for all the 290 DNA samples collected (**Appendix A**), as well as for cloning procedures mentioned in later sections (Sambrook, 2001).

The Epoch Microplate spectrophotometer (BioTek, USA) together with the Gen5 software (BioTek, USA) were used to determine each DNA sample's concentration and purity in this study. Absorbance is measured at 260 nm and 280 nm, while purity is estimated from the Optical Density (OD) ratio of 260/280nm. 2µL of samples or blanks were placed in the available Take3 plate for reading. All samples and blanks were loaded in duplicates. On the Gen5 software, the option New plate is selected, followed by the option DNA quantification for dsDNA. The first two wells were always filled with distilled water replicates, followed by unknown samples in duplicates. Duplicates were done to ensure a proper reading, especially when detection may be hindered by air bubbles, misplacing the sample drop or other

factors. The average of the duplicate readings will be calculated to give the final result. Before loading the samples, the Take3 plate was carefully cleaned with kimwipes (Kimberly-Clark, Texas) and 70% alcohol, in order to remove any traces of contamination. The latter step was repeated after use as well, to keep the plate clean and contamination free.

## 3.7 Cloning of the positive control

### 3.7.1 Introduction

Molecular cloning was created to allow multiplication of organisms, cells or DNA fragments. Cloning of a segment of DNA was performed in this section, with four essential steps. The cloning strategy started with fragmentation of the DNA strand, ligation of the desired sequence in a cloning vector, transfection of the cloned plasmid into cells of choice and screening/selecting the successful clones. Once the successful clones were obtained, they were cultured in agar and broth in order to increase the number of clone copies. These copies were then extracted, analysed and stored for future use in experiments (Sambrook, 2001).

A PPUKM patient previously screened/sequenced for HLA-B alleles was confirmed to have the HLA-B\*58:01 allele, and was used as a positive control in this study. However, a limited amount of extracted DNA was available from UMBI's biobank, thus the need for cloning arose. The patient could not be contacted again for withdrawal of more blood for more DNA sample. This positive control was used as reference in the future HRM method, in order to compare unknown samples to its DNA sequence as a method of screening. Thus an infinite amount of the positive control needed to be produced for this project and all future screenings to be done.

Cloning was done by using the Qiagen PCR Cloning Plus Kit (Qiagen, Germany), with several rounds of optimisation to the given steps. This kit had several advantages and all the needed reagents provided for quick and efficient cloning to take place. The transition from PCR product to plated cells took only 40 minutes and the kit provided a ready-to-use Ligation Master Mix, *E. coli* competent cells, pDrive cloning vector and Super Optimal broth with Catabolite repression (SOC) medium. Ligation takes 30 minutes, followed by transformation and plating within 10 minutes. This kit was tested against other competitors in the market and proved to be faster, easier and highly efficient cloning. For a simple comparison, topoisomerase-mediated cloning kits took more than 70 minutes from PCR product to

plated cells. TA-based cloning kits took more than 5.5 hours and conventional ligase cloning took more than 7.5 hours to reach the plating stage. The pDrive vector provided numerous handy features for subsequent cloning analysis, namely, several unique restriction enzyme recognition sites, universal sequencing primer sites and promoters for in vitro transcription. Moreover, both ampicillin and kanamycin selection is possible, along with blue/white screening for recombinant colonies. The Ligation Master Mix was premixed, with all conditions ready for optimized hybridization and cloning. The provided Qiagen EZ Competent cells did not require the usual time-consuming recovery incubation time in SOC medium for high efficiency transformation ( $>10^8$  colony forming units (CFU) per microgram DNA). The procedure was simple to follow; where the PCR product is directly mixed with the pDrive Cloning Vector and Ligation Master Mix and incubated. This ligation mixture was then added to the competent cells, for transformation to take place, and immediately plated onto agar plates. However, a few tweaks and optimisation steps were also done along the way and this will be fully explained below in the cloning protocol (Qiagen, Germany)

### 3.7.2 Preparation of cloning materials and reagents

#### 3.7.3 LB agar preparation

LB agar was always prepared in the end, after all the other reagents were prepared and diluted appropriately. These reagents (ampicillin, IPTG and X-Gal) may be prepared the day before and stored accordingly. However, LB agar was always autoclaved in the morning, poured in the afternoon and stored at 4°C on the same day after solidification. Pouring of agar on the same day prevents contamination and degradation of its content from keeping it too long in the oven. The 500 mL Schott bottle (Merck, Germany) to be used for LB agar preparation was autoclaved, filled with distilled water to remove all contaminants. The water was poured out right before agar preparation. The same set of Schott bottles were always kept for LB agar preparation during the whole cloning phase, in order to avoid contamination.

15.0 g of LB agar powder (First Base Laboratories, Malaysia) was added to a clean, empty, sterile 500 ml Schott bottle. The bottle was closed, an autoclave tape stuck on top of the cap and was vigorously shaken to mix the contents. After autoclave, the LB agar bottle was placed in a hybrid oven at 60°C in order to keep it in liquid state. The laminar flow cabinet (Esco Micro Private Limited, Singapore) was wiped down with 70% ethanol. All the three reagents were thawed out, vortexed and placed in a ice box in the laminar flow cabinet. A full, unopened, sterile, clean stack of petri dishes were placed in the cabinet and labelled on the bottom with the date and type of plate to be prepared. The Schott bottle with LB agar was removed from the oven, wiped with 70% ethanol and placed in the laminar flow cabinet for cooling down. The cooling method used here was by rolling the bottle horizontally on the cabinet's surface slowly. This has to be done slowly to avoid the formation of bubbles. The LB agar was cooled down until body temperature, where you can hold the bottle in your hands without being burned. This has to be monitored carefully as over cooling may lead to clumps being formed in the agar bottle. Once the agar was at the right temperature, add the ampicillin, IPTG and X-Gal. Mix by rolling the bottle horizontally for 10 times and proceed to pouring the agar plates.

The agar was poured swiftly and quickly into the centre of the petri dishes and filled up to half the quantity of the dish. In cases where bubbles were seen in the plate, the latter was gently swirled until the bubbles migrated to the dish's outer border. A sterile, autoclaved 10  $\mu$ L tip may also be used to pop the bubble quickly. Once all the plates were poured, they were arranged in order and the lid slanted on them, covering half of the plate. The plates were left to solidify for a minimum of 30 minutes and a maximum of 1 hour. After solidification, the plates were closed, stacked upside down and placed in the plastic that initially contained the petri dishes. These plates were stored at 4°C-8°C and used within 1-2 weeks. Another batch of LB agar plates containing ampicillin only (100 $\mu$ g/mL LB agar) was also prepared for subcultures to be done after successful transformation.

### 3.7.4 Cloning protocol

#### 3.7.4.1 Amplification of the positive control's exon 2 and 3

The positive control's DNA obtained was first amplified with the two different primers optimized separately, i.e, P2 and P3, followed by gel electrophoresis for verification .The PCR products were purified with the QIAquick PCR Purification kit (Qiagen, Germany) and sent for Sanger sequencing (First Base laboratories, Malaysia) for verification of the desired products. Once this step was cemented, preparations for ligation started.

#### 3.7.4.2 Ligation

Ligation started with the calculation of the amount of PCR product needed. A molar ratio of 10 times molar excess was used as recommended by the kit's protocol. Both P2 and P3 primers gave products at 149-250 bp, thus 26 ng of the purified PCR products were used for the first step. The 2x Ligation Master Mix, pDrive Cloning Vector DNA and distilled water provided were thawed and immediately placed on ice. Solutions need to be properly mixed to avoid uneven distribution of their components. A ligation-reaction mixture was then prepared according to Table 3.2 below. The ligation mixture is added last to the labelled PCR tube used for this step.

Table 3-2 Detailed volume of the ligation mixture components to be prepared.

Component	Volume/reaction ( $\mu\text{L}$ )
<b>pDrive Cloning Vector (50ng/<math>\mu\text{L}</math>)</b>	1
<b>PCR product</b>	1-4
<b>Distilled water</b>	Variable
<b>2x Ligation Master Mix</b>	5
<b>Total volume</b>	<b>10</b>

Once the ligation mixture was added to the tube, it was briefly and gently mixed by pipetting it up and down for 10 times. Incubation was then done for 2 hours at 16°C in the Eppendorf Mastercycler Nexus Gradient machine (Eppendorf, Germany) which was previously used for PCR. Ligation time was increased to 2 hours after optimisation in order to generate a 2-3 fold increase in recombinants. Incubation was followed immediately with transformation.

### 3.7.4.3 Transformation

#### 3.7.4.3.1 Transformation preparation:

This step will be done with the provided Qiagen EZ Competent Cells, stored at  $-80^{\circ}\text{C}$ . These cells were extremely sensitive to temperature and mechanical stress and should not be thawed at any point prior to transformation. When being used, these cells need to be mixed by gentle flicking and no rough handling like vortexing or pipetting. SOC medium was thawed and warmed to room temperature before use. The latter was stored at  $-15^{\circ}\text{C}$  to  $-30^{\circ}\text{C}$  after use. Fresh LB agar plates were prepared before transformation, containing ampicillin ( $100\mu\text{g}/\text{mL}$  LB agar), IPTG ( $50\mu\text{M}$ ) and X-Gal ( $80\mu\text{g}/\text{mL}$ ).

#### 3.7.4.3.2 Transformation protocol

One tube of Qiagen EZ Competent Cells was left to thaw on ice. This tube was observed carefully and once thawed, transformation must be done immediately. The SOC medium was thawed and warmed to room temperature.  $2\mu\text{L}$  of the ligation-reaction mixture was added to one tube of Qiagen EZ Competent Cells, mixed gently by slow flicking and incubated on ice for 5 minutes. The tube was heated to  $42^{\circ}\text{C}$  in a water bath for 30 seconds and then incubated on ice for another 2 minutes.  $250\mu\text{L}$  of SOC medium was added to the tube and gently flicked. This mixture was immediately plated onto the previously prepared LB agar plates. This mixture was pipetted onto the agar plates in two different volumes of  $50\mu\text{L}$  and  $75\mu\text{L}$ , and spread by using a sterile plastic spreader (Apical Scientific, Malaysia). This was the optimum volume used for good colony formation throughout the plate after several optimisation steps. The plates were incubated at room temperature until the transformation mixture was absorbed in the agar and then sealed with parafilm (Apical Scientific, Malaysia). The plates were then incubated upside down at  $37^{\circ}\text{C}$  for 16 hours, followed by 4 hours of cold incubation at  $4^{\circ}\text{C}$ . The cold incubation enhanced the blue color development and helped in blue colonies differentiation from the white colonies. This whole process, starting from ligation, to transformation, plating and blue-white colony formation was depicted in Figure 3.6 below.

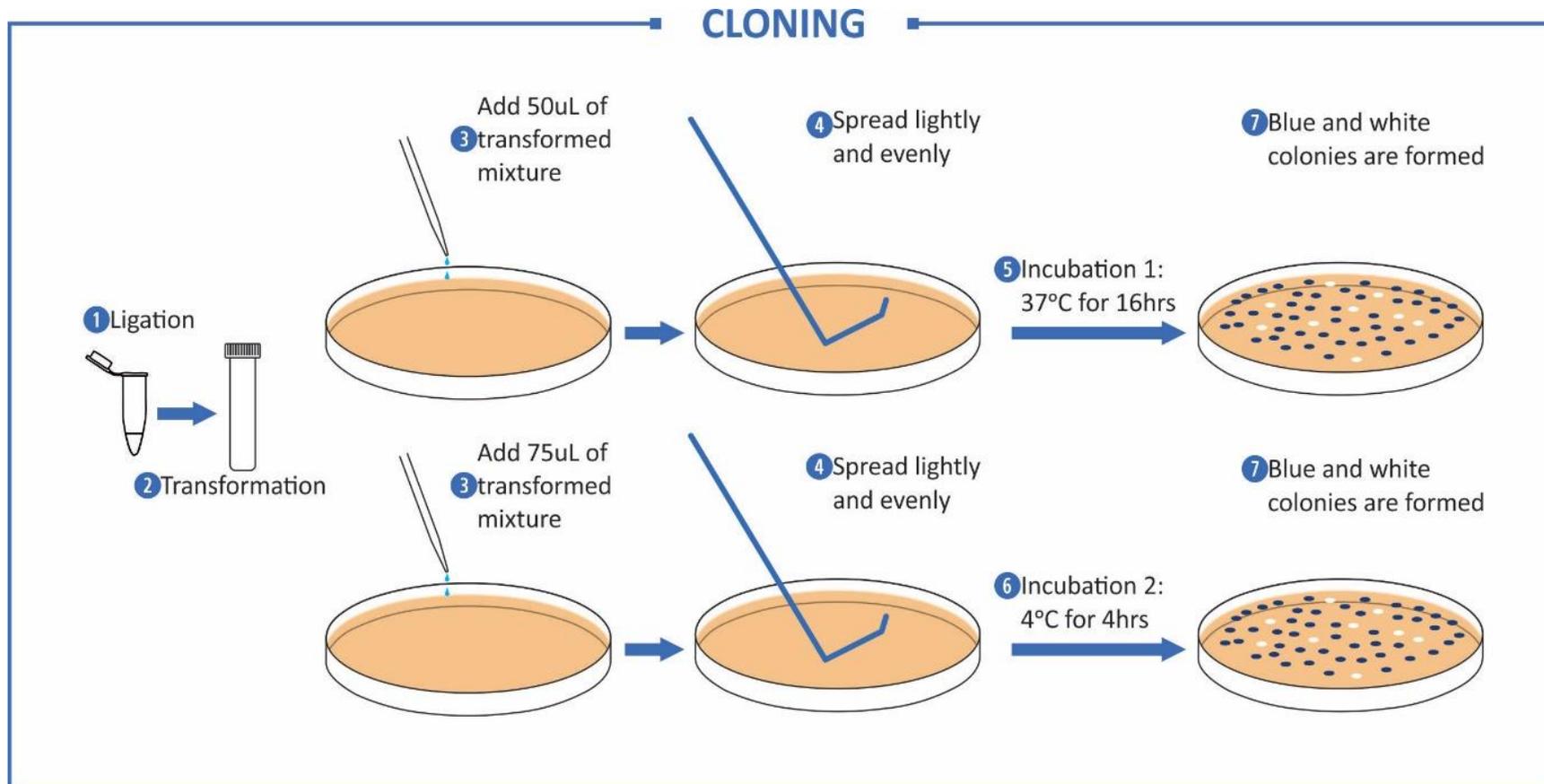


Figure 3-6 First part of cloning process depicted, starting from ligation, transformation, plating, spreading and finally blue-white colony formation. (DIY)

### 3.7.4.4 Subculture of white colonies in LB agar plates with ampicillin

The white colonies obtained in the different transformation plates were sub cultured into LB agar plates containing ampicillin only (100µg/mL LB agar). The procedure starts with wiping down the laminar flow cabinet with 70% ethanol and exposure to UV light. The transformed plates (in its plastic bag) along with all the equipment needed were all wiped with 70% ethanol and placed in the cabinet. Inoculating loops were sterilized by using the loop sterilizer (Merck, Steri 350) in the cabinet. Using a cool, sterile inoculating loop, one white colony was touched and streaked onto LB agar plates with ampicillin. The quadrant streak technique was used, to allow a sequential decrease in the amount of bacteria present throughout the plate (shown in Figure 3.7 below). This resulted in a few isolated colonies on the last two quadrant. The loop was sterilized and cooled down in between different quadrant streaking. Once the full plate's surface was streaked properly, it was sealed with parafilm, inverted and incubated at 37°C for 16 hours.

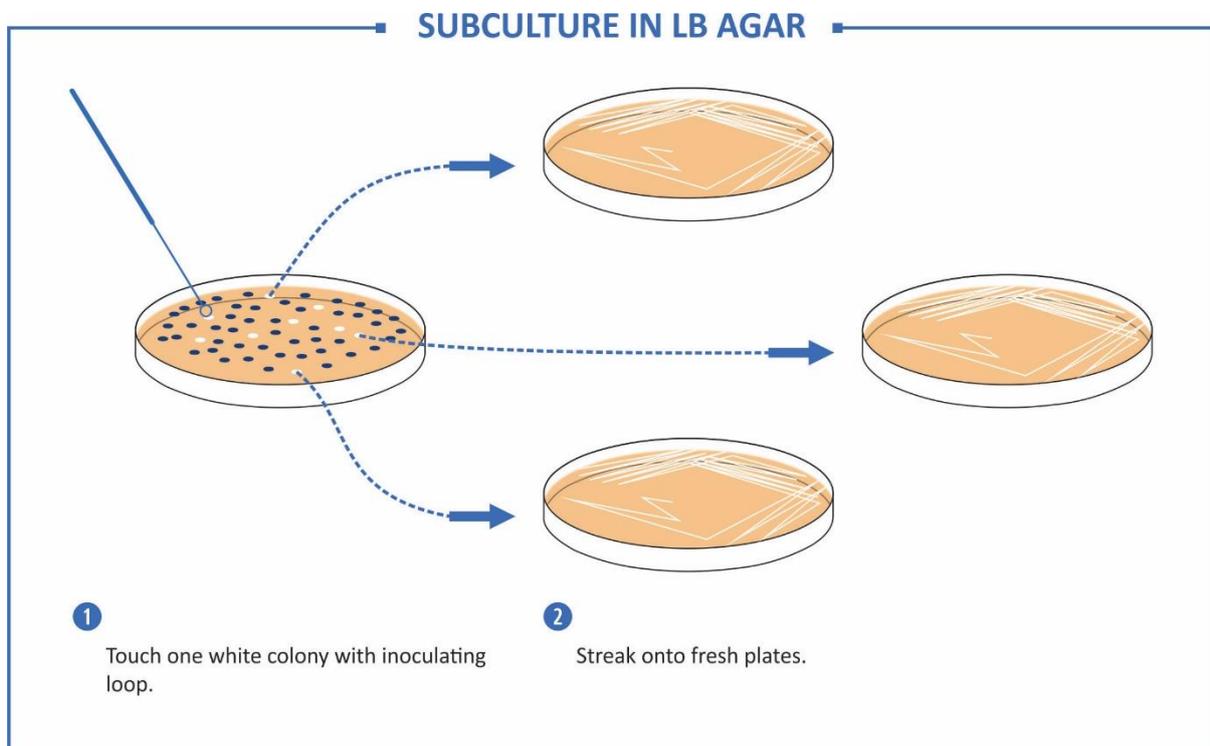


Figure 3-7 Streaking of the white colonies by using the quadrant streak technique. (DIY)

### 3.7.4.5 Subculture of white colonies in LB broth

#### 3.7.4.5.1 Subculture preparation:

This step was done to increase the number of plasmids for subsequent plasmid extraction and analysis. LB broth was prepared in a sterile 500 mL Schott bottle by adding 12.5 g of LB broth powder and 500 mL of distilled water. The bottle was autoclaved and stored at 4°C before use. Subculture was done in five 250mL conical flasks, which were initially washed, dried and sealed on top with aluminium foil. These were sent for autoclave and dried overnight in a hybrid oven before use.

#### 3.7.4.5.2 Subculture protocol:

Subculture was done on the bench, with a lighted Bunsen burner to avoid contamination. The bench and all the equipment to be used were wiped with 70% ethanol. The conical flasks were placed around the lighted Bunsen burner and 50 mL of LB broth were added to each of them by using a sterile measuring cylinder. Five conical flask were used each time to extract a suitable amount of plasmid for downstream use and application. 50µL of the ampicillin stock previously prepared was added to each conical flasks and swirled multiple times for mixing. The aluminium foil cover was always left on top of the conical flasks to prevent contamination. A white colony was touched on the agar plate by using a sterile inoculating loop. This step was then repeated for the next four conical flasks but different white colonies were used from different plates, in order to increase the chances of successful cloning. The aluminium foil was placed loosely enough to allow some air circulation for proper bacterial growth. The flasks were incubated for 16 hours in a shaker incubator (PMI-Labortechnik, WIS-20R) at 37°C and at a speed of 180 rpm. After incubation, the optical density was read to check for optimal bacterial growth. All the subculture steps were depicted in Figure 3.8 below.

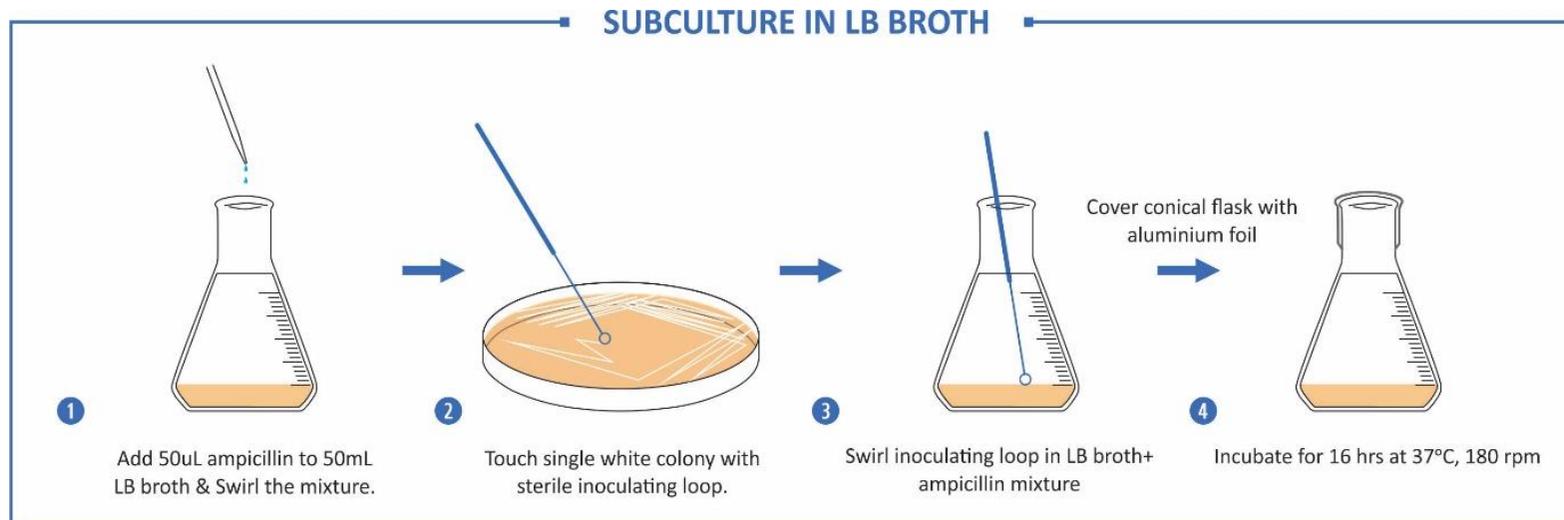


Figure 3-8 Subculture of the successful white clones in LB broth for further increase in copy number and subsequent extraction of plasmids. (DIY)

### 3.7.4.5.3 Measuring the Optical density of LB broth cultures:

This step was performed after every single LB broth subculture to ensure proper bacterial growth in optimal conditions. Optical density (OD) is measured by using a spectrophotometer and used to get a measure of concentration of bacteria in suspension. This method uses the principle of light scattering as it passes through a cell suspension. A greater scatter indicates the presence of more bacteria or other contaminants. Bacteria at their mid-log phase of growth were measured at a wavelength of 600nm (OD600). An optimum O.D reading of 0.5 to 0.8 must be respected for each subculture with this specific spectrophotometer (D'Ans et al., 1972). This optimum reading was obtained after several *E. coli* subcultures in LB broth.

#### 3.7.4.5.3.1 OD measurement protocol:

Five plastic cuvettes were cleaned with distilled water and wiped with kimwipes. Care must be taken with the cuvette quality, cleanliness and the absence of bubbles of air in the cuvette. The spectrophotometer was switched on and calibration allowed for a few minutes. The wavelength was adjusted to 600nm. A blank reading was always performed first, with 3mL of LB broth only. Fill clean cuvettes with 3mL of the different bacterial subcultures and proceed to reading them one after the other. Note all the OD down in a notebook. The optimum OD for all subcultures must fall in the range of 0.5-0.8.

#### 3.7.4.5.4 Purification of plasmid from LB broth

Plasmids were purified from the LB broth by using the PureYield™ Plasmid Miniprep system kit (Promega, USA). For the first time use of the kit, these two steps must be followed; the lysis buffer was incubated at 37°C for 30 minutes and shaken well afterwards. 95% ethanol was added to the column wash solution as indicated on the bottle.

##### 3.7.4.5.4.1 Extraction protocol:

3.0 mL of bacterial culture grown in LB broth was transferred to a 15 mL falcon tube and pelleted down at 5000rpm for 15 minutes. The supernatant was discarded and 600 µL of 1x TE buffer was added to the pellet. The pellet was resuspended numerous time with a pipette to obtain a homogenous solution. This mixture was then transferred to a 1.5 mL microcentrifuge tube. 100 µL of Cell Lysis Buffer was added to the tube and mixed by inverting it 10 times. A color change from opaque to clear blue must be observed for complete lysis to occur. This step must be completed within 2 minutes, in order to avoid excessive lysis. After proper lysis, 350 µL of cold Neutralization solution (4°C-8°C) was added to the tube and mixed thoroughly by inverting the tube. A centrifugation step was then carried out at 10,000 rpm for 3 minutes. All centrifugation steps were carried out at 10,000 rpm. A PureYield™ Minicolumn was placed in a PureYield™ collection tube and labelled with the sample code. A maximum of 900 µL of the clear supernatant was transferred via pipette to the minicolumn. The cell debris pellet must not be disturbed to avoid contamination and a decrease in purity of the extracted plasmid. Centrifugation was carried out again for 15 seconds. The flowthrough was discarded and the minicolumn placed back into the collection tube. 400 µL of the Column Wash Solution was added to the minicolumn and centrifuged for 30 seconds. The microcolumn was then transferred to a sterile 1.5 mL microcentrifuge tube and 15µL of Elution Buffer was added to the centre of the filter paper. The tube was incubated for 10 minutes at room temperature and centrifuged for 30 seconds. The plasmid DNA eluted was vortexed briefly and stored on ice. The plasmid DNA concentration was measured by Nanodrop analysis as previously described in **Section 3.6.4**.

### 3.7.4.6 Validation of cloning step/Plasmid analysis:

#### 3.7.4.6.1 Restriction digestion of plasmid

Restriction enzymes (REs) were endonucleases that recognize short, DNA sequences and cleave them at specific recognition sites. Optimal enzymatic activity and fidelity depends on proper storage, assay optimisation, temperature, pH, salt concentration and ionic strength. Double restriction digestion was performed on the extracted plasmid based on the provided restriction sites on the pDrive cloning vector (Qiagen, Germany). BamHI and XhoI (Promega, USA) were used for double restriction digestion in order to precisely excise the cloned target. A MULTI-CORE™ buffer from the digestion kit was used as it provides broad compatibility and higher activity with many REs, after optimisation. Plasmid DNA concentration and purity must be carefully calculated and monitored before digestion. Sterility of reagents and all tools need to be maintained properly. To protect REs from extreme conditions and maintain their sensitivity, they must be stored in freezers, were always placed on ice when in use and were added to the digestion mixture last. All reagents were thoroughly thawed, mixed and centrifuged before use. After the first step, showed in **Table 3.3** below, all reagents were mixed by continuous pipetting. The final mixture containing the REs (after step 2 in **Table 3.3**) must be mixed by gentle pipetting to avoid enzyme inactivation.

#### 3.7.4.6.1.1 Setting up the Restriction Enzyme Digestion:

A digestion mixture is set up, according to the details in Table 3.3, in a sterile and labelled 1.5mL microcentrifuge tube and mixed gently by pipetting up and down. The tube is then centrifuged for a few seconds, sealed with parafilm on top, placed in a tube holder and incubated for 4 hours in a water bath at 37°C. Experimental controls were also used, especially during optimisation in order to understand and explain inconsistencies in results. Controls used in this experiment were the purified cloned target DNA and the full uncut plasmid, in order to compare all the fragment sizes. The next step consists of analysing the digestion mixture by gel electrophoresis, along with the two controls. 9µL of the digestion mixture is added to another sterile tube, along with 1µL of 10x Blue Juice, and mixed by gentle pipetting. The same amount were added for the two controls and loaded in a 1% agarose gel with bigger wells.

Table 3-3 Components of the restriction digestion mixture to be added in two steps.

Step number	Reagent	Volume (µL)
1	Sterile deionized water	~15.0 (varies)
	RE 10 x buffer	2.0
	Acetylated BSA (10µL/µg)	0.2
	DNA (1µg/µL)	-
2	Xho1 (10u/µL)	0.5
	BamH1	0.5
Final volume		20.0

#### 3.7.4.6.1.2 Sanger sequencing

Plasmids were sent for Sanger sequencing to First Base Laboratories Sdn Bhd, with M13F (-20) as the sequencing primer.

## 3.8 High Resolution Melt (HRM) method

### 3.8.1 Introduction

The HRM screening done in this research used the detailed melting profile of a positive control as reference and screened numerous samples alongside in one run. One HRM run could accommodate a maximum of 15 samples, along with one positive control. Any sample which matched the melting profile of the positive control showed presence of the targeted allele. Small mutations will be translated into a change in melt curve shape or a shift of the whole HRM melt profile.

### 3.8.2 HRM mix preparation

The whole HRM mix preparation stage was done in a biological safety cabinet (Esco Technologies Inc, USA), in order to avoid all kinds of contamination. A standard 20 $\mu$ L HRM reaction was prepared by adding 10 $\mu$ L of 2x SensiFAST HRM mix (Bioline, USA), 0.4 $\mu$ M of forward and reverse primers, 5 $\mu$ L of RNAase-free water (Qiagen, Germany) and 500ng of DNA template. A master mix was prepared for 3 reactions for every DNA sample to be screened, i.e. duplicates for one DNA sample with its non-template control (NTC) (shown in Table 3.4 below). This was done to minimize variations, errors and contamination as the HRM method has a high sensitivity. This master mix was vortexed and centrifuged briefly and pipetted in the wells of the eco plate (Illumina, USA). Once all the wells were filled with the required amount of HRM master mix, DNA samples were added last to each well. The plates were sealed by first placing the eco-seal on top of the eco-plate without touching its sticky surface. Once the seal was in place, a sterile ruler was used to seal the edges tightly by running it along all the edges. The plate must always be kept upright and undisturbed on a flat surface. In cases where the eco plate was shaken and drops accumulated on the eco-seal on top, it must be discarded as the sample will be contaminated by the glue present on the eco-seal. The eco-plate was then centrifuged at 4000rpm for 2 minutes and placed in the Eco Real-time PCR machine for the HRM run (Illumina, USA).

Table 3-4 Breakdown of the HRM reaction contents with their detailed volume used and the final concentration needed.

HRM components	Volume (µL)	Final concentration
2x SensiFAST HRM mix	10.0	1x
Forward primer (10 µM)	0.80	400nM
Reverse primer (10 µM)	0.80	400nM
DNA sample	Varies	500 ng
RNAase free water	Varies	-
<b>Total</b>	20.0	-

The cycling conditions for the initial PCR and following HRM step were listed down in Table 3.5.

Table 3-5 The fully detailed cycling steps for the initial PCR and following HRM step done for all samples during HLA-B\*58:01 screening.

HRM/PCR Stage	Cycling steps	Temperature (°C)	Time (seconds)
<b>PCR stage</b>	Polymerase activation	95	120
	Denaturation (35 cycles)	95	5
	Annealing (35 cycles)	45/60	10
	Extension (35 cycles)	72	15
	Final extension	72	600
<b>HRM stage</b>	1 <sup>st</sup> step	95	15
	2 <sup>nd</sup> step	55	15
	3 <sup>rd</sup> step	95	15

### 3.8.3 HRM run set up

The Eco Real Time machine and accompanying laptop were turned on 30 minutes before the run, while starting preparation for the HRM mix, as it required a setup time before any run. Once the machine was ready, the blinking red lights on the top left corner turned green and the Eco Software opened indicated that the instrument was connected, on the bottom right corner. The options HRM was selected, followed by the setup of the thermal profile and according to the previously mentioned cycling conditions in Table 3.5. The third step was to design the plate layout. Each sample, including the positive control, were in batches of 3 wells, with the first one being the NTC, and followed by 2 wells with the sample/positive control labelled as U for unknown. This was done until the whole eco-plate was labelled, all under one assay type. The sample codes were added to their respective wells and the same was done for the first set of trio wells for the positive control. All the tabs were double-checked, the eco-plate inserted in the machine and the run started. A live monitor run will then be displayed to check the melting stages for all the samples. After the run was completed, the result will be checked and analysed using the Eco Study software (Illumina, USA).

### 3.8.4 HRM evaluation

#### 3.8.4.1 Specificity, Sensitivity, Positive predictive value and Negative predictive value

The basic measures to quantify the diagnostic accuracy of a test include sensitivity and specificity. The sensitivity of a diagnostic test quantifies its ability to correctly identify subjects with the disease condition. The specificity is the ability of a test to correctly identify subjects without the condition. The PPV and NPV are the other two basic measures of diagnostic accuracy. They are related to sensitivity and specificity through disease prevalence ( $\Pi$ ). The PPV is the probability that the disease is present given a positive test result and the NPV is the probability that the disease is absent given a negative test result (Wong & Lim, 2011). In this study, the HRM method was designed to screen for the HLA-B\*58:01 allele in 145 gout patients and 145 healthy volunteers. The Sanger sequencing method was used to validate the HRM results in this study and is hence used as the gold standard. The Sanger sequencing method was done on the 28 patients with signs of SCARs and ADRs and on 14 healthy volunteers. To determine the validity and accuracy of the HRM method designed, its sensitivity and specificity needs to be measured. This is best illustrated by using a two-by-two table, where the results from a new diagnostic test (HRM) is compared to a gold standard (Sanger sequencing). In cell a, we enter those in whom the test in question (HRM) correctly showed the presence of the HLA-B\*58:01 allele (as determined by the gold standard). In other words, the HRM result is positive, as is the gold standard (Sanger sequencing). These are the true positives (TP). In cell b, we enter those who have positive results for the test in question (HRM) but do not have allele according to the gold standard test (Sanger sequencing). The newer test (HRM) has wrongly diagnosed the disease: These are false positives (FP). In cell c, we enter those who have the allele on the gold standard test (Sanger sequencing) but have negative results with the test in question (HRM). These are false negatives (FN). In cell d, we enter those who do not have the allele as determined by the gold standard test (Sanger sequencing) and are also negative with the newer test (HRM). These are true negatives (TN). (Wong & Lim, 2011).

Table 3-6 Calculation of sensitivity and specificity in a two-by-two table, where TP stands for True positive, FP for False positive, FN for False negative and TN for True negative. (Wong & Lim, 2011).

	Allele present	Allele absent
Test positive	a (TP)	b (FP)
Test negative	c (FN)	d (TN)
Calculation	Sensitivity: $a/(a+c)$	Specificity: $d/(b+d)$

Hence,

$$\text{Sensitivity} = a/(a+c)$$

$$\text{Specificity} = d/(b+d)$$

$$\text{Positive predictive value} = a/(a+b)$$

$$\text{Negative predictive value} = d/(c+d)$$

### 3.8.4.2 Statistical Analysis

Fisher's exact test and One-way analysis of variance (ANOVA) were used in this study to analyse the data generated by using the SPSS statistics software version 22 for Windows (IBM Corp., New York, USA). One-way analysis of variance (ANOVA) was used when there are more than two independent groups being compared the one-way ANOVA is used if the parametric assumptions are satisfied—that is, interval-scale variable approximately normally distributed. Fisher's exact test was used to determine if there are non-random associations between two categorical variables. It therefore examined the significance of the association (contingency) between the two kinds of classification. Fisher's exact test was used here as it is more accurate than the chi-square test or G-test of independence when the expected numbers are small (McCrum-Gardner, 2008). Hence, with all the data generated in the study, seven different hypotheses were formulated for statistical analysis, as shown in Table 3.7 below. All the data collected in this study was converted to numerical data for statistical analysis. The null hypotheses stated that there is no difference or association between the data being compared, while the alternative hypothesis states that there is a significant difference.

Table 3-7 Different statistical tests used paired to the different formulated hypotheses to be tested.

Statistical tests used	Hypothesis formulated
Fisher's exact test	Is there an association between the HLA-B*58:01 allele's distribution between the healthy volunteer cohort and the gout cohort.
	Is there an association between HLA-B*58:01's presence and the different ethnic groups in the gout cohort.
	Is there an association between HLA-B*58:01's presence and the different ethnic groups in the healthy volunteer cohort.
One-way ANOVA	Is there a link between HLA-B*58:01's presence and the different types of allopurinol-induced hypersensitivity reactions seen.

## 3.9 Next Generation Sequencing

### 3.9.1 Introduction

The whole NGS run was broken down into three different sections, namely; sample selection, sample DNA amplification and preparation and finally library preparation. These processes will be broken down into further steps and explained with the help of Figures/flowcharts in the following subsections. The general flow of the NGS method was depicted in Figure 3.9 below.

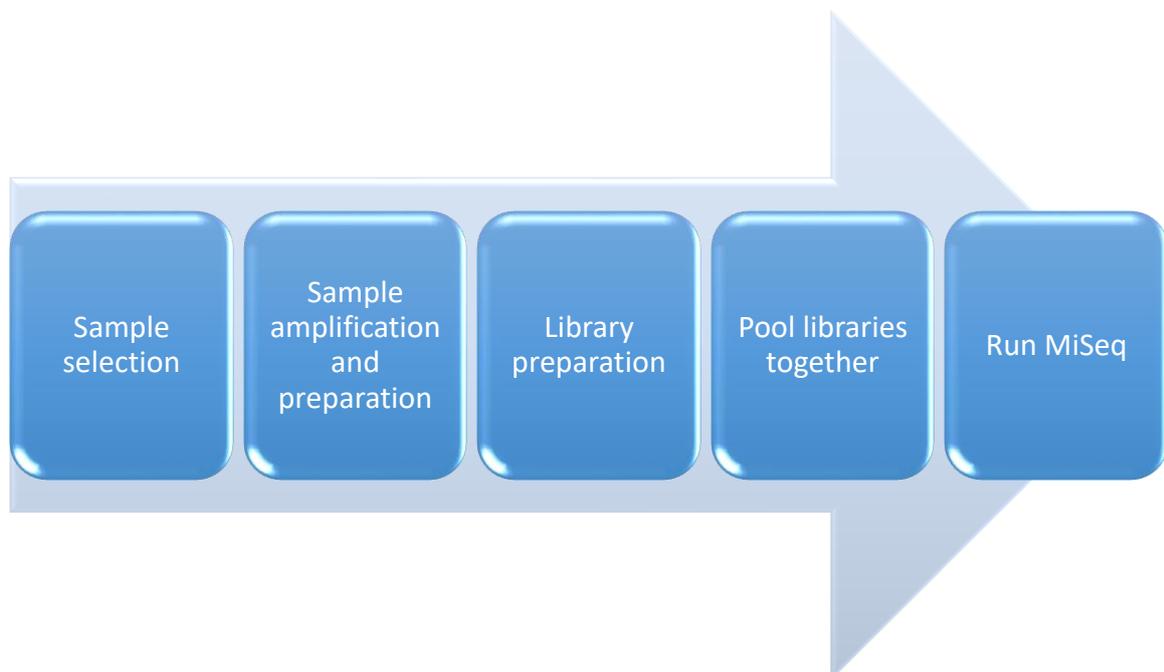


Figure 3-9 General flow of the NGS method, starting from samples collection to the end step of running all the samples on the MiSeq machine.

### 3.9.2 NGS samples selection

Due to the limited number of gout samples and the slow progression of sample collection in 2017, only six gout patients were included in this NGS run. Only these six samples showed signs of SCARs or other ADRs in the whole batch of 60 gout samples collected. The positive control used for all the HRM screening was included in those six samples, in order to check the different variants and SNPs present in it. It was considered to be more valuable to use gout samples with ADRs/SCARs in order to investigate the disease-allele-ADR association better.

### 3.9.3 NGS Sample DNA amplification and preparation

The 6 samples were amplified and prepared according to the standards for the NGS library preparation. The flow of this section was shown in Figure 3.10 on the next page. The six gout samples were processed first, followed by nine batches of 10 epilepsy samples. The whole HLA-B region was amplified, followed by validation by the gel electrophoresis method and PCR product purification. The concentration and purify of the purified PCR product were checked first by using the Nanodrop analysis method and then by using a Qubit Fluorometer for higher accuracy. The sample was then normalized to 0.2ng/μL and this was checked via the Qubit fluorometer again, before moving on to the library preparation. This whole process was depicted in the flowchart below for a simplified breakdown of steps. All the seven steps were done within 3.5 hours and timed in order to fit everything in the first half of the day (shown in Figure 3.11 on the next page). The second half of the day goes to NGS preparation as shown in the following sub-section.

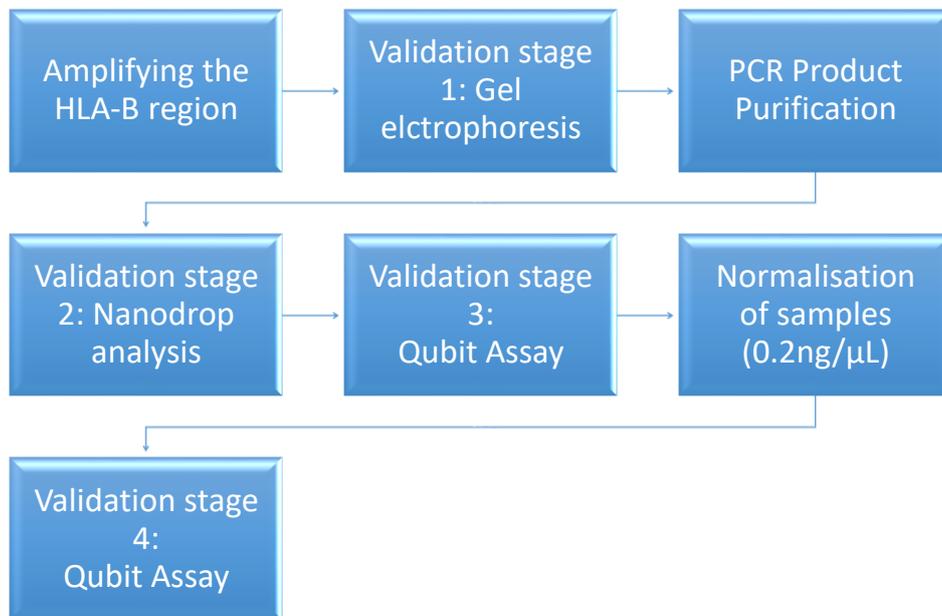


Figure 3-10 The simplified flow of the seven stages for sample amplification and preparation.

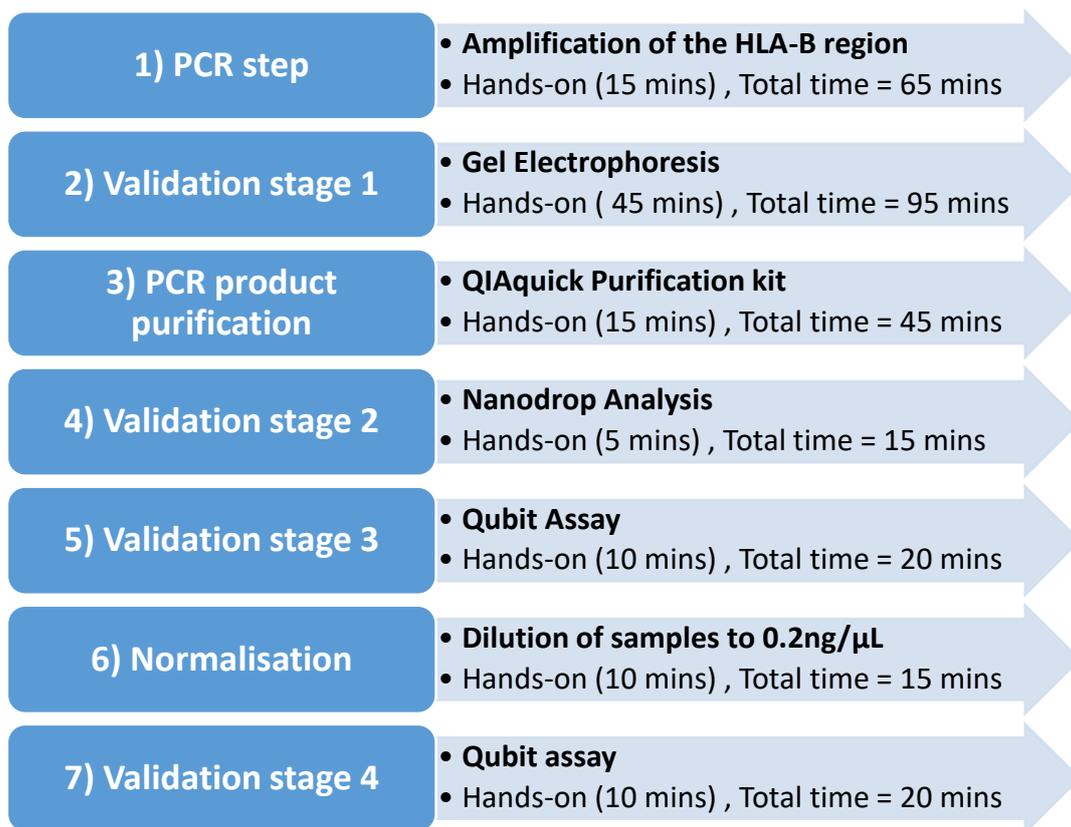


Figure 3-11 Breakdown of the steps along with their hands-on time and running time required, in order to effectively fit all of them in the first half of the NGS preparation day.

### 3.9.3.1 PCR with HLA-B primer and Gel electrophoresis

The samples were first amplified by using a specific set of primer acquired from a previous study where HLA typing was performed for the whole HLA-B region by the NGS method. Moreover, the NGS specialist from Science Vision used and validated this set of primer for a few projects before (Hosomichi et al., 2013). The primers' properties (shown in Table 3.8 below) were first analysed by using the primer designing tool on NCBI and then sent for synthesis at First Base Laboratories Sdn Bhd (Malaysia).

Table 3-8 Full details for the HLA-B forward and reverse primers used for the NGS run.

Primer	Sequence (5' to 3')	Length (bp)	Tm (°C)	GC content (%)	Self complementarity	Self 3' complementarity
<b>HLAB-F</b>	AGGTGAATGGCTCTGAAAA TTTGTCTC	27	57.6	40.74	6.00	2.00
<b>HLAB-R</b>	AGAGTTTAATTGTAATGCTG TTTTGACACA	30	56.0	30.0	4.00	4.00

The primers were first optimised sequentially on the Veriti 96 well thermal cycler (serial number 4375788) (Applied Biosystems, USA) through several steps. The annealing temperature was first determined by performing a gradient PCR with KAPA HiFi Hot Start Ready Mix PCR Kit (Kapa Biosystems Inc., USA), followed by an increase in DNA and primer concentration. The full optimised conditions for the final PCR were shown in Table 3.9 below, along with the cycling conditions in Table 3.10. The first set of PCR products were sent for sanger sequencing to check the proper targeting and amplification of the HLA-B region.

Table 3-9 Full details for the HLA-B amplification for the NGS run, including component breakdown, their volume and final concentration.

Component	Volume of reagents (μL)	Final concentration
2X KAPA HiFi HotStartReadyMix	25.0	1X
HLAB-F primer (10μM)	6.00	1.2μM
HLAB-R primer (10μM)	6.00	1.2μM
DNA template (100 ng/μL)	3.00	6 ng
PCR-grade water	10.0	N/A
<b>Total</b>	50.0	N/A

Table 3-10 Thermal cycler program for HLA-B amplification on the Veriti 96 well thermal cycler (Applied Biosystems, USA)

Step	Temperature	Duration (sec)	Cycles
Initial denaturation	95	180	1
Denaturation	98	20	30
Annealing	64	15	
Extension	72	60	
Final extension	72	240	1

Gel electrophoresis was performed for validation of the PCR step, by using a 0.7% agarose gel and TBE buffer. The gel was run at 110V for 50 minutes and the gel was checked under UV light in a gel reader. Only 5μL of the PCR product was loaded onto the gel, along with 0.5μL of Blue Juice gel loading buffer (10X) (Thermo Fisher Scientific, USA). The remaining 45μL was kept on ice for use in subsequent steps. Once the PCR product showed the correct size when extrapolated with the ladder, it was validated and sent for Sanger sequencing at First Base Laboratories Sdn Bhd.

### 3.9.3.2 Purification of PCR product

The remaining 45 $\mu$ L of PCR product was purified here, by using the QIAquick PCR Purification kit (Qiagen, Germany). Before starting, 100% ethanol was added to Buffer PE (see bottle label for volume). All centrifugation steps were carried out at 17,900 x g (13,000 rpm) in a conventional table-top microcentrifuge at room temperature. 5 volumes (225 $\mu$ L) of Buffer PB was added to 1 volume (45 $\mu$ L) of the PCR product and mixed by pipetting up and down. A QIAquick column was placed in a provided 2 ml collection tube and labelled accordingly. The sample from the first step was added to the QIAquick column and centrifuged for 60 seconds, in order to bind DNA to the column. The flow-through was discarded and the QIAquick column placed back in the same tube. The wash step consisted of adding 750  $\mu$ l Buffer PE to the QIAquick column, followed by centrifugation for 60 s. The flow-through was discarded and the QIAquick column placed back in the same tube again. The QIAquick column was centrifuged once more in the provided 2 ml collection tube for 1 min, in order to remove residual wash buffer. Each QIAquick column was then placed in a clean, labelled 1.5 ml microcentrifuge tube. To elute DNA, 20 $\mu$ l of Buffer EB was added to the center of the QIAquick membrane and the column was centrifuged for 1 min. A smaller volume of Buffer EB was added, compared to the recommended volume, in order to increase the eluted DNA's concentration. The eluted DNA was placed on ice, vortexed briefly and its concentration and purity quantified by Nanodrop analysis.

### 3.9.3.3 Qubit quantification

The Qubit dsDNA High Sensitivity (HS) Assay Kit was used along with a Qubit Fluorometer (model and company) in order to get a more accurate DNA concentration. This assay ensured higher selectivity for double-stranded DNA over RNA and showed high accuracy for sample concentrations ranging from 10 pg/ $\mu$ l to 100 ng/ $\mu$ l. This assay tolerated traces of common contaminants, such as salts, free nucleotides, solvents, detergents or protein and did not disrupt the reading in any way. The assay was performed at room temperature and the signal had a stability of 3 hours. The kit contained the assay reagent, dilution buffer, prediluted DNA standards and 0.5 mL tubes for the assay.

The standards were prepared first and were used to calibrate the Qubit Fluorometer. The required number of 0.5 mL tubes were set up, according to the number of samples and standards. Labelling of these tubes were done on the top, as labels on the side of the tube might interfere with the sample reading and signal. Calibration required insertion of the standards into the instrument in the correct order. All the reagents, standards and dye were thawed before starting. The dye was covered with aluminium foil as it was light sensitive. The dye and buffer were left at room temperature, vortexed briefly and spun down before use.

The master mix was prepared by adding the buffer and dye, for the required number of samples, two standards and an allowance of 1 extra sample for pipetting errors. For example, for a set of 6 samples, a master mix was prepared for 9 samples, including the two standards. The volume of buffer added for 9 samples was 1791 $\mu$ l (199 $\mu$ l for each sample), along with 9  $\mu$ l of the dye. This mixture was vortexed, spun down and aliquoted into the 8 different labelled tubes, including the standards. The samples and standards were vortexed, spun down and added in their respective tubes. Sample tubes contained 198 $\mu$ l of the master mix, along with 2  $\mu$ l of the sample to be measured. Standard tubes contained 190  $\mu$ l of the master mix, along with 10  $\mu$ l of each standard. The standards were stable for 2 hours after preparation. New standards were prepared after 2 hours, for measurement of a new batch of samples.

All the tubes were vortexed briefly, spun down and then incubated for 2 minutes at room temperature. On the home screen of the Qubit 2.0 Fluorometer (Invitrogen, USA), the option DNA was selected, followed by dsDNA High Sensitivity selection for the assay type. The 'Read standards' screen was then displayed and selected. The S1 tube was first inserted into the instrument, the lid closed, and then read for around 3 seconds. The same procedure was performed for S2 and the results were shown on the screen. Samples were read next, inserted one by one and the concentration noted down in ng/μl. Once all the samples were read, their concentration recorded were added to the excel sheet for normalization.

#### **3.9.3.4 Normalisation of samples to 0.2ng/MI**

Once all the samples were measured, the values were added to the prepared excel sheet for normalisation to 0.2 ng/μl. A minimum of 2 μl of the samples were used, in order to have enough DNA molecules in tube to proceed for the future steps. EB buffer from the QIAquick PCR Purification kit (Qiagen, Germany) was used to dilute the samples. The sample was added to the required volume of EB buffer, vortexed briefly and spun down. Once those samples were diluted, NGS sample preparation ensued right after, due to the instability of the DNA samples in such a small concentration. The final concentration of 0.2 ng/μl was checked on the Qubit 2.0 Fluorometer for all the samples.

### 3.9.4 NGS Library Preparation

The Nextera XT DNA Library Prep (Illumina, USA) manual was used along with the help of a NGS professional from the ScienceVision Company in Malaysia. Several optimisation steps and precautions were included throughout the manual in order to obtain the best results and minimize chances of errors. Some methods provided in the Nextera XT DNA Library Prep (Illumina, USA) were changed, in order to accommodate the availability of kits and to minimize errors and time consumed.

The library preparation can be simplified into a workflow (Figure 3.12 on the next page) for a set of 6 to 10 samples, to increase accuracy and reduce all possible errors in handling and lab work.

Careful planning was required for each step as the library preparation was a highly intensive lab session where the probability of errors was elevated. By setting up a daily preparation schedule (shown in Figure 3.13 on the next page) and dissecting each step in the manual beforehand with the help of a NGS specialist, all the small possible errors and precautionary measures were highlighted properly. NGS was a highly sensitive method and any small errors may deviate the results in undesired ways. Every single detail, such as the number of times a mixture was pipetted up and down, or the method of mixing by flicking only were kept constant for all the 6 samples. A simple deviation from the original method can result in different outcomes or outputs. Once the 6 samples were prepared, they were normalized, pooled together and finally run on the MiSeq machine (shown in Figure 3.14)

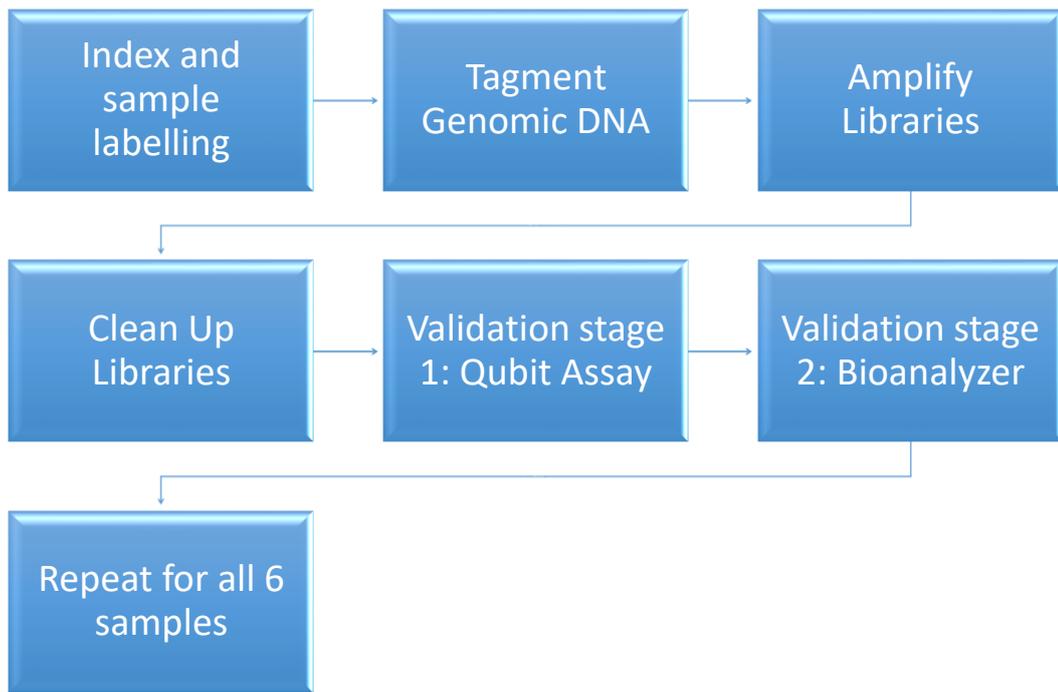


Figure 3-12 NGS library preparation workflow, which was carried out and repeated for every single sample.

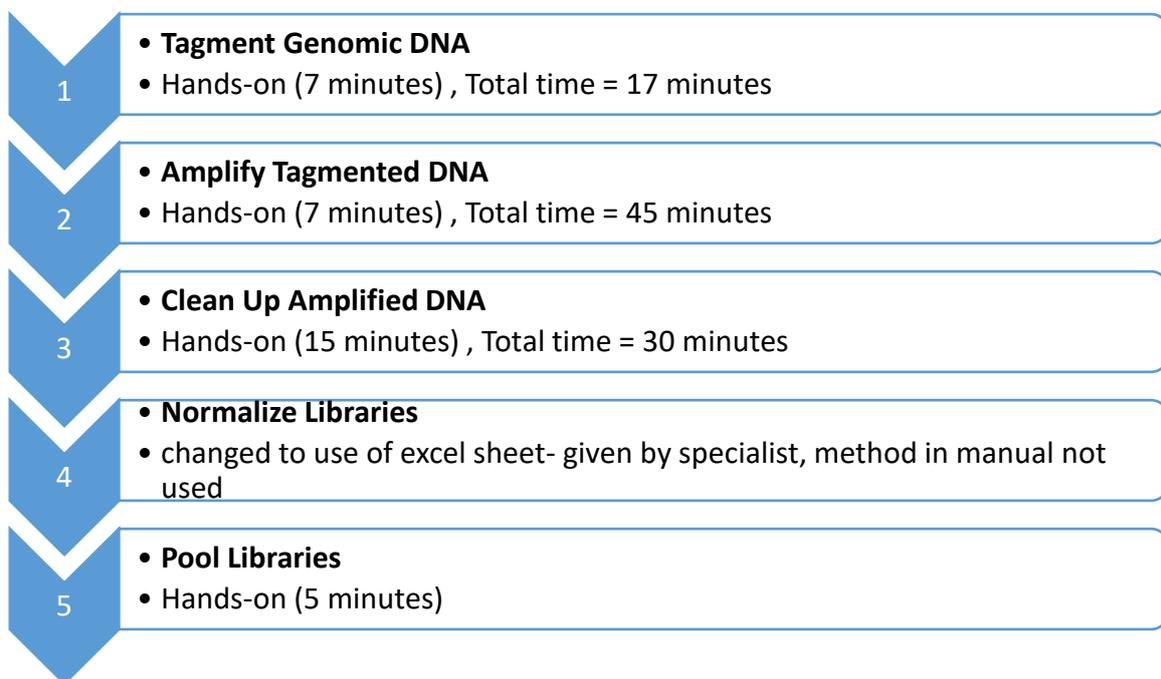


Figure 3-13 Breakdown of each step with the required time for hands-on preparation and the actual running time.

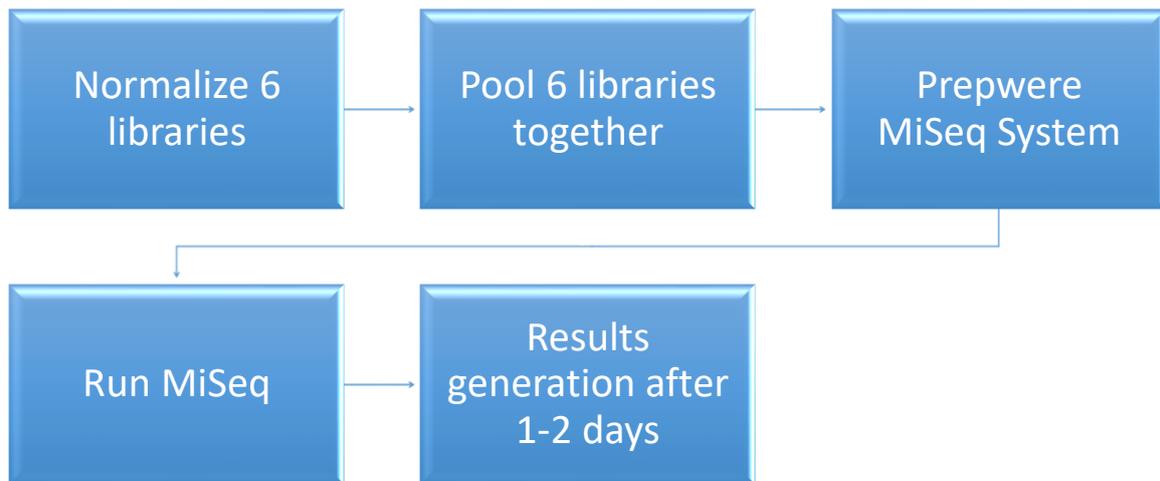


Figure 3-14- Last step for the NGS preparation and the final running of all the libraries on the MiSeq system and following result generation.

The following sub-sections will lay down all the details for each single step of the NGS library preparation until the final NGS run on the MiSeq machine.

### 3.9.4.1 Index and sample labelling

An IEM file was generated on Excel with a set of all the index primers to be used for the 96 NGS samples. They were listed in a combination of i5 and i7 index primers, where two primers were only used in combination once for one sample. This excel sheet will be inputted directly in the MiSeq system and should not be modified in any way, to avoid errors during the final NGS run. No index primers should be modified and the file should be saved in csv format for compatibility with the MiSeq system.

The sample ID were just filled in for the 6 samples as shown in Table 3.11 below, where the index primer combinations for the 6 gout samples were shown. PC5801 stands for the positive control containing the HLA-B\*58:01 and the code HLAG was used for the other 5 samples, along with their sample numbers. The full table was included in Appendix B.

Table 3-11 Extract from the IEM index labelling sheet.

Sample_ID	I7_Index_ID	Index	I5_Index_ID	index2
PC5801	N701	TAAGGCGA	S517	GCGTAAGA
HLAG12	N702	CGTACTAG	S517	GCGTAAGA
HLAG19	N703	AGGCAGAA	S517	GCGTAAGA
HLAG26	N704	TCCTGAGC	S517	GCGTAAGA
HLAG28	N705	GGACTCCT	S517	GCGTAAGA
HLAG29	N706	TAGGCATG	S517	GCGTAAGA

### 3.9.4.2 Tagment Genomic DNA

#### 3.9.4.2.1 Preparation required

The following tagmentation program was saved on the thermal cycler (Applied Biosystems, USA); preheat lid option was selected, followed by a heating step of 55°C (optimum temperature) for 5 minutes and ending with a hold step at 10°C. ATM reagent must be stored at -20°C at all times as it was quite temperature sensitive. The tube must be removed quickly, used instantly and stored back at -20°C immediately. The reagent does not freeze and was mixed by simply inverting the tubes a few times. TD reagent was also stored at -20°C and thawed by rolling the tube between two hands. All buffers used were vortexed and spun down before use.

#### 3.9.4.2.2 Procedure

10 µL of TD was added to 5µL of the normalized gDNA of a sample in a labelled PCR tube. This mixture was pipetted up and down for 10 times in order to mix it properly. All preparations were done at rtp on the bench here. 5µL of ATM was added to the tube and the latter was flicked hard 3 times, in order to mix it properly. The flicking step was critical for success of tagmentation step and bubbles must be seen in the tube. Centrifugation was carried out for 1 minute on a bench-top centrifuge, until all the bubbles disappeared. The tube was then placed in the thermal cycler and the tagmentation program was started. As soon as the temperature dropped to 10°C, the tube was removed from the thermal cycler. 5µL of NT was added to the tube and the mixture was pipetted up and down for 10 times, followed by 3 flicks. Centrifugation was carried out for 1 minute, followed by an incubation step at room temperature for 5 minutes.

### 3.9.4.3 Amplification of libraries with indexes

#### 3.9.4.3.1 Preparation required:

The amplification program shown in Table 3.12 was first saved on the thermal cycler for this particular step (Applied Biosystems, USA).

Table 3-12 Amplification conditions for libraries

Step Number	Temperature (°C)	Time (seconds)
<b>1</b>	72	180
<b>2</b>	95	30
<b>3 (12 cycles of)</b>	95	10
	55	30
	72	30
<b>4</b>	72	300
<b>5</b>	10	Hold

The required index primers were chosen according to the sample code on the IEM excel sheet and removed. Both the index primers and NPM (PCR master mix) were thawed with hands and left on ice. While waiting for the PCR to finish, the KAPA pure beads need to be taken out, thawed at room temperature for 30 minutes, vortexed thoroughly and wrapped in aluminium foil to prevent exposure to light. Index primer caps must always be changed to a new orange one after use, in order to prevent contamination by misplacing of caps.

#### 3.9.4.3.2 Procedure:

5µL of Index 1 (i7) was added to a labelled PCR tube and its cap replaced with a new orange one. 5µL of Index 2 (i5) was then added to the same tube and its cap replaced with a new orange one. 15µL of NPM master mix was added to the tube and mixed by pipetting it up and down a few times. The mixture from the tagmentation step, all 25µL of it, was added to this tube and pipetted up and down 10 times for mixing. Pipetting here needs to be done slowly, in order to avoid bubbles. Other methods of mixing such as vortexing or flicking were prohibited here. The tube was finally centrifuged, placed on the thermal cycler and the PCR started.

### 3.9.4.4 Clean Up libraries

#### 3.9.4.4.1 Preparation:

RSB was always thawed at room temperature and stored at 2°C to 8°C afterwards. KAPA pure beads were used here and guidelines for its use was procured from its Technical Data Sheet (TDS). The amount of KAPA beads to be used was calculated from Table 2 on page 5 of the TDS. The optimum beads-to-sample volumetric ratio was 0.6X for fragments  $\geq 450$ bp to be retained during clean up. Hence, for a total of 50  $\mu$ L of PCR mix, 30  $\mu$ L of KAPA beads were used. KAPA beads were aliquotted into a few 1.5 mL micro centrifuge tube, in order to prevent contamination during use. Fresh 80% ethanol was prepared from absolute ethanol for each new batch of samples. One sample required a total of 500 $\mu$ L of 80% ethanol and this was multiplied by the number of samples to be processed in each batch.

#### 3.9.4.4.2 Procedure:

The PCR product (all 50 $\mu$ L) from the previous step was centrifuged, spun down and transferred to a labelled 1.5mL micro centrifuge tube. KAPA beads were vortexed thoroughly and 30 $\mu$ L was added to the 1.5 mL micro centrifuge tube. This mixture was pipetted up and down 20 times gently, in order to avoid splashing on the walls of the tube and therefore loss of the mixture. An incubation at rtp was allowed for 5 minutes. The tube was then placed in a magnetic stand gently and after 2 minutes a clear liquid was obtained. The tube can be moved around slightly to ensure that all beads were gone from the clear liquid. The tube was kept on the magnetic stand and all the clear liquid was removed by using 100  $\mu$ L pipette tips. Two wash steps were performed with 200  $\mu$ L of freshly prepared 80% ethanol. After the first addition of 200  $\mu$ L of ethanol, it was left in the tube for 30 seconds and then removed with a 100  $\mu$ L pipette tip. The cap of the microcentrifuge tube was not closed to avoid moving the beads around. The second wash step was performed in the same way quickly. The residual ethanol was removed from the tube by using a 20  $\mu$ L pipette tip. No splashing of liquid must be allowed on the walls of the tube. In case of liquid presence on the walls of the tube, it must be spun down and magnetised again. The tube was air-dried on the magnetic stand for a maximum of 15 minutes. This step was monitored carefully by observing the beads' appearance. Shiny beads indicated incomplete drying, whereas a mat looking beads indicated dry beads. Cracked beads indicated over-drying and must be pipetted more to be mixed in the next stage. The tube was removed from the

magnetic stand and placed on a normal tube holder. 52.5µL of the RSB was added to the tube and pipetted 40 times for proper mixing. The tube was incubated for 2 minutes at rtp and then placed on a magnetic stand. After 2 minutes a clear liquid must be visible. 50µL of the supernatant was transferred to a clean, labelled 1.5 mL micro centrifuge tube. The tube was placed on ice and the required amount was removed for Qubit assay and bioanalyzer. This tube was sealed and stored at -20°C after analysis and only removed when it needs to be normalized and pooled for the final MiSeq run.

#### **3.9.4.5 Qubit Quantification**

The same steps were performed as previously mentioned

#### **3.9.4.6 Run Bioanalyzer**

The Agilent High Sensitivity DNA Kit (Agilent Technologies, USA) used contained a set of 10 DNA chips, one electrode cleaner, one new syringe and two spin filters. The 4 reagents in the kit included the DNA ladder, four vials of DNA markers (35/10380bp), one vial of DNA Dye Concentrate and two vials of DNA Gel Matrix.

##### **3.9.4.6.1 Setting up the Assay Equipment and Bioanalyzer**

The chip priming station and the bioanalyzer were set up before the chip preparation. Ensure that the syringe was replaced at the chip priming station with new kit. The base plate of the chip priming station was next adjusted, followed by the syringe clip and the bioanalyzer's chip selector. The vortex mixer was set up with the required speed and time and the software (2100 Expert Software) was started on the computer before chip loading.

#### 3.9.4.6.2 Setting up the Chip Priming Station

**Replacing the syringe:** The old syringe was unscrewed, released and discarded from the lid of the chip priming station. The plastic cap of the new syringe was removed and placed into the clip. The latter was slid into the hole of luer lock adapter and was screwed tightly to the chip priming station.

**Adjusting the base plate:** The chip priming station was opened by pulling the latch. Using a screwdriver, the screw at the bottom of the base plate was opened. The base plate was lifted, inserted again in position and the screw retightened. **Adjusting the syringe clip:** The lever of the clip was released and slid down to the lowest position.

#### 3.9.4.6.3 Setting up the Bioanalyzer and Vortex Mixer

The chip selector was first adjusted by opening the lid of the bioanalyzer and making sure that the electrode cartridge was inserted in the instrument. If not, the latch was opened, the pressure cartridge was removed and the electrode cartridge was inserted. Any remaining chip was removed and the chip selector was adjusted to position (1). The vortex mixer, IKA Model MS3 (IKA, Germany), was adjusted to a speed of 2400rpm.

#### 3.9.4.6.4 Essential Measurement Practices

All reagents were handled and stored according to the instructions on the label of the individual box. Avoid sources of dust or other contaminants. Foreign matter in reagents and samples or in the wells of the chip will interfere with assay results. All reagents and reagent mixes need to be kept refrigerated at 4 °C when not in use. All reagents and samples were allowed to equilibrate to room temperature for 30 min before use. Dye and dye mixtures were protected from light by an aluminium foil wrap and the latter was only removed when pipetting. The dye was at risk of decomposition when exposed to light and will cause a reduced signal intensity. The pipette tip was always inserted to the bottom of the well when

dispensing the liquid. Placing the pipette at the edge of the well may lead to poor results. A new syringe and electrode cleaners were used with each new kit. Loaded chips need to be used within 5 min after preparation due to possible evaporation of reagents, which might lead to poor results. The Agilent 2100 Bioanalyzer must not be touched during analysis and must never be placed on a vibrating surface.

#### 3.9.4.6.5 Preparing the Gel-Dye Mix

After completing the initial steps, the assay can be prepared, the chip loaded, and the assay started, as described in the following procedures.

##### 3.9.4.6.5.1 Precautions

**Handling DMSO:** Kit components contain DMSO. Because the dye binds to nucleic acids, it should be treated as a potential mutagen and used with appropriate care. Hand and eye protection must be worn and good laboratory practices must be followed when preparing and handling reagents and samples. Solutions must be handled with particular caution as DMSO was known to facilitate the entry of organic molecules into tissues. The prepared gel-dye mix was sufficient for 5 chips and was used within 6 weeks of preparation. The gel-dye mix was protected from light and stored at 4 °C when not in use for more than 1 hour.

##### 3.9.4.6.5.2 Procedure

The blue- capped High Sensitivity DNA dye concentrate and red- capped High Sensitivity DNA gel matrix were allowed to equilibrate to room temperature for 30 minutes. The dye concentrate were protected from light with an aluminium foil wrap. The blue- capped vial with High Sensitivity DNA dye concentrate was vortexed for 10 seconds and spun down. The DMSO was completely thawed for this step. 15 µl of the blue capped dye concentrate was pipetted into a red- capped High Sensitivity DNA gel matrix vial. The dye concentrate was stored at 4 °C in the dark again after use. The tube was capped and vortexed for 10 seconds.

Proper mixing of gel and dye was inspected visually. The complete gel- dye mix was transferred to the top receptacle of a spin filter. The spin filter was placed in a microcentrifuge and spun for 10 minutes at room temperature at 2240 g. The filter was discarded according to good laboratory practices and the tube labelled with the date of preparation.

#### 3.9.4.6.6 Loading the Gel-Dye Mix

##### 3.9.4.6.6.1 Precautions

Before loading the gel-dye mix, the base plate of the chip priming station was set in position (C) and the adjustable clip was set to the lowest position. When pipetting the gel-dye mix, make sure not to draw up particles that may sit at the bottom of the gel-dye mix vial. The tip of the pipette was inserted to the bottom of the chip well when dispensing. This prevented a large air bubble forming under the gel-dye mix. Placing the pipette at the edge of the well may lead to poor results.

##### 3.9.4.6.6.2 Procedure

Allow the gel- dye mix to equilibrate to room temperature for 30 minutes before use. Protect the gel- dye mix from light during this time. Take a new High Sensitivity DNA chip out of its sealed bag and place the chip on the chip priming station. Pipette 9.0  $\mu\text{l}$  of the gel- dye mix at the bottom of the well marked and dispense the gel- dye mix. Set the timer to 60 seconds, make sure that the plunger was positioned at 1 ml and then close the chip priming station. The lock of the latch will click when the Priming Station was closed correctly. Press the plunger of the syringe down until it was held by the clip. Wait for exactly 60 seconds and then release the plunger with the clip release mechanism. Visually inspect that the plunger moves back at least to the 0.3 ml mark. Wait for 5 s, then slowly pull back the plunger to the 1 mL position. Open the chip priming station. Pipette 9.0  $\mu\text{L}$  of the gel- dye mix in each of the wells marked.

#### 3.9.4.6.7 Loading the Marker

5  $\mu\text{L}$  of green- capped High Sensitivity DNAMarker was pipetted into the well marked with the ladder symbol and into each of the 11 sample wells. No wells were left empty as this may impede the chip's run and results.

#### 3.9.4.6.8 Loading the Ladder and the Samples

1  $\mu\text{L}$  of the yellow- capped High SensitivityDNA ladder vial was pipetted in the well marked with the ladder symbol. In each of the 11 sample wells, 1  $\mu\text{L}$  of sample (used wells) or 1  $\mu\text{L}$  of marker (unused wells) were pipetted in. The chip was placed horizontally in the adapter of the IKA vortex mixer and care was taken not to damage the bulge that fixes the chip during vortexing. The chip was vortexed for 60 seconds at 2400 rpm. The chip was then inserted in the Agilent 2100Bioanalyzer and run within 5 minutes of preparation. Once the ladder and samples were loaded, the chip was inserted in the Agilent 2100 Bioanalyzer and run. After assay completion, the used chip was immediately removed and disposed. The electrodes were then cleaned by filling the electrode cleaner with 350 $\mu\text{L}$  of deionized analysis-grade water, inserting it into the bioanalyzer and leaving it there for 10 seconds with the lid closed. The electrode cleaner was removed and the lid was left open for 10 seconds to allow the water on the electrodes to evaporate. The results were recorded in an excel sheet along with the final qubit concentration for each samples.

#### 3.9.4.7 Normalize and pool all libraries together

Once all the 6 samples were prepped, their final qubit concentration and their length were added to the Library Quantification sheet in order to normalize all of them to 2nM. All the 96 samples were diluted to 2nM by adding EB buffer (Qiagen, Germany) , followed by 0.1% Tween20(Sigma-Aldrich, USA) and 10  $\mu$ L from the sample's stock. A stock of 5mL EB buffer and 5  $\mu$ L of Tween 20 was prepared and vortexed. Different volumes were added to the 96 empty, labelled 1.5 mL microcentrifuge tubes according to the calculations in the excel sheet, followed by 10  $\mu$ L of each sample into their respective tubes. The contents of the tubes were all pipetted up and down 10 times, for thorough mixing. A standard volume of 10  $\mu$ L was used for all samples so as to have enough sample DNA molecules in the final reaction mix. All the samples were then pooled together in one 1.5mL microcentrifuge tube by adding 5  $\mu$ L out of each sample's tube (2nM) and a qubit assay was done to check the concentration.

### 3.9.4.8 MiSeq run preparation

#### 3.9.4.8.1 Denaturation and dilution of the samples' library

The pooled samples were denatured and then placed in the machine to run. For the denaturation step 5 $\mu$ L of the 2nM pooled library was added to 5  $\mu$ L of 0.2 N NaOH in a clean 1.5mL microcentrifuge tube. The latter was vortexed briefly, spun down for a few minutes and incubated for 5 minutes at room temperature. 990  $\mu$ L of pre-chilled HT1 was added to the tube containing the denatured library to make 1 mL of a 10 pM denatured library.

#### 3.9.4.8.2 Denaturation and dilution of thePhiX control

For this library, a low-concentration PhiX control (Illumina, USA) spike-in of 1% was used as sequencing control. Dilution was first performed by adding 2 $\mu$ L of 10 nMPhiX library to 3 $\mu$ L of 10 nM Tris-CL, pH 8.5 with 0.1% Tween 20. This resulted in a final concentration of 4 nM of the PhiX control. Secondly, denaturation was performed, whereby 5  $\mu$ L of 4 nM of the PhiX library was added to 5  $\mu$ L of 0.2 N of NaOH. The latter was vortexed, spun down and incubated at room temperature for 5 minutes. 990  $\mu$ L of pre-chilled HT1 was added to the tube containing the denatured PhiX library to make 1 mL of a 20 pM denatured library. This was further diluted by adding 375  $\mu$ L of the PhiX library to 225  $\mu$ L of pre-chilled HT1, in order to make 600 $\mu$ L of a 12.5 pMPhiX library. The last dilution was performed as the MiSeq Reagent Kit v3 was being used.

#### 3.9.4.8.3 Combine Library and PhiX control

594 $\mu$ L of the final library was added to 6  $\mu$ L of the final PhiX library prepwered in a clean 1.5mL microcentrifuge tube. This tube was set on ice until the MiSeq system was ready for NGS run.

#### 3.9.4.8.4 Final MiSeq run

The MiSeqsystem (Illumina, USA) was cleaned and prepped for the final NGS run with the help of an experienced lab technician at UMBI. The initial sample index sheet, with all the sample codes and corresponding indexes used, was inserted to the MiSeqsystem. The MiSeq run took around 17 hours to complete for 150 cycles. An account was created on Illumina's BaseSpace, where all the results can be uploaded to and checked later.

### 3.9.5 Bioinformatics Analysis

#### 3.9.5.1 Introduction

The major challenge in NGS reared its head after the actual NGS run. Genomic variant discovery seemed simple in theory, where you mapped reads to a reference sequence, removed mismatches and then deduced the variant. However, multiple sources of error were encountered, such as amplification bias during library prep in the wet lab, machine errors during library sequencing, software errors or mapping of artifacts during read alignment. The correct manipulation of the output data and the use of an appropriate analysis pipeline for variant calling were of the utmost importance to fully exploit the NGS output data (McKenna et al., 2010). The purpose of this bioinformatics analysis section was to confirm and validate the presence of HLA-B\*58:01 in the 6 samples sequenced. The specific HLA variants and SNPs were obtained for each sample after full bioinformatic analysis. The output data obtained from the NGS run was in form of organized and distinct subsets called read groups. The latter consisted of the intersection of libraries and lanes generated throughout the NGS run. This data was inputted into the Genome Analysis Toolkit (GATK) pipeline (McKenna et al., 2010) to get the variants and SNPs for each sample, to enable in depth analysis of the HLA-B alleles present.

GATK was labelled as the “gold standard” for managing large NGS data and for targeted sequencing as well. GATK provided rich parameters for variant calling in HLA regions, as previously proven by numerous studies. One of the advantages of this pipeline was the availability of hard filtering, where specific tailored parameters can be set by the user for specific DNA regions of interest, which in turn ameliorated the accuracy of variant calling. The Variant Quality Score Recalibration (VQSR) step generated an adaptive model based on metrics, which differentiated between strand bias and true variants (McKenna et al., 2010). The GATK pipeline (version 3.8) was used for analysis of the NGS results and the full bioinformatics analysis workflow consisted of 7 main steps as shown in the flowchart below (Figure 3.15 on the next page).

The dataset for the 6 samples first went through a pre-processing stage, where only data of good quality was filtered and adaptors were removed, before being aligned to the human genome (Hg19) by using the Burrows-Wheeler Aligner (BWA). BWA was used as it was a robust read-mapping algorithm, which can be utilized to map sequenced data to a reference genome. Output from the BWA can be paired with minimal processing by using utility software packages such as Picard and SAMtools and can be inputted directly in the GATK pipeline to complete the bioinformatics analysis. Duplicates were then marked to avoid repetition and bases were recalibrated to eliminate errors and systemic bias. Variant calling was finally performed, along with variant annotation by using 19 different databases (McKenna et al., 2010). Downstream analysis was done manually, where the SNPs for each samples were compiled in an excel sheet and their related HLA allele mined from dbSNP and NCBI databases. Novel SNPs were also identified and marked for all the 96 samples.

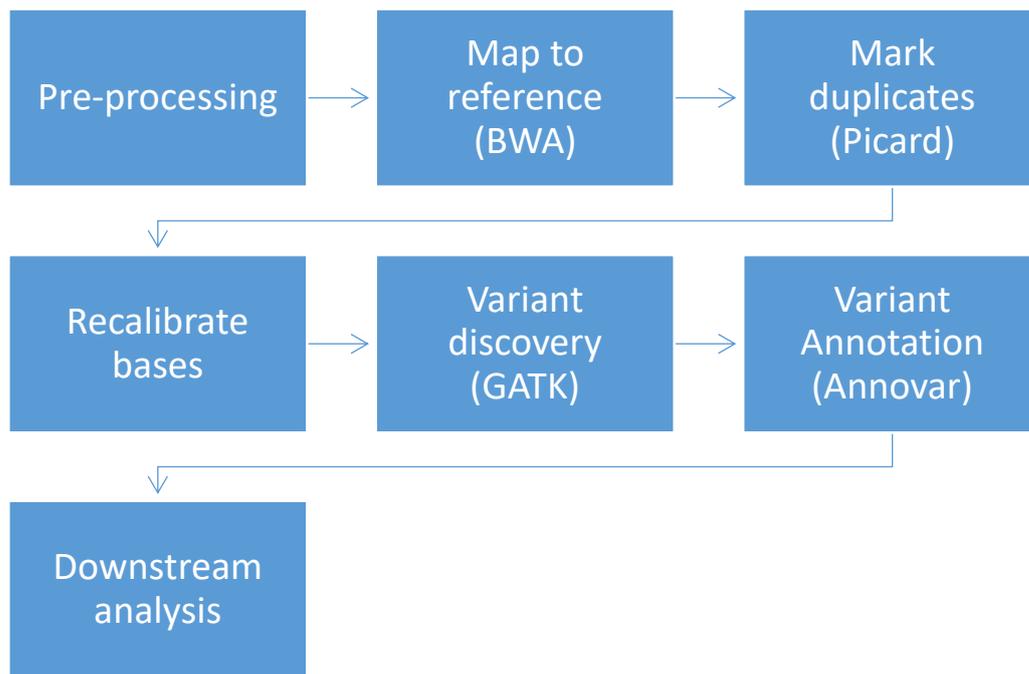


Figure 3-15 Simplified flowchart of the whole bioinformatics pathway used to analyze the NGS run results generated.

### 3.9.5.2 Stages of bioinformatics analysis

The whole bioinformatics analysis was done and optimized for one sample first and then automated to run for all the remaining 95 samples at one go. The Figure 3.16 below shows the breakdown of the initial workflow with more details concerning the input data, its format and the output generated.

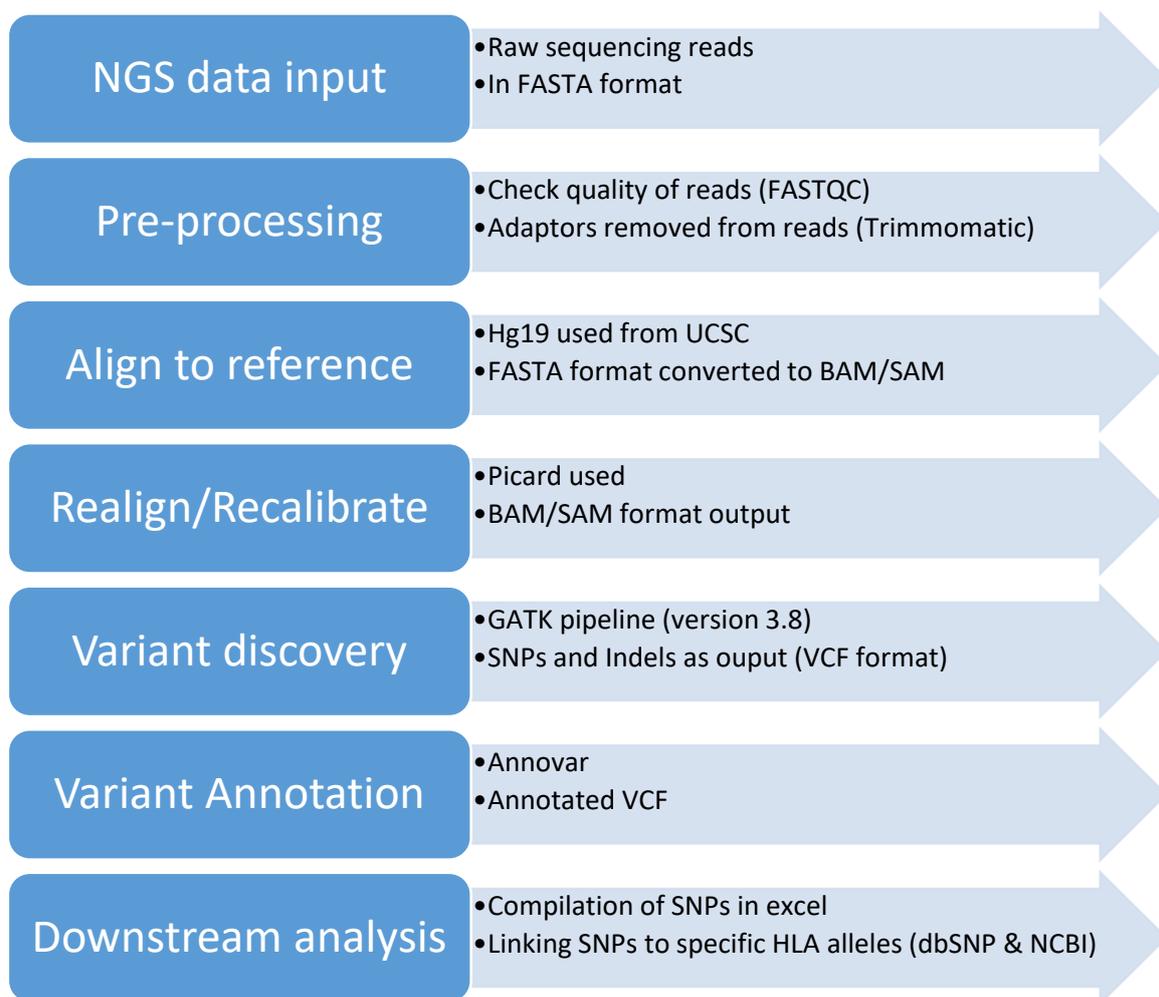


Figure 3-16 Breakdown of the full bioinformatics analysis pathway with the input format, software used and output generated.

#### 3.9.5.2.1.1 Pre-processing

The bioinformatics analysis starts with pre-processing, where the raw sequence data obtained in FASTA format was processed before obtaining data suitable for analysis. The reads were trimmed and filtered in this step by using the Trimmomatic software, in order to remove errors. This software was acknowledged as a flexible read trimming tool for Illumina NGS data and results in high quality filtered reads only.

The quality of the data was first checked by the standard tool on the Illumina platform, known as FASTQC. The report generated by FASTQC needs to be interpreted correctly, so as to check presence of good quality data and then proceed with subsequent steps. The NGS data was first opened for one sample in FASTQC and Phred scores were analysed to check the base call quality. Higher scores represented higher reliability of base calls. The background of the graph showed the division of very good quality calls in green, followed by reasonable calls in yellow and calls of poor quality in red. The box plots should be in the green or yellow parts of the graph to avoid excessive trimming and ultimately a loss in the sequence length. Overrepresented sequences were analyzed in the last tab in order to check for the presence of contamination. This section showed a list of all sequences which make up more than 0.1% of the total number of sequences. This list will initially show a few contaminants which were identified as Illumina adapters. A list of FASTQC files and results with phred scores were compiled in an excel sheet before removing the adapters.

Standard trimming was performed to remove the adapters and this resulted in a 50% loss of the reads and phred scores of 30. The stringent parameters were adjusted to a score of 20, which resulted in only 25% loss of reads. The following commands were used for trimming of read sequence for the sample S10 as shown in the commands executed below. Only S10 will be used in all the succeeding commands and steps as an example, and the same will be repeated for all the 96 samples.

### Commands used for pre-processing:

```
trimmomatic PE -phred33 -threads 20 10_S10_L001_R1_001.fastq
10_S10_L001_R2_001.fastq 10_S10_L001_R1_001_paired.fastq
10_S10_L001_R1_001_unpaired.fastq 10_S10_L001_R2_001_paired.fastq
10_S10_L001_R2_001_unpaired.fastq
ILLUMINACLIP:/home/nottingham/01_Preprocessing/NexteraPE-PE.fa:2:30:10
LEADING:3 TRAILING:3 SLIDINGWINDOW:4:20 MINLEN:36
```

#### 3.9.5.2.1.2 Mapping to a reference human genome

After pre-processing, the individual read pairs were mapped to a human reference genome, a synthetic single-stranded representation of a commonly used human genome, in order to provide a common coordinate framework for all genomic analysis. The second latest database UCSC Hg19 was used instead of NCBI. The Burrows-Wheeler Aligner (BWA) was first used to align the samples to the human reference genome Hg19. The data was converted to BAM files, sorted and then indexed by using SAMtools. A set of SAM, BAM, SAI, BAI and BWA index files will be generated in total during the mapping step. The final results will come out as mapping statistics, in percentage, for all the samples when mapped to the reference Hg19.

Commands used for mapping:

```
# Create BWA index files  
bwa index -p hg19bwaidx -a bwtsv hg19.fa  
  
# Mapping using BWA  
bwa aln -t 20 hg19bwaidx sequence1.fq.gz > sequence1.sai  
bwa aln -t 20 hg19bwaidx sequence2.fq.gz > sequence2.sai  
bwa sampe hg19bwaidx sequence1.sai sequence2.sai sequence1.fq.gz sequence2.fq.gz  
> sequence12_pe.sam  
  
# Convert SAM to BAM  
samtools view -bS sequence12pe.sam > sequence12pe.bam  
  
# Sort BAM files  
samtools sort -O bam -o sequence12pe.sorted.bam -T temp sequence12pe.bam  
  
# Index BAM files  
samtools index sequence12pe.sorted.bam
```

#### 3.9.5.2.1.3 Mark Duplicates

Marking of duplicates consisted of identification of read pairs that originated from duplicates of the same original DNA fragments by artefactual processes. This step will tag all but the single read pair within each set of duplicates, causing the marked pairs to be ignored by default during the variant discovery step. Picard was used to first sort the data (BAM files) and then mark all duplicates.

Commands for marking duplicates:

```
# Add RG Groups  
picard AddOrReplaceReadGroups I=S10.sorted.bam O=S10.sorted.addRG.bam RGID=4  
RGLB=lib1 RGPL=illumina RGPU=unit1 RGSM=20  
  
# MarkDuplicates  
picard MarkDuplicates I=./samplefile/Sorted.bam O=./samplefile/DupMarked.bam  
M=./samplefile/Duplicates.metrics CREATE_INDEX=true  
VALIDATION_STRINGENCY=SILENT
```

#### 3.9.5.2.1.4 Recalibrate Bases

In this step, machine learning was applied in order to detect and correct patterns of systemic errors in the base quality scores, which were the confidence scores generated by the sequencer for each base. Base quality scores were crucial to determine possible variant alleles correctly and to eliminate any systematic bias in the data. The possible sources of biases were biochemical processes during library preparation and sequencing, manufacturing defects in the chips used or instrumentation defects. Recalibration consisted of the collection of covariate measurements from all base calls, building a model from the latter's statistics and applying base quality adjustments to the dataset based on the resulting model. Once the recalibration rules were finalized, they were applied to the original dataset in order to generate a recalibrated dataset. The indels were realigned, followed by base recalibration as shown below in the commands executed.

#### Commands for recalibrating bases:

```
#SplitNCigarReads  
gatk -T SplitNCigarReads -R ucsc.hg19.fasta -I S10.duplicate.bam -o S10.split.bam -rf  
ReassignOneMappingQuality -RMQF 255 -RMQT 60 -U ALLOW_N_CIGAR_READS  
  
#Create targets for indel realignment  
gatk -T RealignerTargetCreator -R ucsc.hg19.fasta -I S10.split.bam -o  
S10.split.bam.intervals -nt 30 -known  
Mills_and_1000G_gold_standard.indels.hg19.sites.vcf -known  
1000G_phase1.indels.hg19.sites.vcf  
  
#Perform indel realignment  
gatk -T IndelRealigner -R ucsc.hg19.fasta -I S10.split.bam -targetIntervals  
S10.split.bam.intervals -known Mills_and_1000G_gold_standard.indels.hg19.sites.vcf -  
known 1000G_phase1.indels.hg19.sites.vcf -o S10.processed.bam  
  
#Perform BQSR  
gatk -T BaseRecalibrator -I S10.processed.bam -R ucsc.hg19.fasta -knownSites  
1000G_phase1.indels.hg19.sites.vcf -knownSites  
Mills_and_1000G_gold_standard.indels.hg19.sites.vcf -knownSites  
dbnp_138.hg19.vcf -o S10.recalibrate.table
```

### 3.9.5.2.1.5 Variant Discovery

Variants were discovered next by using the HaplotypeCaller (HC) from the GATK pipeline and then filtered. The output will be raw and filtered VCF files for each samples.

#### Commands for variant discovery:

```
# HaplotypeCaller
gatk -T HaplotypeCaller -R ucsc.hg19.fasta -I S10.recalibrate.bam -
dontUseSoftClippedBases -stand_call_conf 20.0 -o S10.variants.vcf

# Filter Variants
gatk -T VariantFiltration -R ucsc.hg19.fasta -V S10.variants.vcf -window 35 -cluster 3 -
filterName FS -filter "FS > 30.0" -filterName QD -filter "QD < 2.0" -o
S10.filtered.variants.vcf
```

### 3.9.5.2.1.6 Variant Annotation

For the final step, Annovar was used to annotate the predicted variants by using these 19 databases and the output was annotated VCF files for each samples.

- 1) refGene
- 2) knownGene
- 3) ensGene
- 4) dbnsfp33a
- 5) avsift
- 6) dbscsnv11
- 7) intervar\_20170202
- 8) cg69
- 9) esp6500siv2\_all
- 10) exac03
- 11) gnomad\_exome
- 12) gnomad\_genome



# CHAPTER 4: RESULTS

#### 4.1 Sample collection, DNA extraction, demographic data and clinical data.

A total of 290 samples were collected for this study, where half were gout patients, while the other half were healthy volunteers without any gout history. The crucial details needed for each participant were their concentration, purity, age, race, gender, drugs they took, and the hypersensitivity reactions observed in them. All this data was shown in Appendix B. A summary of the most important demographic and clinical data were shown in Table 4.1 below.

Table 4-1 Summary of the demographic and clinical data for the 145 gout patients and 145 healthy control samples.

Descriptions	Gout patients	Healthy volunteers
<b>Sample size</b>	n = 145	n = 145
<b>Gender</b>	<b>Male: Female</b> <b>134: 11</b> <b>(92.4%: 7.6%)</b>	<b>Male: Female</b> <b>68: 77</b> <b>(46.9%: 53.1%)</b>
<b>Race</b>	Indian: Malay: Chinese: Others 2: 115 :27 : 1 (1.3% : 79.3% : 18.6% : 0.7%)	Indian: Malay: Chinese: others 3: 107 :31 : 4 (2.07% : 73.8% : 21.4% : 2.76%)
<b>Age mean</b>	55 years	51 years
<b>Clinical Diagnosis</b>	Gout: 145	None
<b>SCARs or ADRs observed</b>	28 patients (19.3%)	None

Table 4.1 showed that all the demographic data were similar, except for the gender imbalance in the gout patients and volunteers cohort. There were 92.4% male gout patients, compared to 46.9% healthy volunteers, thus showing a significant difference. A higher percentage of female healthy volunteers were found (53.1%) compared to the very low percentage of females in the gout patient cohort (7.1%). Ethnic groups were closely matched between the gout patients and healthy volunteers, as seen for the Malays,

Chinese, Indians and other ethnic groups. The mean age calculated for both cohorts were close together and within the desired range of 5 years. 28 patients from the gout cohort (19.3% of cohort) were less than 40 years old, with 21 of them being at least 30 years old. The women in the gout cohort were all in an age range of 55 to 71 years old, except for one patient aged 37 years old. Clinical data showed that none of the healthy volunteers had gout and any type of SCARs/ADRs related to the latter. The patient cohort all had confirmed clinical gout and 28 (19.3%) of them had experienced SCARs or ADRs during their gout treatment. All DNA samples were extracted by using the same salting-out method and yielded good concentration and purity, as shown in Appendix B. Most of the samples had good concentration of extracted DNA of around 1000-3000 ng/ $\mu$ L, which was more than enough for all experiments, optimisations and storage of extra DNA for future use. Seven samples had a smaller concentration of extracted DNA, such as HLAG1 which had a concentration of 359.1 ng/ $\mu$ L, as shown in Table 4.2 below. The common factors for these samples are; their gender (male), higher mean age of 64 years old and more Malays. Nevertheless, this concentration was enough for all planned experiments and extra care was taken not to waste samples. The purity of all samples, measured by a ratio of 260/280nm, were within the optimum range of 1.7-2.0, showing good quality and pure DNA samples were obtained (Jorgez et al., 2006).

Table 4-2 Seven samples identified with a smaller DNA concentration.

HLA Code	Concentration (ng/ $\mu$ L)	Purity (260/280 nm)	Age	Race	Gender	Drugs	Hypersensitivity observed
<b>HLAG 1</b>	359.1	1.94	60	Indian	Male	ALLO	No ADRs noted
<b>HLAG 3</b>	223.5	1.93	74	Malay	Male	ALLO	No ADRs noted
<b>HLAG 7</b>	307.2	1.9	65	Malay	Male	ALLO	No ADRs noted
<b>HLAG 30</b>	354	1.95	33	Malay	Male	ALLO	No ADRs noted
<b>HLAG 31</b>	404.6	1.97	82	Malay	Male	ALLO	No ADRs noted
<b>HLAG 32</b>	227.1	2.00	64	Malay	Male	ALLO	No ADRs noted
<b>HLAG 43</b>	251.5	1.89	69	Chinese	Male	ALLO	No ADRs noted

Once sample collection and DNA extraction were successfully performed and validated according to all the required criteria set at the start of this research, the next steps could then be done. These 290 DNA samples were the starting material of this whole research and special attention was paid to their quantity and quality. It took a total of 3 years to collect the samples in batches, extract the DNA out of them and collect specific clinical data required from two different hospitals.

There was a whole range of ADRs and SCARs observed in the 28 patients who experienced side effects to gout treatment, as shown in Table 4.3. ADRs started from mild rashes and itchiness, to generalized rashes all over the body, to vasculitis/maculopapular rash and progressed to more severe ADRs. The more frequent form of ADRs was a mild allergy where rashes and/or itchiness occurred and presented itself in 46.4% of the 28 ADRs patients. Stevens Johnson Syndrome (SJS) came next at 10.7% occurrence followed by Exfoliative dermatitis, DRESS syndrome and allopurinol allergy reported. Vasculitis rash, maculopapular rash with skin eruptions, allopurinol hypersensitivity syndrome (AHS), anaphylaxis, transaminitis and eye angioedema all had the lower occurrence rate at 3.60% of all the ADRs. All of these patients experienced allergies to allopurinol in different forms and severity. Interestingly, samples HLAP3 (generalised rash) and HLAP51 (exfoliative dermatitis) experienced allergy to both allopurinol and febuxostat. Febuxostat was used as an alternative treatment after patients experienced allergy to allopurinol, but this treatment resulted in an allergic reaction too. When these ADRs and SCARs were placed in the whole sample size of 145 gout patients, the occurrence percentage decreased significantly as shown in Table 4.3. The percentage of affected patients in the total sample size was a fifth of what was observed within the ADR patients only, where mild allergies occurred in 8.97% of patients, followed by 2.07% for SJS, 1.38% for DRESS/exfoliative dermatitis/allopurinol allergy. Table 4.4 shows all the ADR/SCAR samples, along with their allergy details, age, gender and ethnicity. A mean age of around 57 years old was calculated for the 28 samples, with 7 samples being less than 50 years old and the rest being older (age range: 32 to 81 years old). The male to female ratio is 25:3, showing a high male preponderance. The ethnicity ratio for Malay: Chinese: Indian: others is 25:2:0:1, with a clear majority being Malay. Further investigation was done via Sanger sequencing and NGS in the next sections in order to find the source of these ADRs/SCARs in all the patients affected.

Table 4-3 Documentation of ADRs/SCARs observed in the 28 gout patients along with the percentage of patients affected.

<b>ADRs/SCARs observed in patients</b>	<b>Samples affected</b>	<b>Number of samples affected</b>	<b>Percentage affected in ADR patients only (%)</b>	<b>Percentage affected in total sample size (n=145) (%)</b>
<b>Vasculitis rash</b>	HLAG12	1	3.60	0.69
<b>Mild allergy- itchiness or rash</b>	HLAG13, HLAG18, HLAG19, HLAG28, HLAG53, HLAG64, HLAG66, HLAP2, HLAP3 HLAP25, HLAP37, HLAP46, HLAP60	13	46.4	8.97
<b>Maculopapular rash and skin eruptions</b>	HLAP56	1	3.60	0.69
<b>Exfoliative dermatitis</b>	HLAG26, HLAP51	2	7.10	1.38
<b>DRESS Syndrome</b>	HLAG29, HLAG56	2	7.10	1.38
<b>Stevens Johnson Syndrome (SJS)</b>	HLAG38, HLAG39, HLAP6	3	10.7	2.07
<b>Allopurinol hypersensitivity syndrome (AHS)</b>	HLAG46	1	3.60	0.69
<b>Anaphylaxis and allergy</b>	HLAG68	1	3.60	0.69
<b>Allergy and Transaminitis</b>	HLAP52	1	3.60	0.69
<b>Severe allergy, eye angioedema</b>	HLAP64	1	3.60	0.69
<b>Only allopurinol allergy reported</b>	HLAG48, HLAG69	2	7.10	1.38

Table 4-4 Twenty eight ADR/SCAR samples linked to their detailed allergic reactions, gender, ethnicity and age.

Sample Code	ADR observed	Age (years)	Gender	Ethnicity
HLAG12	Vasculitis rash	81	Male	Malay
HLAG13	Mild allergy- itchiness	63	Male	Malay
HLAG18	Mild allergy- rash	45	Male	Malay
HLAG19	Mild allergy- itchiness	32	Male	Malay
HLAG26	Exfoliative dermatitis	32	Male	Malay
HLAG28	Mild allergy- itchiness	64	Female	Malay
HLAG29	DRESS syndrome	68	Male	Chinese
HLAG38	SJS	72	Male	Malay
HLAG39	SJS	67	Female	Other
HLAG46	Allopurinol hypersensitivity syndrome- red eye + severe itchiness	53	Male	Malay
HLAG48	Allergy to allo	52	Male	Malay
HLAG53	Rashes	73	Male	Malay
HLAG56	DRESS + rash	64	Male	Malay
HLAG64	Rashes	66	Male	Malay
HLAG66	Itchiness	80	Male	Malay
HLAG68	Allergy to allopurinol	68	Male	Malay
HLAG69	Allergy to allopurinol	49	Male	Malay
HLAP2	Generalised rash	60	Male	Malay
HLAP3	Allergy to allopurinol- generalised rash + face swelling. Allergy to Febuxostat – rash	63	Male	Malay
HLAP6	SJS ( mouth and genitalia)	39	Male	Malay
HLAP25	Generalized skin rash	42	Male	Malay
HLAP37	Generalized rash	58	Male	Malay
HLAP46	Rash on face	37	Female	Malay
HLAP51	Allopurinol- exfoliative dermatitis + febuxostat allergy	66	Male	Malay

<b>HLAP52</b>	Allergy to allopurinol	40	Male	Malay
<b>HLAP56</b>	Generalized rash+ skin eruptions (raised and flat eruptions), maculopapular rash	41	Male	Malay
<b>HLAP60</b>	Generalised rashes	62	Male	Chinese
<b>HLAP64</b>	Severe allergy to allopurinol-angioedema + eye swelling	55	Male	Malay

## 4.2 Polymerase Chain Reaction (PCR)

### 4.2.1 Single PCR with P2 and P3

Sensitivity, specificity and accuracy of the HRM method started with a proper optimized PCR for the specific primers designed. Primers targeting exon 2, named P2, and primers targeting exon 3 (P3) were optimized by standard PCR with target sizes of 149bp and 249bp respectively to show bright bands (shown in Figure 4.1). Optimisation of the standard PCR reactions was done systematically by changing one condition at a time, assessing the improvement in results and moving on to a second factor. The first condition to be optimised was the annealing temperature by doing a gradient PCR to find the optimum annealing temperature. Other conditions like primer and DNA concentration were then modified one by one to get the best results with bright bands as shown in Figure 4.1.

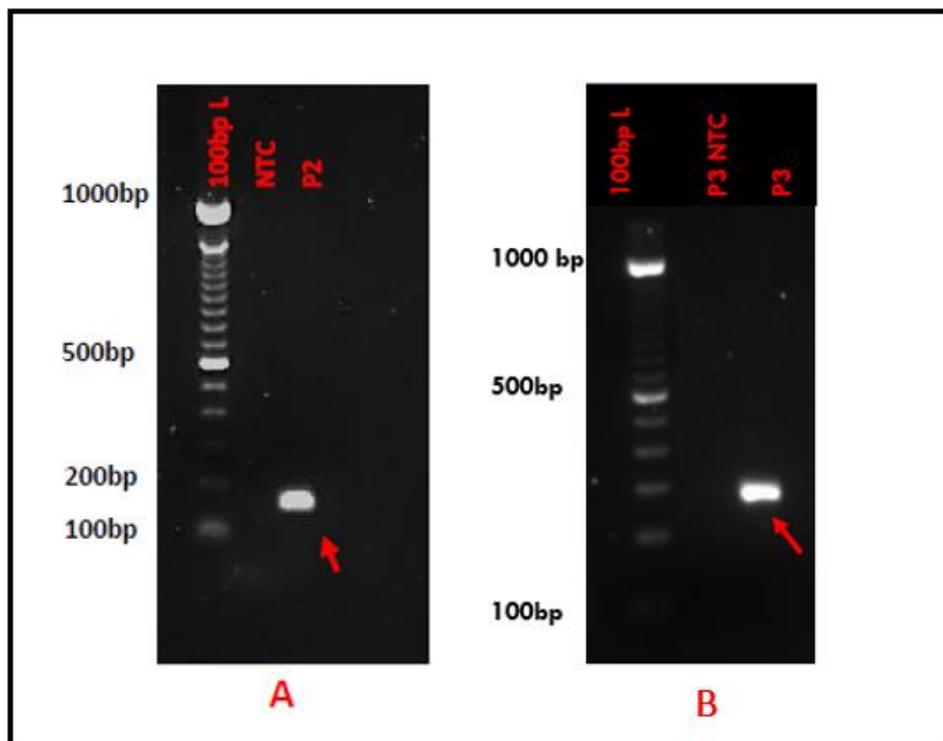


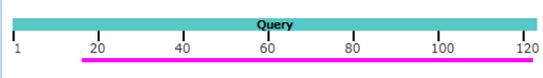
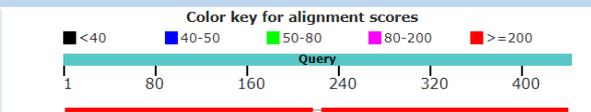
Figure 4-1 Amplification of exon 2 (Figure A) and exon 3 (Figure B) of the HLA-B\*58:01 allele. 100bp ladders are shown starting from the left, followed by the Non-template control (NTC), P2 product at 149bp (Figure A) and P3 product at 249bp (Figure B).

Figure 4.1-A shows the first lane with the 100bp ladder, followed by the Non-template control (NTC) and the P2 product. The NTC shows no amplification and contamination as shown by the absence of bands in its lane. Moreover, the NTC lane has a faint smear below 100bp which indicates the presence of primer dimers. Lane P2 shows a bright, thick band at around 150bp, corresponding to the exon 2 target, amplified at an optimum annealing temperature of 45°C. Similar results for exon 3 amplification was shown in Figure 4.1-B, with the 100bp ladder, followed by the NTC lane and the P3 product amplified at an optimum temperature of 56°C.

The PCR products were sent for Sanger sequencing and analysed on NCBI BLAST. 86% and 97% of the PCR products were found to cover the original HLA-B\*58:01 exons, along with 97.2% and 100% being identical to exon 2 and 3, respectively (shown in Table 4.5). A gap was observed when the samples and exons were aligned, as shown in the graphical representation for P3 in Table 4.5 below and in. These two primers have been used for 4 years and their stability and efficiency remained the same throughout the whole project. This was portrayed in the Sanger sequencing done for numerous samples over the years. The percentage of samples identical to the original exons were further analysed over time and a table with 15 samples' results were shown in Appendix C. More than 15 samples amplified with the P2 primers were sent for Sanger sequencing as the project progressed, but only this small batch was put into a table in Appendix C to show the constant and proper amplification of the target. The calculated average was 85.1% for the percentage coverage on the original exon and 95.15% of the PCR products were similar to the original targeted segment of exon 2. However, as the research progressed, primer P3 showed better sensitivity as shown by the higher percentage similarity (>90%) and minimum fluctuation in Appendix C compared to exon 2 results. Primer P3 had an average coverage percentage of 93% and a higher percentage similarity to exon 3 of 98%. Results shown in Table 4.6 and in the two tables in Appendix C were analysed again in August 2019, in order to keep up with the growing number of HLA alleles entry in the NCBI website. Results in Appendix C show that no matter which samples were used, the targeted segments of exons 2 and 3 were always amplified correctly. More emphasis was placed on the results for exon 3 with primers P3 during the HRM screening later on, as it had a higher stability and efficiency. This

was portrayed in Appendix G, where all the previously identified 28 ADR samples were sent for Sanger sequencing in order to see which HLA-B alleles were involved.

Table 4-5 Sanger sequencing results done on sample HLAG26 for P2 and P3 primers.

Primer/Exon	Graphical representation aligning the sample to exons (Query- sample, last pink/red line- exons)	Percentage covered on original exon (%)	Percentage identical to exons (%)
P2 (exon 2)		86.0	97.2
P3 (exon 3)		97.0	100.0

#### 4.2.2 Multiplex PCR with IC

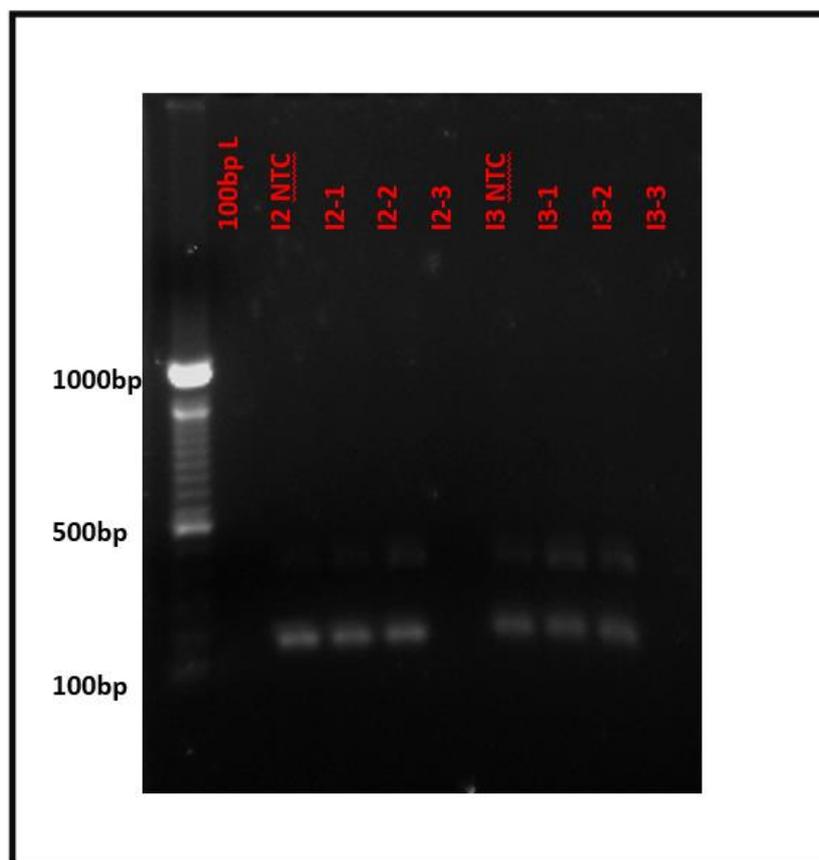


Figure 4-2 Multiplex PCR done in triplicate for I2 and I3, where I2 combined IC+P2 primers and I3 combined IC+P3 primers. IC primers produced a product at around 423bp, P2 around 149bp and P3 around 249bp.

The internal control (IC) primers along with the P2 primers were combined and labelled as I2, and the same was done for IC and exon 3 reactions, labelled as I3. This multiplex step was done as a verification step to check optimum PCR amplification, as  $\beta$ -globin primers acts as a housekeeping gene in molecular studies.

of the human genome. The gel in **Figure 4.2** showed the desired results, with the 100bpL, a clear NTC, followed by triplicates of I2 reactions, I3 NTC and triplicates of I3 reactions. The product for IC was seen faintly around 423bp for all 6 reactions and a clear difference was seen between P2 and P3 products, with the P3 products having a higher molecular weight compared to P2. These 6 reactions were optimised (shown in Appendix D) and stabilized in the presence of 0.5 $\mu$ L of 100% Dimethyl sulfoxide (DMSO), to show that the correct HLA-B region was being targeted by both primers. Thus multiplex reactions (I2 & I3) were optimised at annealing temperatures of 47°C and 50°C respectively, IC concentration of 0.3 $\mu$ M and P2 and P3 concentrations of 0.1 $\mu$ M.

## 4.3 Cloning of the positive control

### 4.3.1 Gel Electrophoresis results for cloning

The HRM screening method used in this study needed a positive control as a reference in order to compare its HRM melt curve to unknown samples being screened. The positive control targeted with primers for exon 2 and 3 were separately cloned in order to obtain an unlimited amount of the reference. Figure 4.3 shows the double restriction digestion of cloned plasmids by BamH1 and Xho 1 enzymes, with their targets as faint bands around 149bp and 249bp for exon 2 and 3 respectively. Bands may seem too faint and further validation was needed. Figure 4.3-A has the first lane containing the 1kb ladder, followed by the uncut plasmid and the purified P2 PCR product previously obtained. The purified P2 PCR product was slightly slanted due to the erroneous running or solidification of the gel. Lane 4 and 5 shows bright and thick bands at around 4000bp (empty vector) and lower fainter bands at around 200bp (exon 2 insert), showing successful cloning. Figure 4.3-B shows the 1kb ladder in lane 1, followed by the four digested plasmids (~4000bp) containing the exon 3 insert (249bp). Exon 3 inserts can be seen slightly higher than the 200bp ladder mark, as compared to exon 2 insert in Figure 4.3-A, which was around the same height or lower than the 200bp ladder mark.

Since the positive control cloned in these plasmids were used as a reference in the HRM method, Sanger sequencing was performed for every single plasmid batch extracted (shown in the next section). The results stayed the same over four years, where the two exons were always present, even after storage in -80°C, countless sub-cultures and extraction cycles. For every batch of plasmids, products were sent for Sanger sequencing before using them for the HRM screening.

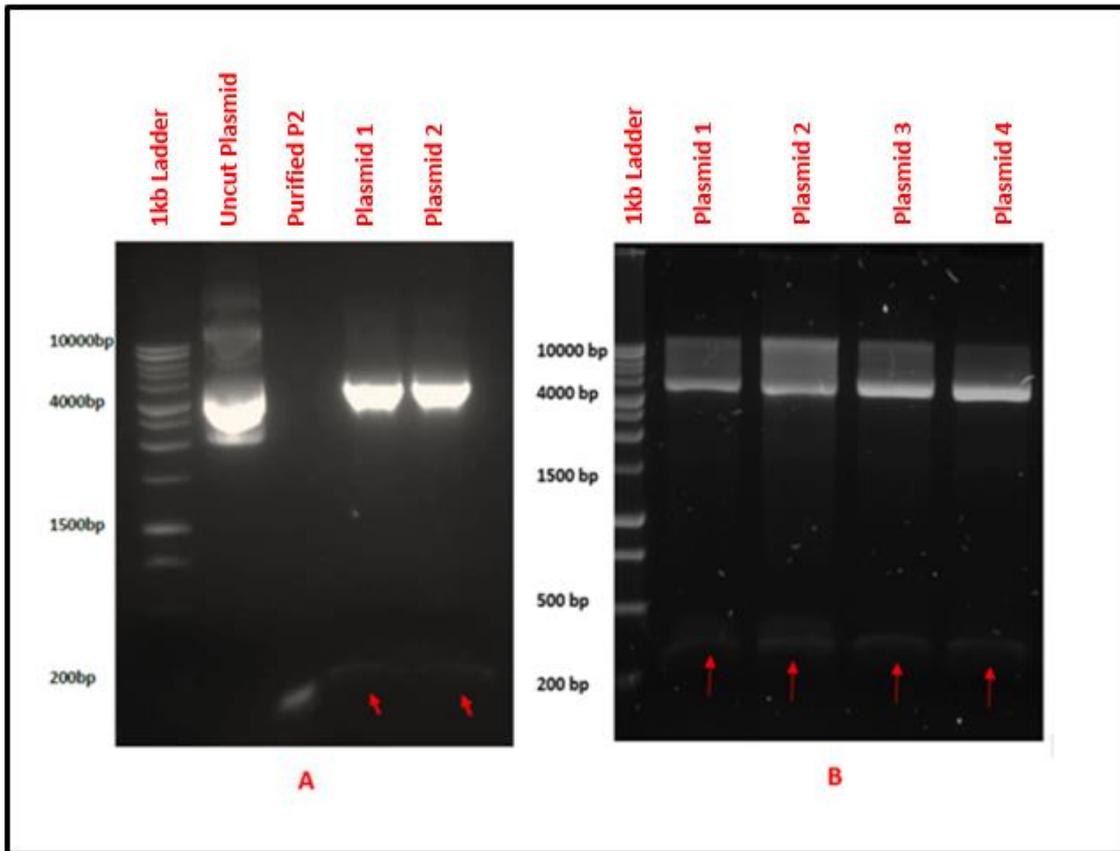


Figure 4-3 Cloning of the exon 2 and 3 of the HLA-B\*58:01 allele into the pDrive vector from Qiagen and subsequent analysis by double restriction digestion by BamH1 and Xho1. Figure 4.3-A shows the presence of the digested target exon 2, which is less than 200bp, for plasmids 1 and 2. Figure 4.3-B shows 4 digested plasmids (~4000bp) with their target exon 3 around 250bp.

### 4.3.2 Sanger sequencing for cloning of exons

The extracted plasmids (from Figure 4.3-A) containing the exon 2 inserts were sent for Sanger sequencing to further validate the cloning results. Comparing the exon 2 sequence of the HLA-B\*58:01 allele to the P2 insert showed a high percentage similarity of 96.6% and 95.5%, for plasmid 1 and 2 respectively, as shown in Table 4.6.

Table 4-6 Sanger sequencing results for the 2 cloned plasmids containing the exon 2 insert.

Plasmid Number	P2 targeted region (bp)	Percentage similarity (%)	Cloning Successful
1	149	96.6	Yes
2	149	95.5	Yes

The same procedure was repeated for the 4 plasmids shown in Figure 4.3-B, for the exon 3 cloning. Exon 3 cloning showed a higher stability and efficiency, with higher percentage similarity between the cloned insert and the HLA-B\*58:01 allele's exon 3. These results again proved the higher efficiency of the P3 primer, compared to the P2 primer. These cloned and tested plasmids will in turn be used as a reference for the HRM method, in every single run. More Sanger sequencing results were collected over time for every single batch of plasmids extracted and analysed as shown in Appendix C & G. Moreover, the specific sequence of P2 and P3 inserts aligned from their sequenced plasmids were entered in the European Bioinformatics Institute (EMBL-EBI at ebi.ac.uk) and linked to its similar HLA-B alleles present.

Table 4-7 Sanger sequencing results for the 4 cloned plasmids containing the exon 3 insert.

Plasmid Number	P3 targeted region (bp)	Percentage similarity (%)	Cloning Successful
1	249	99.0	Yes
2	249	100	Yes
3	249	98.0	Yes
4	249	100	Yes

## 4.4 High Resolution Melt (HRM) method's results

### 4.4.1 Overview of HRM screening

HRM screening was completed for all the 145 gout patient samples and 145 normal control samples for both exon 2 and exon 3 separately, making a total of 580 HRM reactions. Primers P2 and P3 were designed to target exon 2 and 3 of the HLA-B\*58:01 allele and these primers have been validated repeatedly over the study's progression. However, the P3 primer results will be emphasized in this HRM section, due to its higher efficiency and stability. All samples were screened in duplicates and with an accompanying Non Template Control (NTC) for better reliability and accuracy. One HRM run took 1 hour 20 minutes and could contain a maximum of 15 unknown samples along with one positive control reference, targeted for one exon a time. In this section, the main representative results of the HRM screenings will be shown first, followed by different types of analysis undergone, in order to find a trend and characterize what makes a positive or negative HRM screening result. Trends will be analysed by using the melting point ( $T_m$ ), percentage of Guanine and Cytosine (GC) in samples, possible SNPs present and finally by the Sanger sequencing method.

#### 4.4.2 Standard HRM screening result

In this section, the standard HRM screening results generated will be shown for the HLAG1 sample (gout sample) as an example. The full steps taken to analyse the HRM results will be shown, along with all the generated tables, plots and melt curves. Each of these results have a specific role and importance in validating an HRM result. After a HRM run was completed, all the results were instantly generated on the Eco Study software (Illumina, USA). The first step consisted in checking if all the NTC wells for the 15 samples and positive control were empty for all results (specially  $C_q$  value) and hence did not generate any curves or plots. This first step confirmed that the whole assay was free from any type of contamination and confirmed reliability of results to a certain extent. Samples were then allocated different colours in order to facilitate results viewing, differentiation and comparison. One sample's duplicate along with the positive control's duplicate were selected and the amplification plot, melt curves and other results generated were analysed. The first result generated is the  $C_q$  value, where it stands for the cycle in which fluorescence can be detected, also known as quantitation cycle.  $C_q$  is also a measure of amplification used to assess the reproducibility of amplification, a crucial part for HRM runs and analysis. A  $C_q$  value of less than 30 indicates good quality amplification happening in the HRM run.  $C_q$  values higher than 30 show that aberrant and non-specific amplification took place in the HRM run. HRM analysis can be performed even if some products amplify later than expected, as long as all similar samples have consistent amplification curves/ $C_q$  values, NTCs are not contaminated and the product is specific.

Table 4.8 shows the  $C_q$  values recorded for sample HLAG1, along with the positive control, both amplifying exon 3 only. The results for exon 3 will be shown for this whole HRM section as it showed better performance in detecting HLA alleles, compared to exon 2 when analysed via Sanger sequencing (proof shown in Appendix D). The NTCs for both samples are clear of any amplification, showing the absence of any type of contaminants.  $C_q$  values for the gout sample HLAG1 and the positive control both are below 30, showing good quality amplification. Moreover, their small standard values of 0.05 and 0.07, respectively, show that the duplicates' results are close together. This step was repeated for all the samples and the  $C_q$  values were all below 30, allowing further analysis. The melting point

( $T_m$ ) of the HLAG1 sample can also be obtained from the results generated and by extrapolating the highest peak of its melt curve to the x-axis on its difference melting curve in Figure 4.6. HLAG1 has a  $T_m$  of 90.8°C, whereas the positive control has a mean  $T_m$  around 90.75°C, showing a difference of only 0.05°C. However, this small difference can be used in SNPs identification later for more in depth analysis.

Table 4-8 HRM results compilation generated on amplification, with Cq and  $T_m$  values.

Sample Code	Cq value	Cq mean	Std. Dev. Cq	Melting point ( $T_m$ )
HLAG1	20.28	20.25	0.05	90.8
HLAG1	20.22	20.25	0.05	90.8
HLAG NTC	-	-	-	-
PC	17.96	18.01	0.07	90.8
PC	18.06	18.01	0.07	90.75
PC NTC	-	-	-	-

Once this first stage of Cq analysis is done, the next results shown is the derivative melt plot as shown in Figure 4.4. This step consisted of selecting the correct pre-melt and post-melt regions (shown as orange, parallel double-bars in Figure 4.4), in order to isolate the active melt region and align data for proper analysis (Wu et al., 2008).

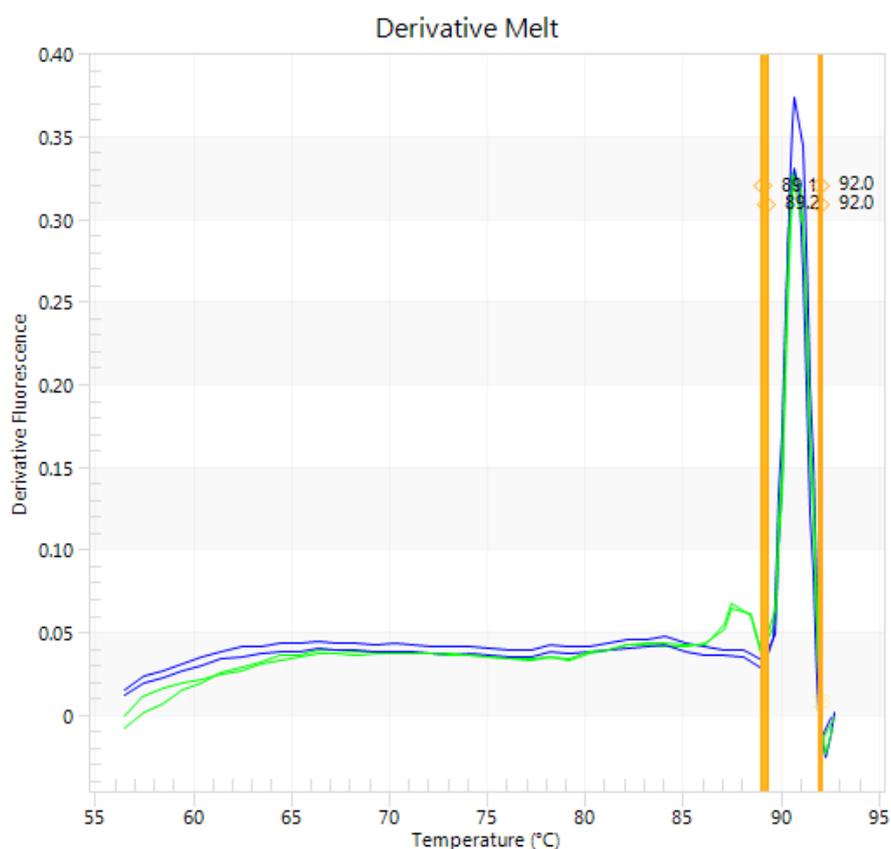


Figure 4-4 Derivative Melt plot for HLAG1 (green colour) and the positive control (blue colour).

Once the pre- and post-melt regions were defined, the normalized melting curve was generated as shown in Figure 4.5. The characteristic decrease in fluorescence with an increase in temperature was seen in the latter. This melting curve was generated after removing the fluorescence variance was eliminated and only the temperature range in between the pre- and post-melt bars was shown. This generates more distinct and specific curves as shown in Figure 4.5. A slight difference in shape in between the HALG1 sample (green curve) and the positive control (blue curve) can already be seen. In order to confirm if the two samples are different, a difference melting curve was generated (shown in Figure

4.6) by subtracting the normalized fluorescence data of the positive control from the HLAG1 sample. The positive control is used as baseline/reference for all the HRM screenings, as it contains the HLA-B\*58:01 allele. In cases where the difference melting curves of unknown samples were similar or very close to that of the positive control, a positive screening result was suspected. This melt curve obtained will be proportional to a sample's DNA sequence and is thus unique for each sample. However, the melt curve shape of the unknown sample must match that of the positive control and be close to the reference baseline curve for a positive result to be called. Figure 4.6 showed that one of the positive control's duplicate peaked downwards and deviated from the first horizontal blue line, thus showing that it's possibly a heterozygous sample. The HLAG1 sample has both green melt curves deviating from the horizontal blue line (positive control reference), with completely different curve shapes, even though they are quite close to the baseline. Hence, the HLAG1 sample was suspected to be negative. However, this must further be verified by Sanger sequencing and NGS later, as the HLAG1 sample's melt curves falls within the downward positive control's melt curve. The unexpected difference melt curve of the positive control must be deciphered in order to understand its peculiar behaviour.

All the HRM results generated by the 145 gout samples and 145 normal samples were analysed in the same way mentioned in this section. The Cq was first validated (<30), a pre-melt and post-melt region defined and the normalized and difference melt curves were finally generated, stored and interpreted. The following section will apply all possible methods of HRM analysis in order to decipher all the sample's results.

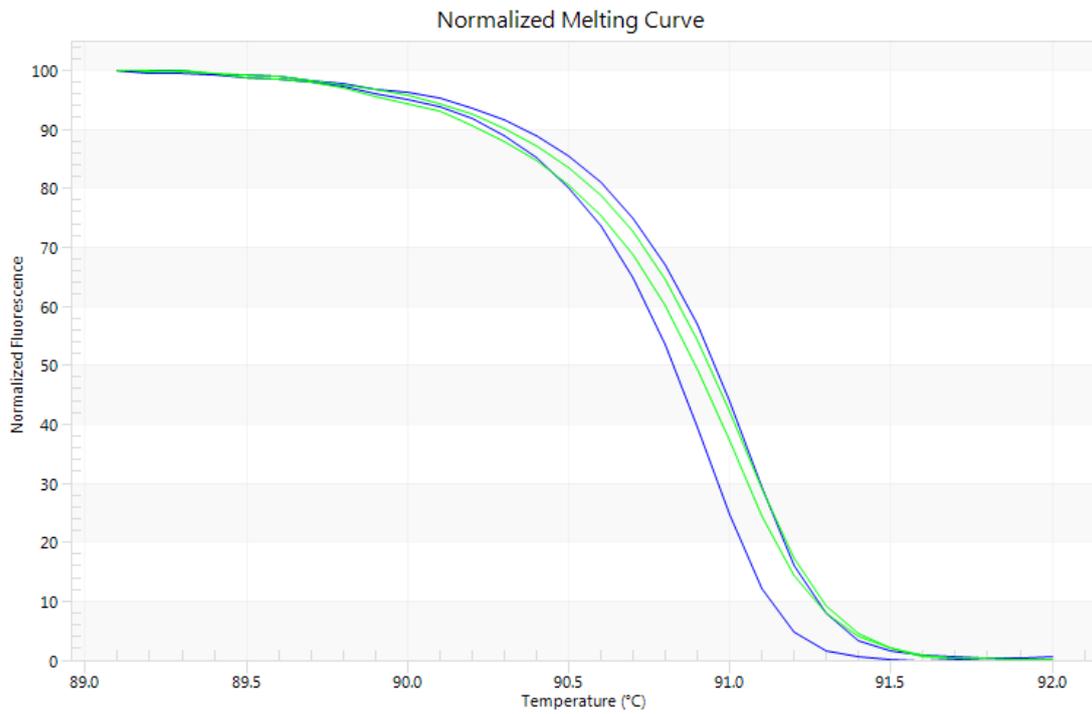


Figure 4-5 Normalized melting curve for HLAG1 (green colour) and the positive control (blue colour).

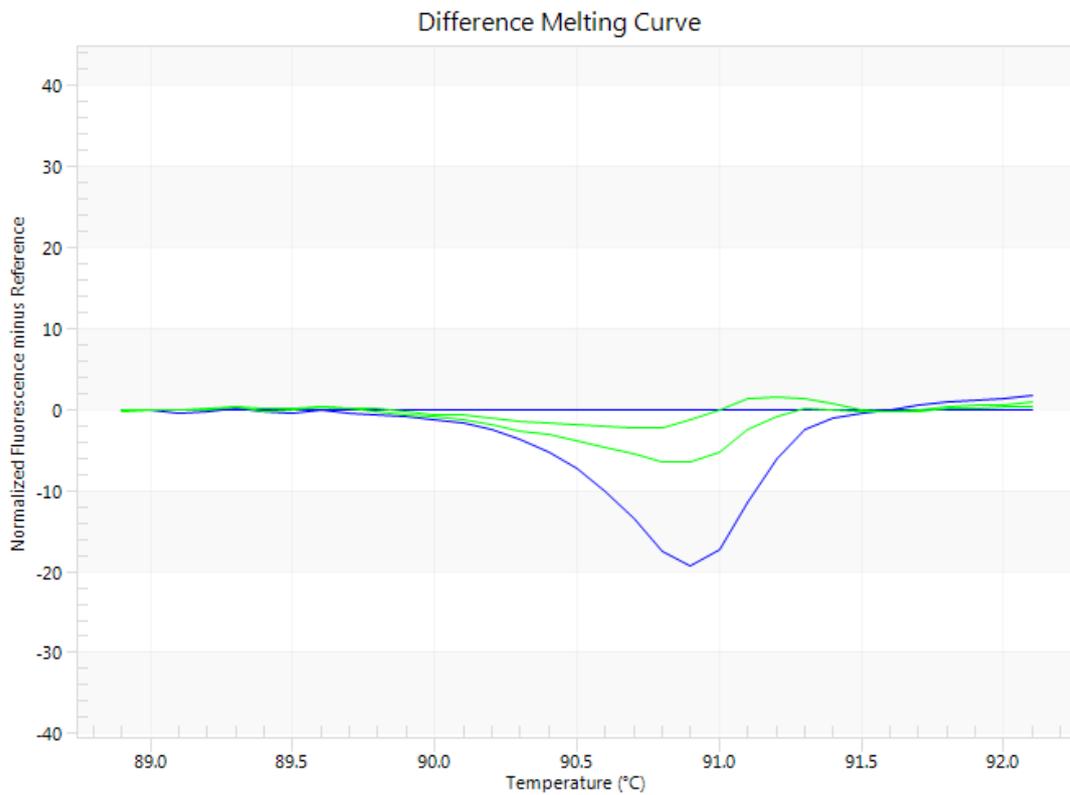


Figure 4-6 Difference melting curve for HLAG1 (green colour) and the positive control (blue colour).

### **4.4.3 Establishing a trend in the HRM results**

#### **4.4.3.1 Introduction**

Since this is a novel study to investigate the role of the HLA-B\*58:01 allele in the Malaysian population, numerous factors will be investigated thoroughly to see which one contributes to the HRM melt curves and results generated. It is crucial to investigate what constitutes a positive or a negative HRM screening result and if only the target allele is involved in the results generated or if there are other factors involved too. The results for both exon 2 and exon 3 generated were analysed at length and exon 3 results were prioritized in this thesis as sequencing results showed more specificity and accuracy in identifying various HLA-B alleles. Moreover, P3 primers showed a better specificity, stability and accuracy in amplifying the HLA-B\*58:01 allele as shown in previous sections. The most important part of the HRM results is the difference melt curve, where a definite answer will be given, whether an unknown sample contains the HLA-B\*58:01 allele or not. The melting profile depends on numerous factors, namely, the length, GC content, DNA sequence, SNPs present and heterozygosity of the amplified target. Even the shape of the melt curves provides other insights into the amplification of the target allele (Reed, 2007)

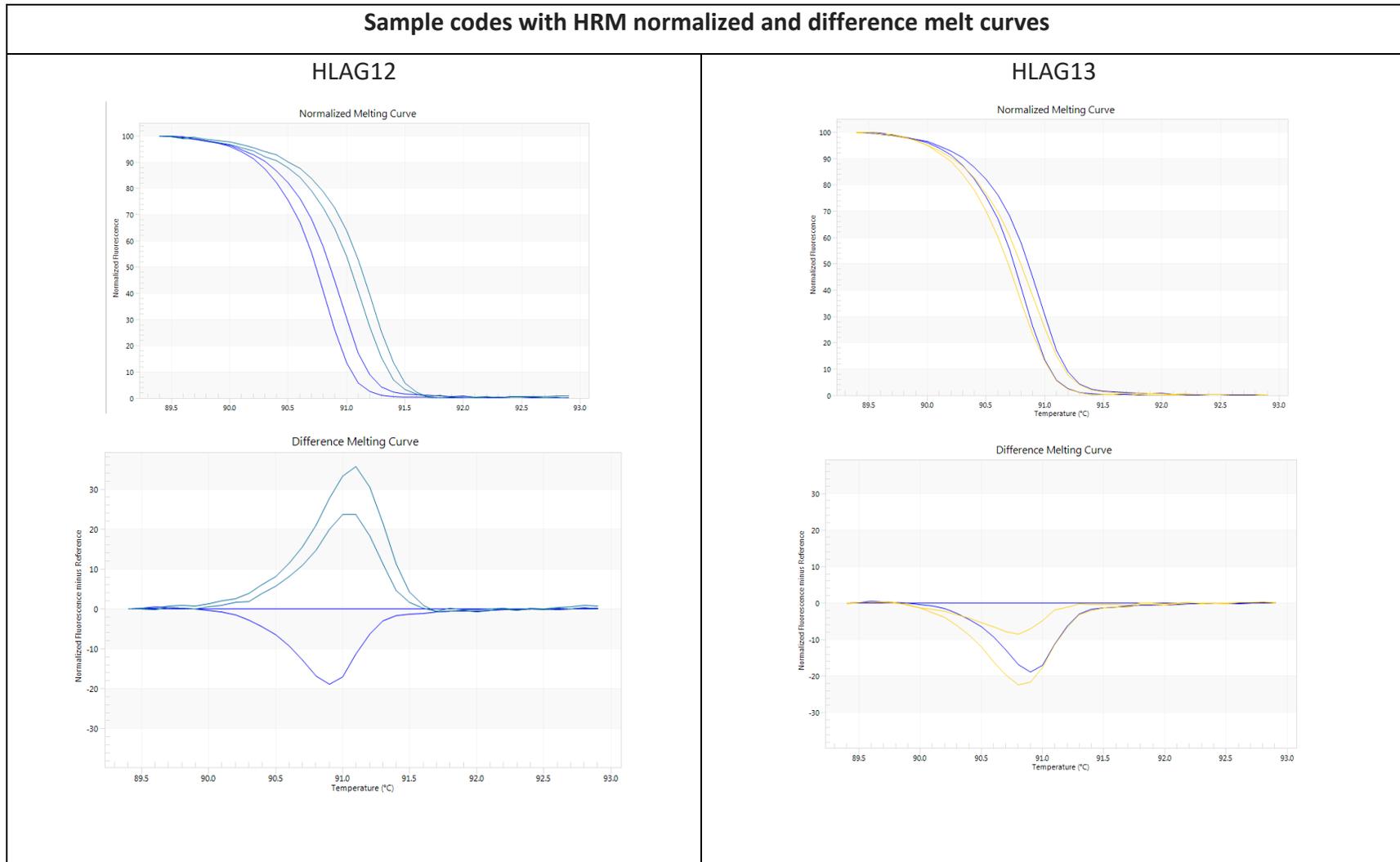
#### **4.4.3.2 HRM difference melt curve analysis**

The first result to be analysed is the trend in the shape of the difference melt curve. The HRM screening method was done on all the 145 gout samples and 145 normal samples, for both exons. A total of 580 HRM difference melt curves were generated and no specific trends were seen. The most intriguing part was that there was no single sample which matched the horizontal baseline and downwards peak of the positive control. HRM melt curves were close to the horizontal positive control baseline, but not a perfect horizontal line. This challenges the initial theory we set at the start of this research and all the previous research seen in literature.

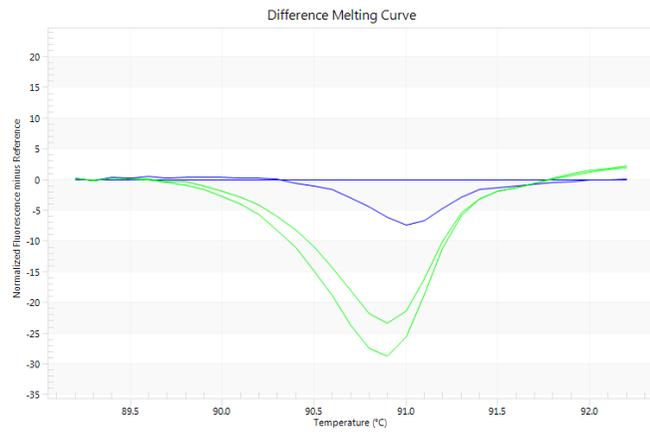
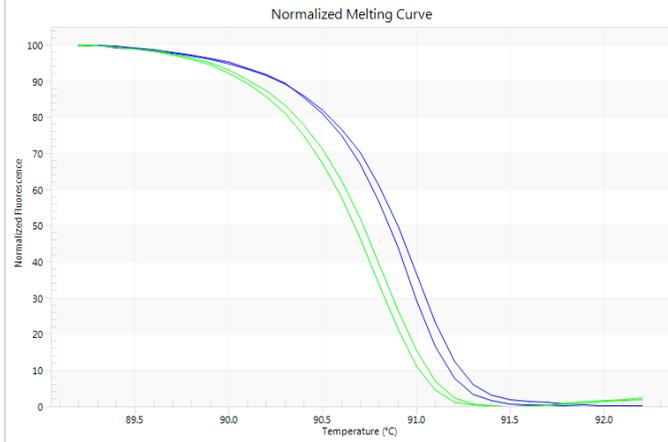
A fluorescence threshold of 20 and -20 was set based on the HRM melt curve limits seen in all the screenings done, in order to call a positive result. Samples with close curve shapes and which fall within the 20 to -20 threshold were called as positive and those beyond this were called negative. Exon 3 screening on the 145 gout patients had 23 samples called as positive and 122 called as negative. Thus 15% of the samples might contain the HLA-B\*58:01 allele in the gout cohort. However, to cement this percentage, numerous bouts of analysis need to be performed as shown in the following sections.

To further investigate this, the normalized and difference melt curves for all the 28 ADR samples will be analysed here (shown in Table 4.9). There was a wide range of ADRs observed and as per the initial theory, some of those samples should contain the HLA-B\*58:01 allele and show matching difference melt curves to the positive control. Table 4.10 shows 8 samples with melt curves close to the positive control's melt curves, namely, HLAG19, HLAG28, HLAG38, HLAG64, HLAG68, HLAG69, HLAP56 and HLAP64. Two samples showed the closest melt curves (HLAG19 and HLAG68), but the shapes and peaks still did not match the positive control perfectly. The latter had their melt curves in between the positive control's melt curves. Samples HLAG69 and HLAP64 also had close duplicate melt curves, with one curve between the positive control's melt curves and the other with an upward, opposite peak to that of the positive control. The other 20 ADR samples all show higher and steeper peaks, both upwards and downwards in the difference melt curves. By taking into consideration only the difference melt curve shapes, it was suspected that the 8 aforementioned samples had the highest chance of containing the HLA-B\*58:01 allele. In order to confirm and validate the presence of the HLA-B\*58:01 allele in any of the 28 ADR samples, they were all sent for Sanger sequencing for deeper analysis.

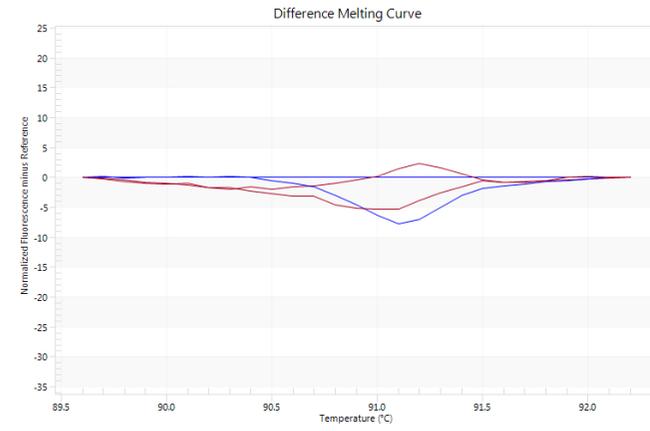
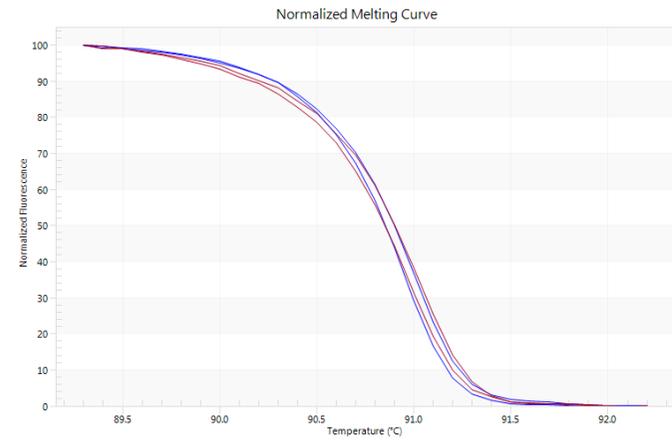
Table 4-9 HRM normalized and difference melt curve compilation for all the 28 gout samples with ADRs. The positive control's duplicate in all the graphs are shown by the dark blue horizontal line, along with the blue downward peak. The gout samples compared to the positive control are all in different colours for easier comparison.



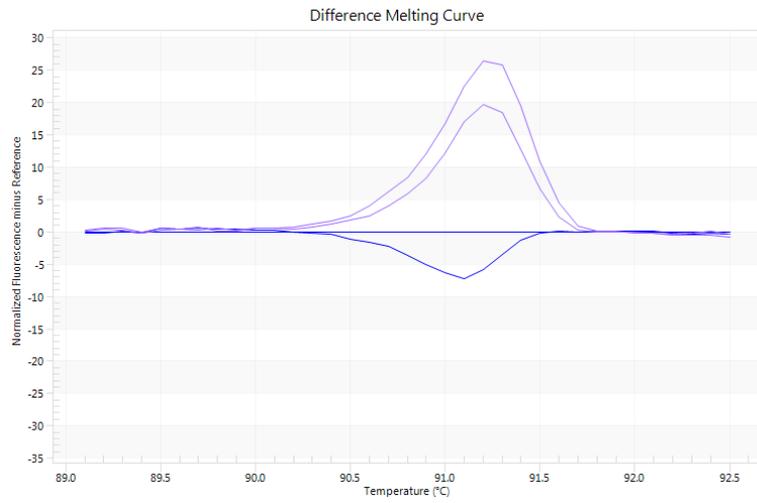
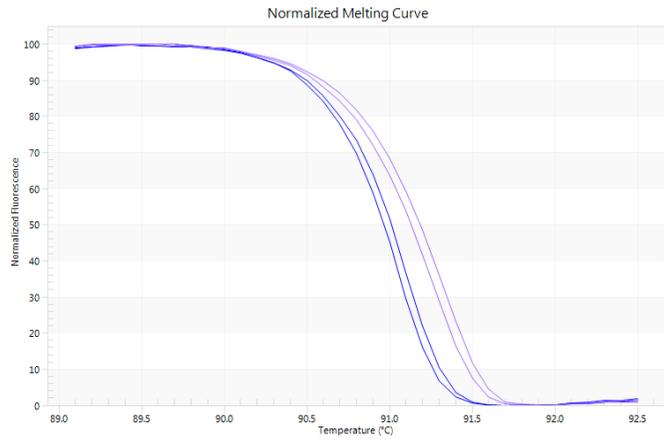
### HLAG18



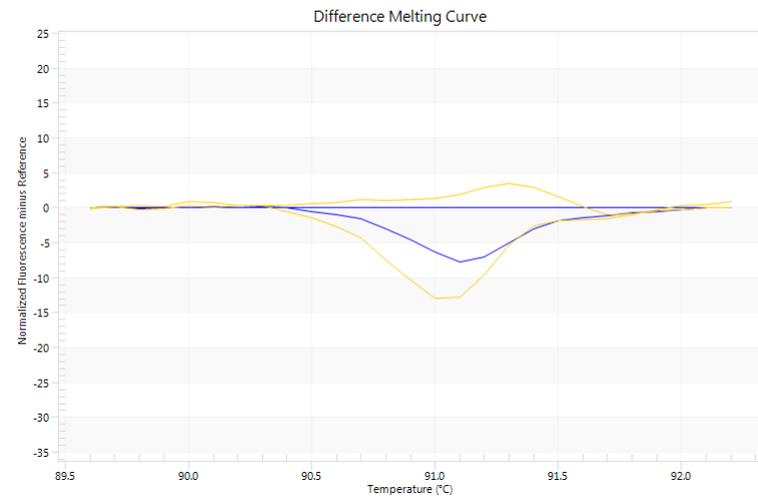
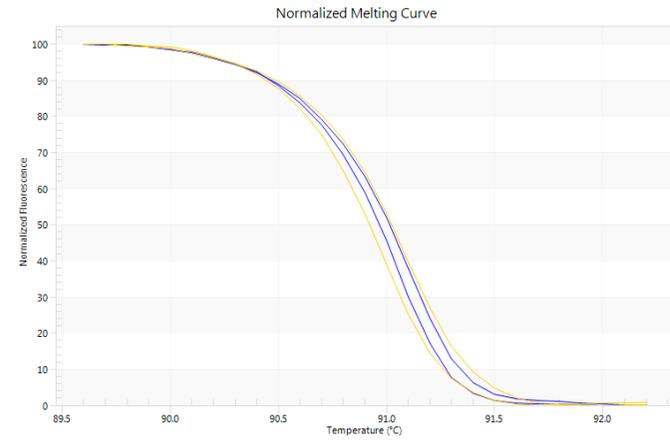
### HLAG19



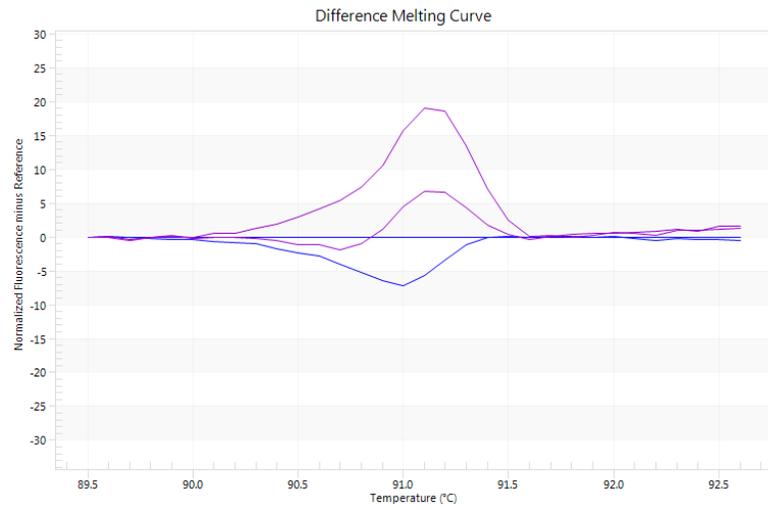
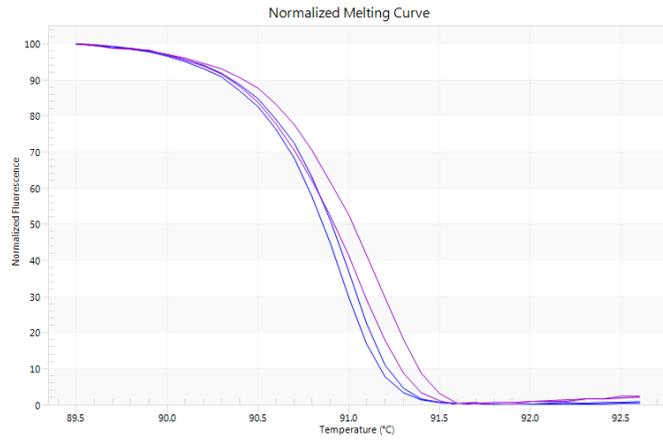
### HLAG26



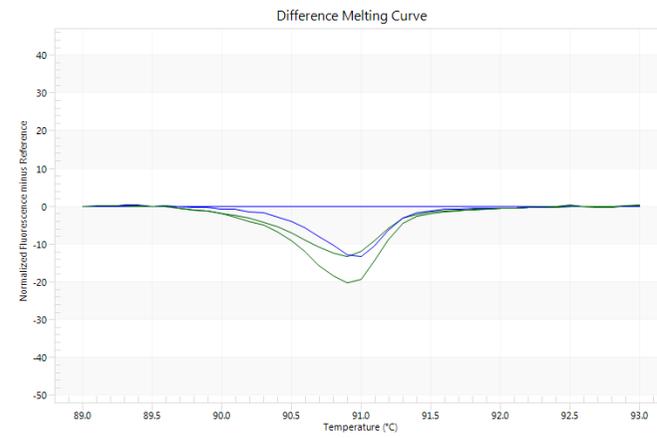
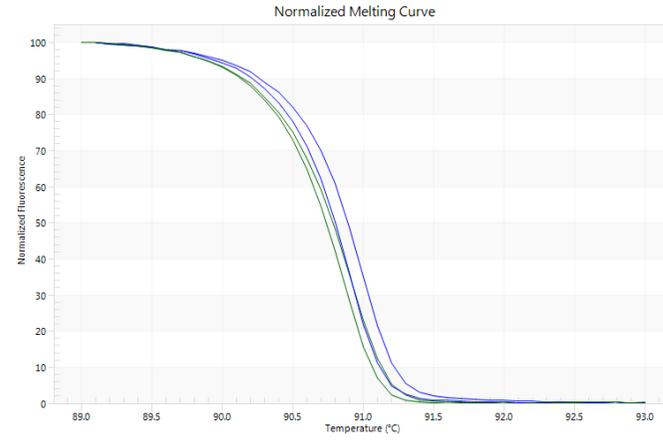
### HLAG28



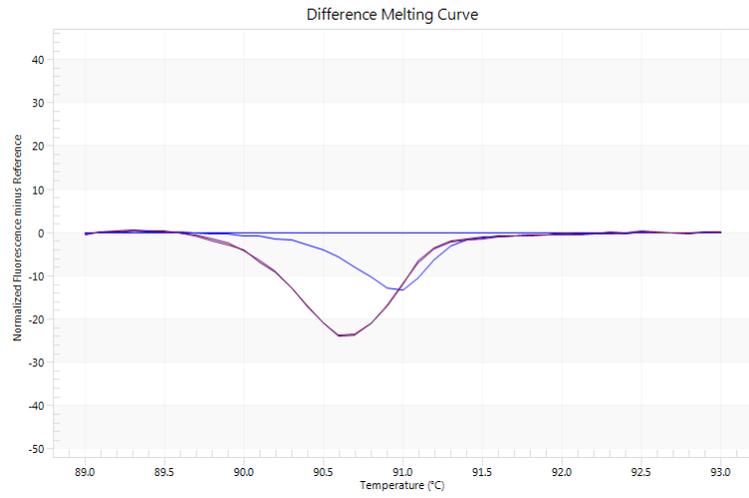
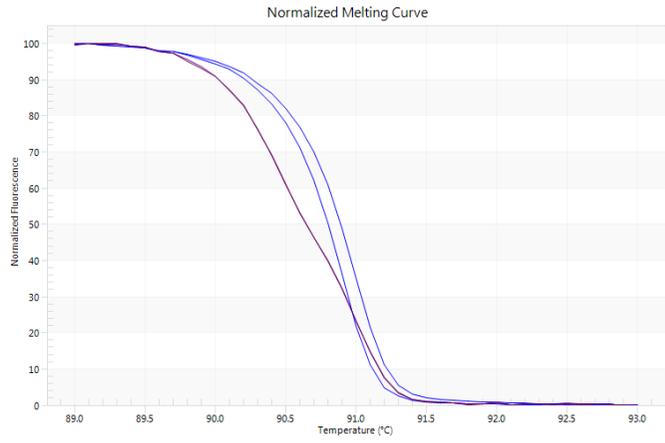
### HLAG29



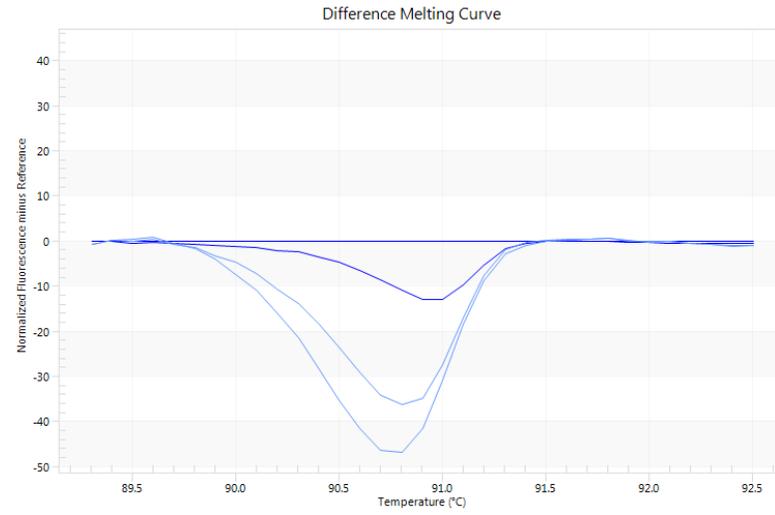
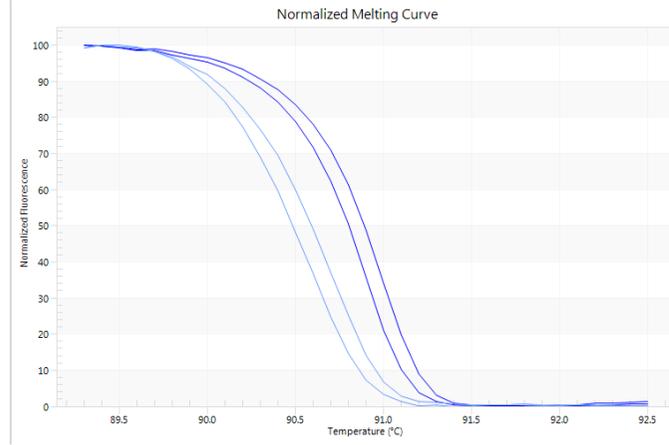
### HLAG38



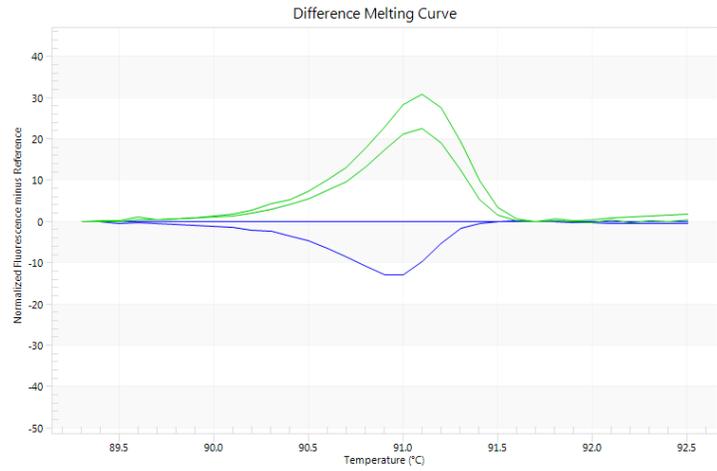
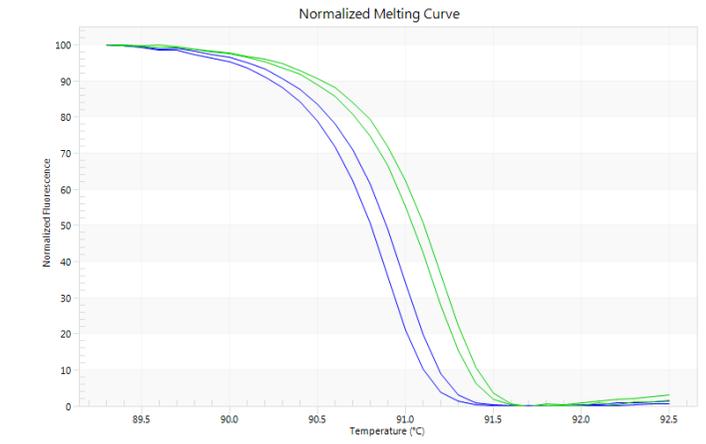
### HLAG39



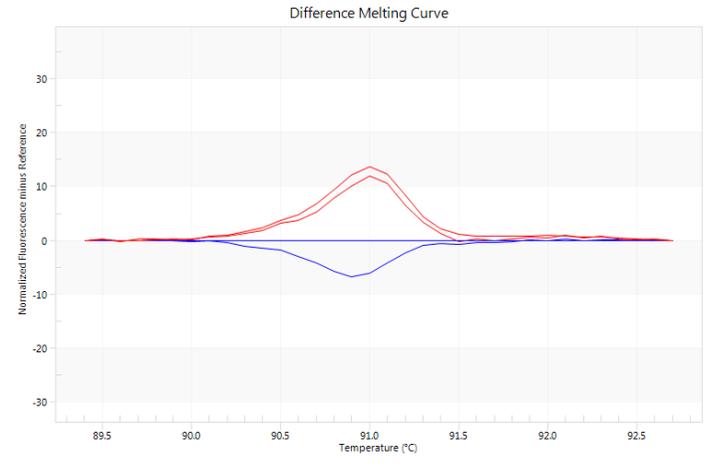
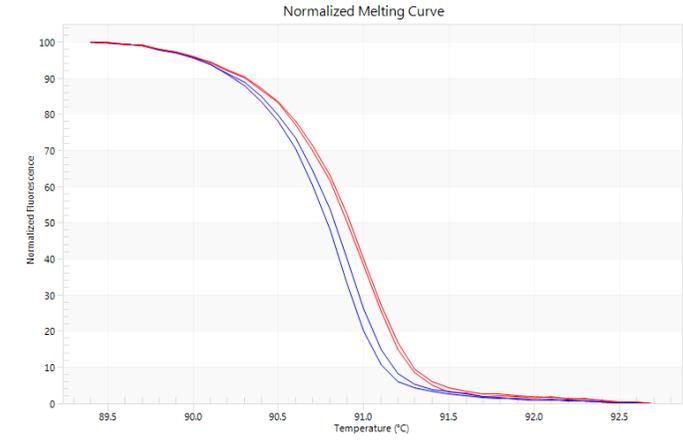
### HLAG46



### HLAG48

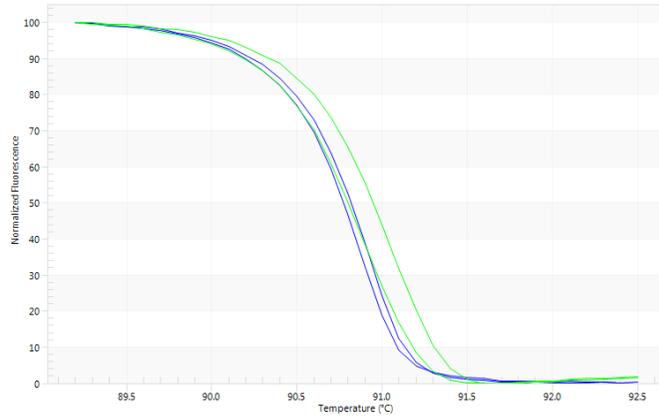


### HLAG53

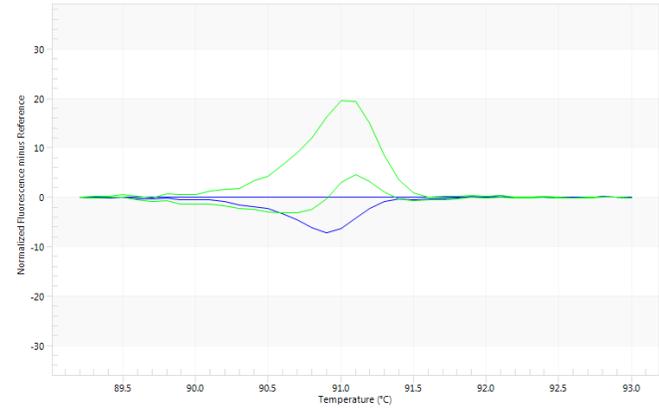


### HLAG56

Normalized Melting Curve

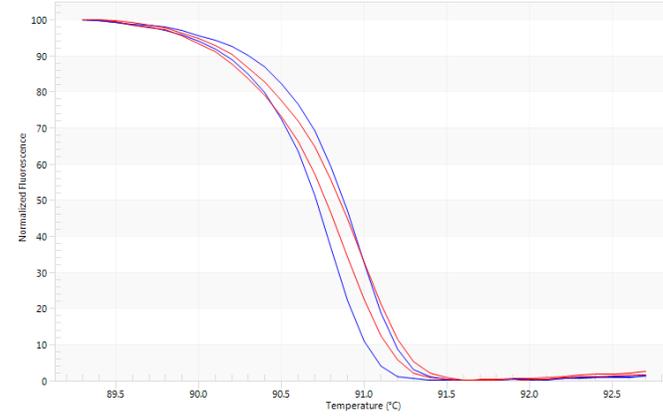


Difference Melting Curve

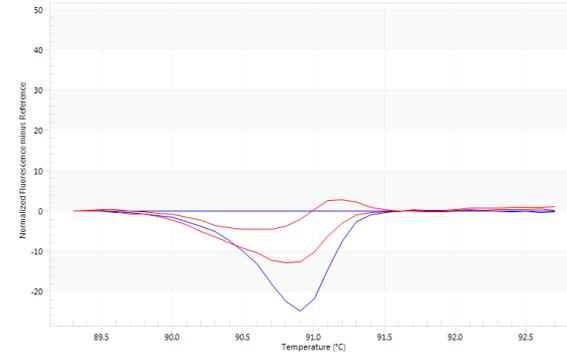


### HLAG64

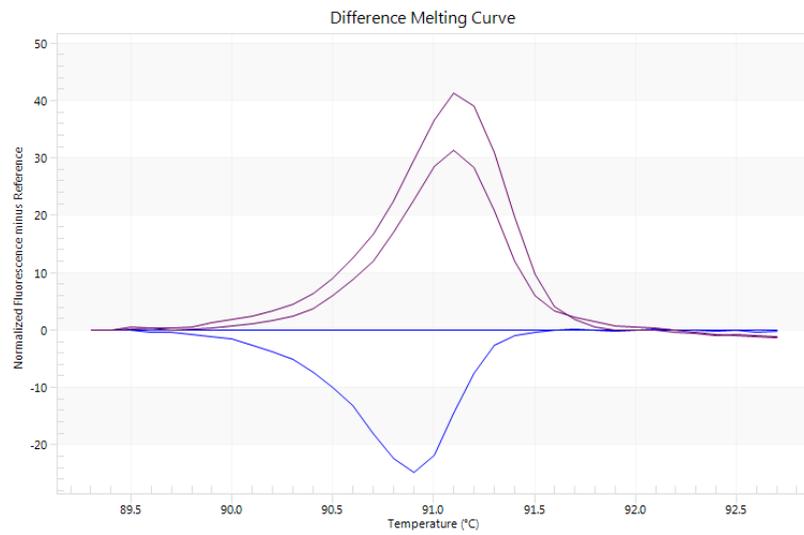
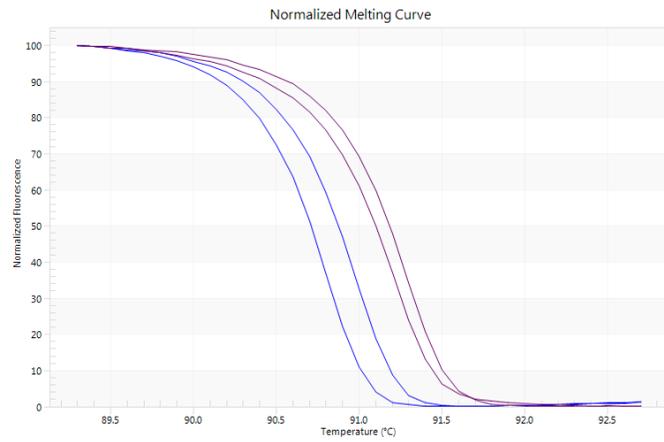
Normalized Melting Curve



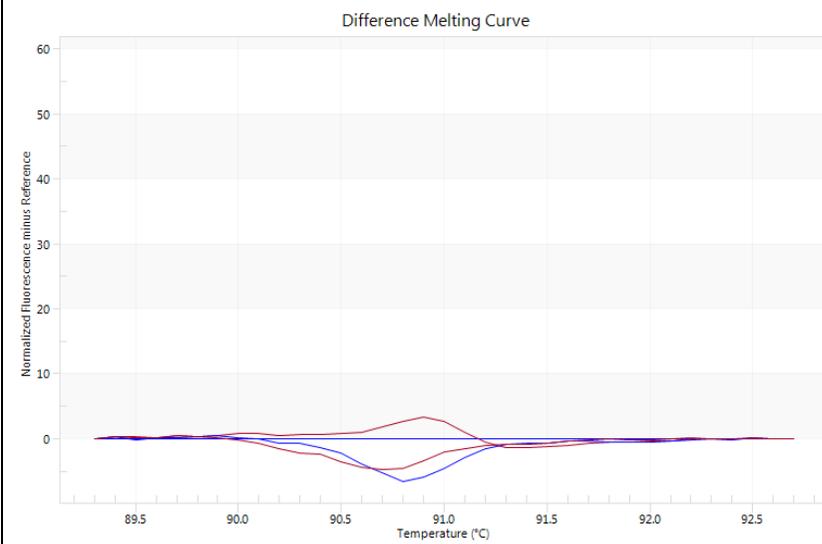
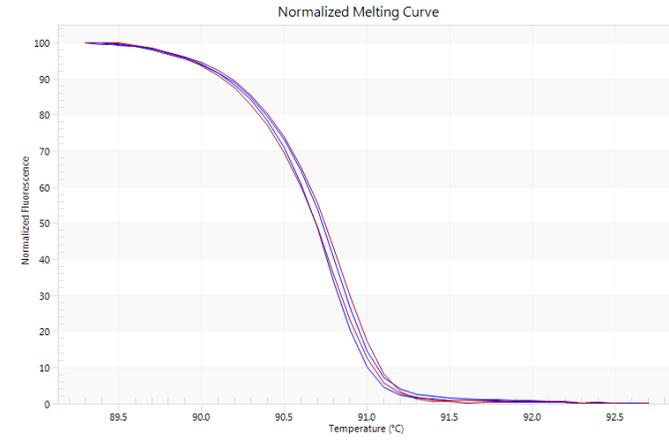
Difference Melting Curve



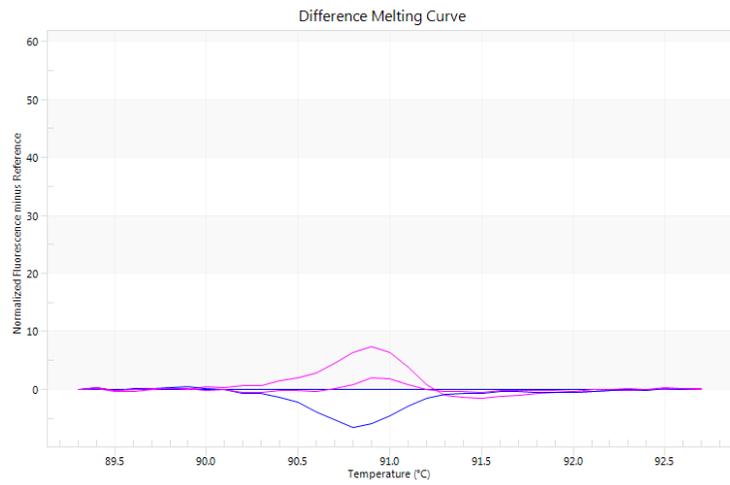
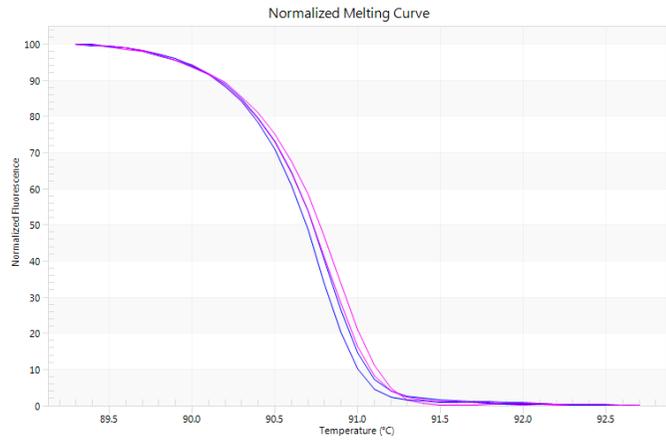
### HLAG66



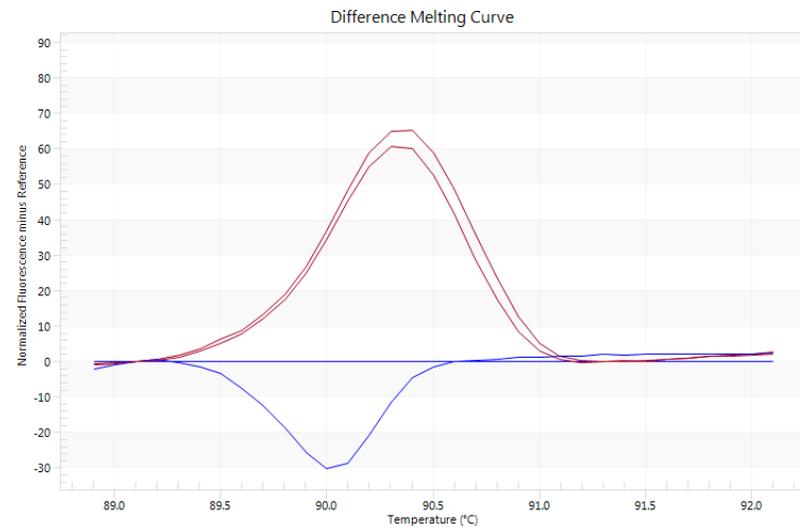
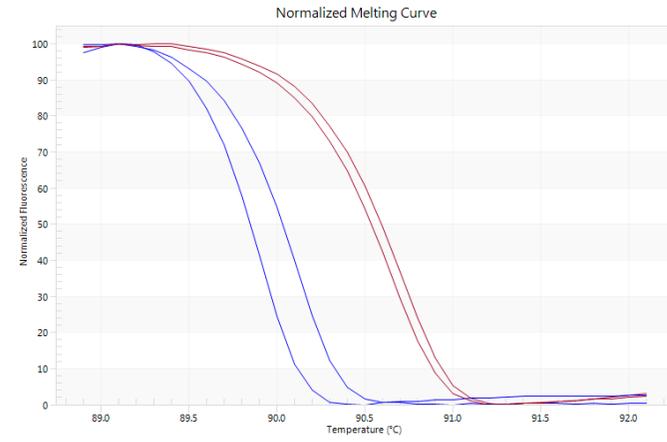
### HLAG68



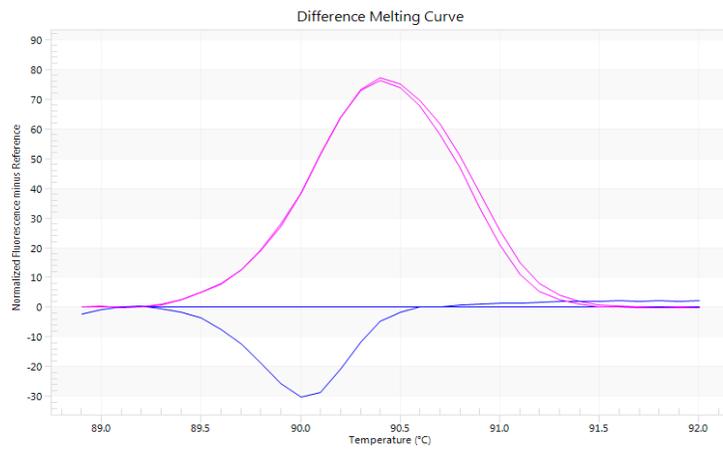
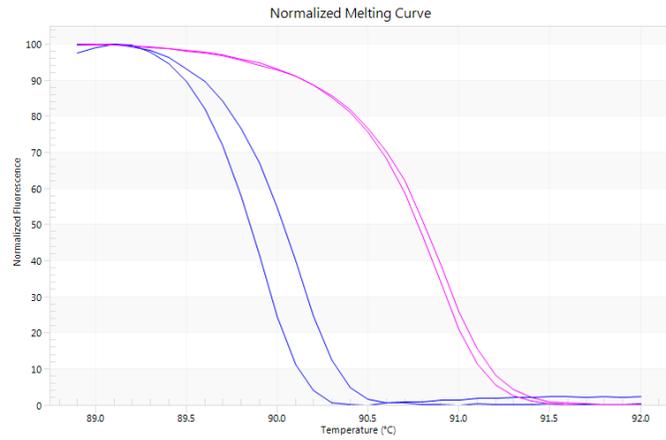
### HLAG69



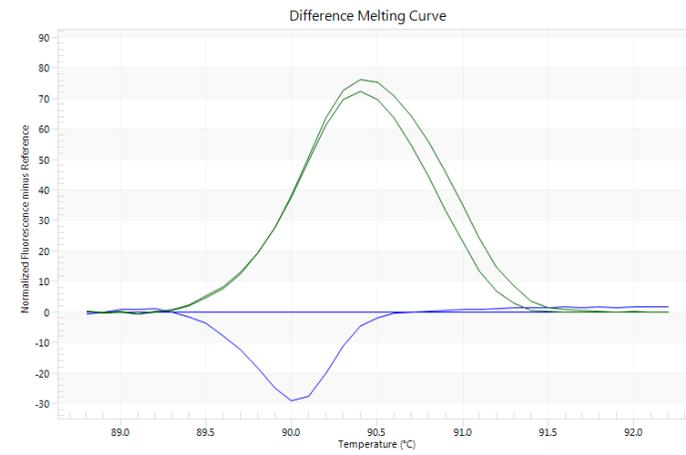
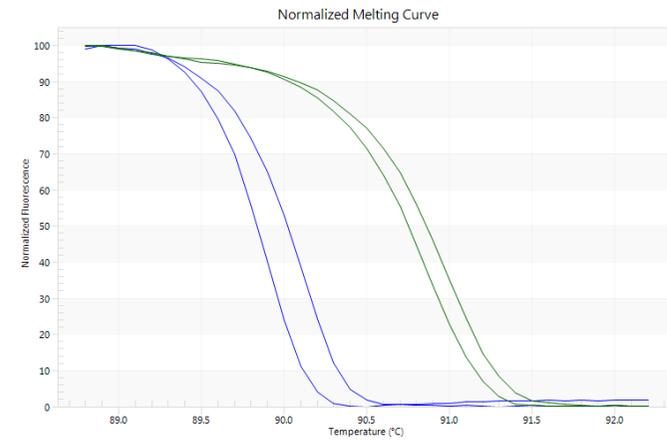
### HLAP2



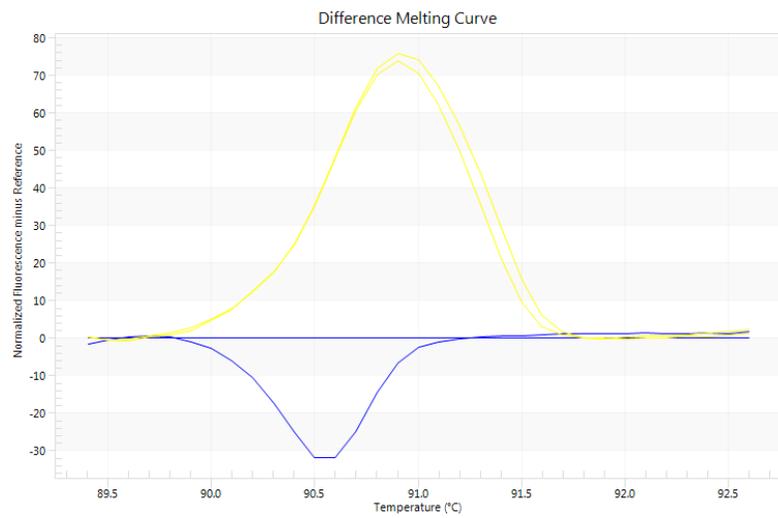
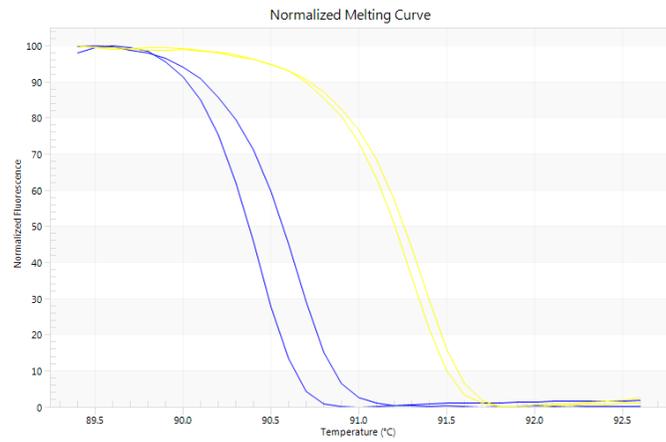
### HLAP3



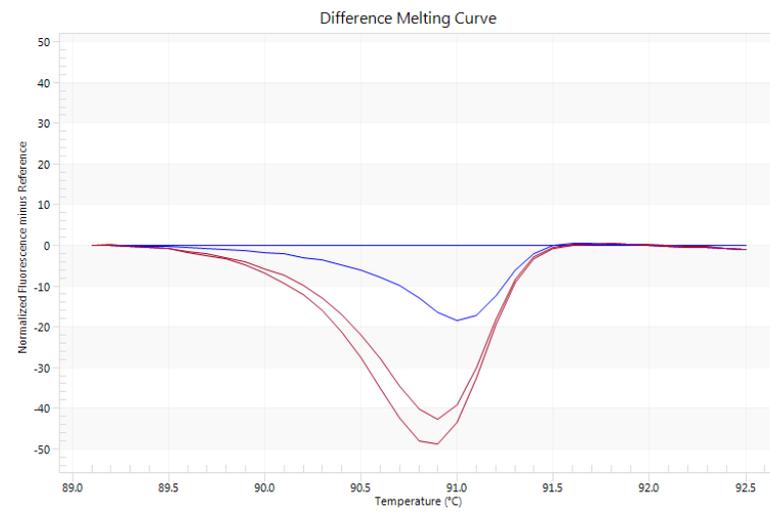
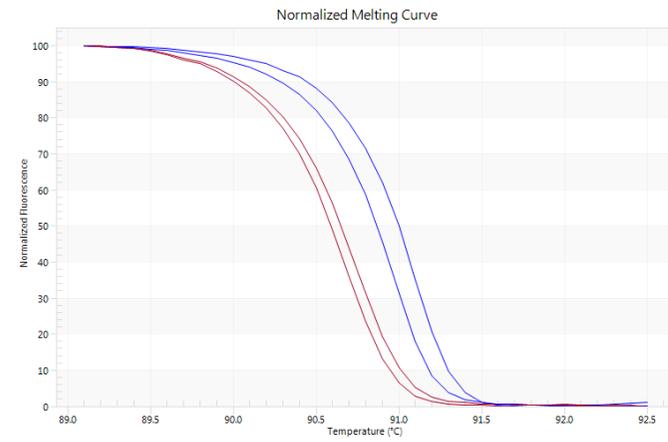
### HLAP6



### HLAP25

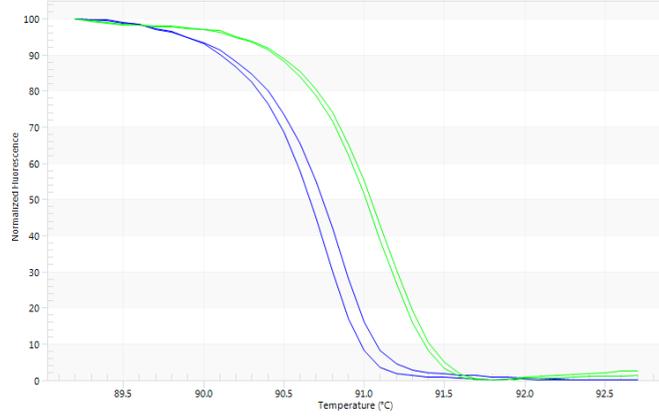


### HLAP37

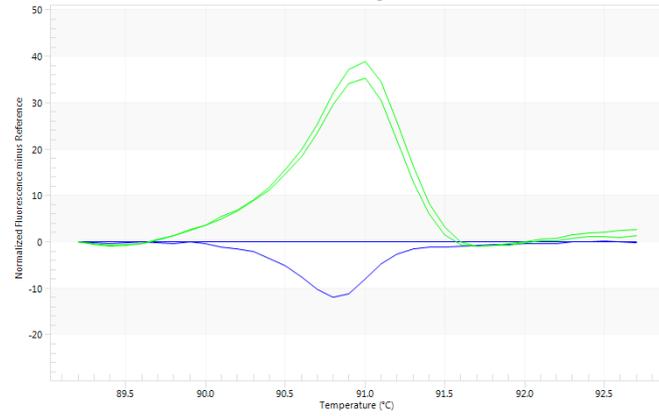


### HLAP46

Normalized Melting Curve

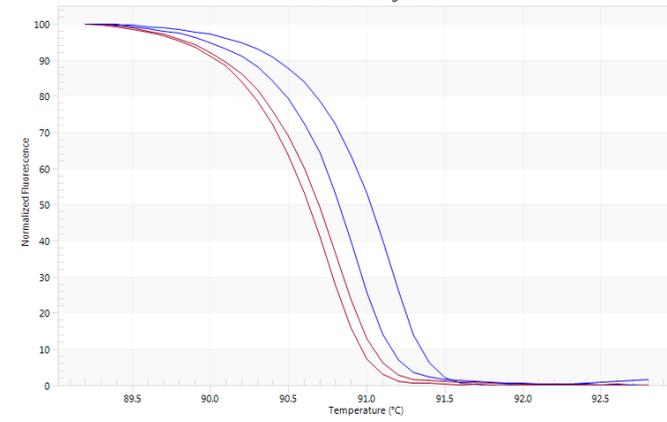


Difference Melting Curve

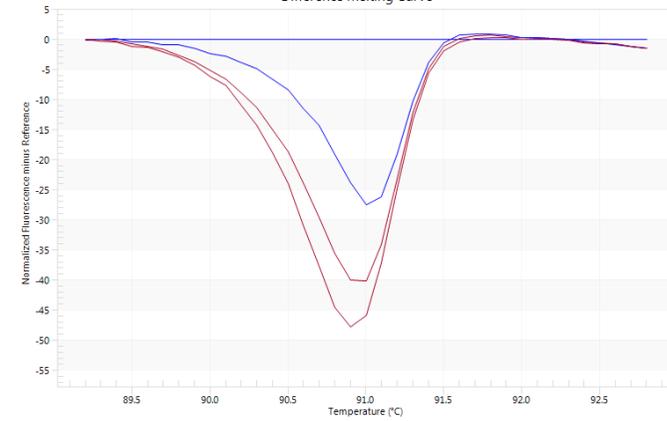


### HLAP51

Normalized Melting Curve

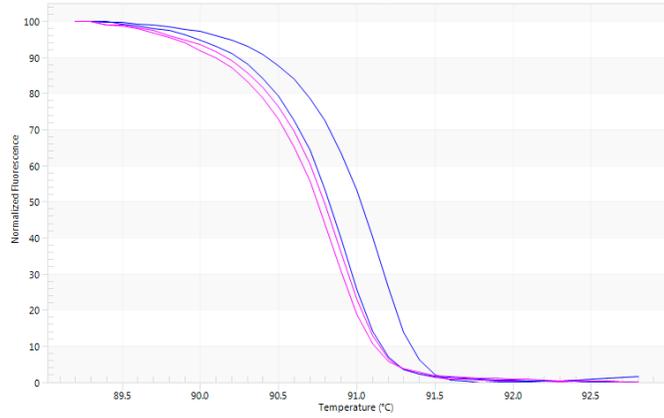


Difference Melting Curve

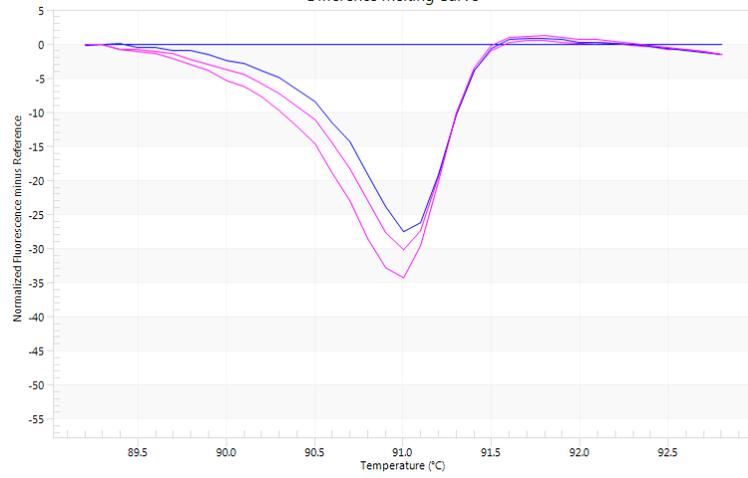


### HLAP52

Normalized Melting Curve

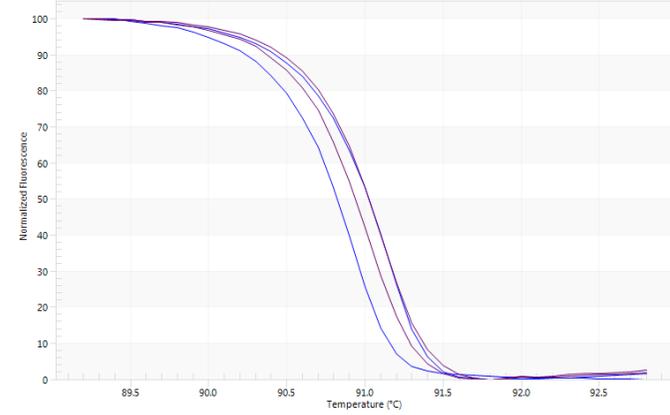


Difference Melting Curve

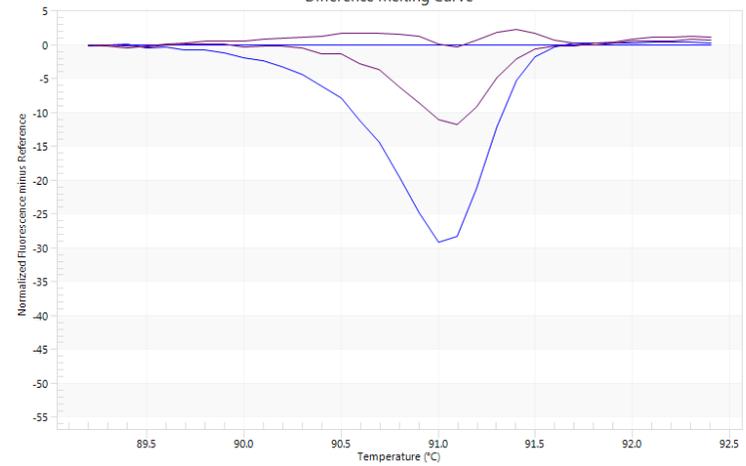


### HLAP56

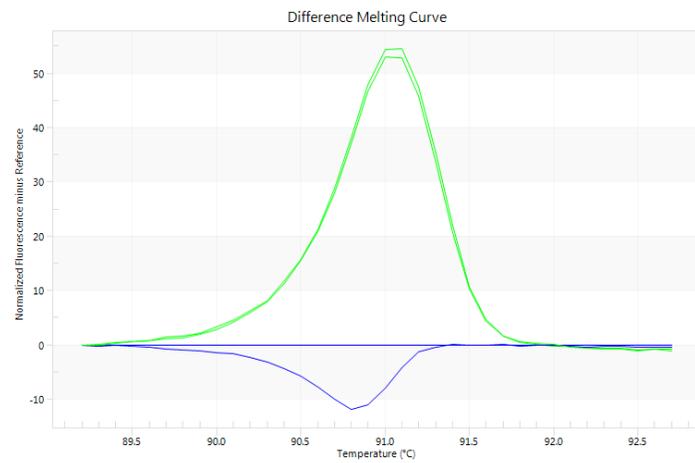
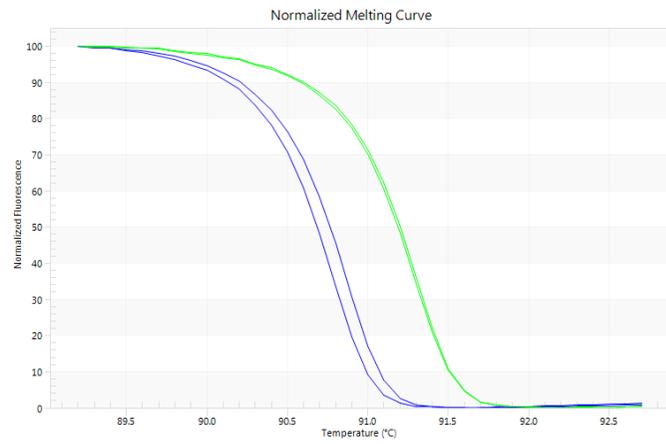
Normalized Melting Curve



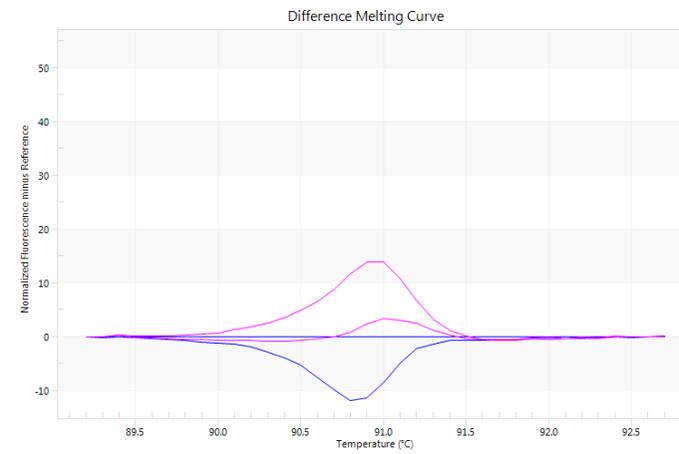
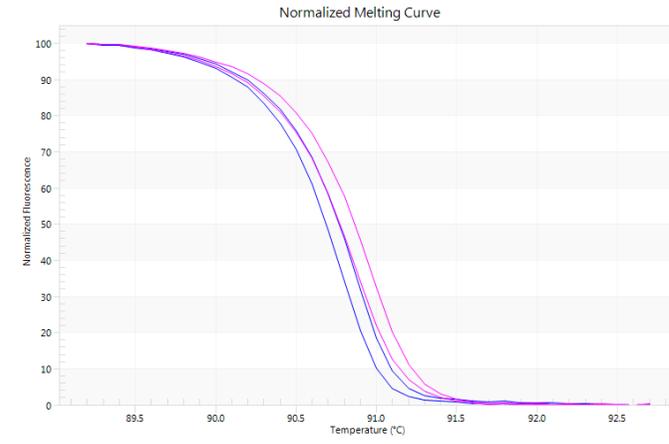
Difference Melting Curve



### HLAP60



### HLAP64



## 4.5 Sanger Sequencing results

### 4.5.1 Sanger sequencing on 28 ADR samples

The results gathered after the full HRM screening of 290 samples did not match the initial theory set and this prompted further and deeper analysis by sequencing. Sanger sequencing results were first analysed on BioEdit Sequence Alignment Editor in order to check the chromatograms generated and modify the DNA sequences. Once the sequence has been refined, it was inputted into two different databases, namely the NCBI website and the European Molecular Biology Laboratory-European Bioinformatics Institute website (EMBL-EBI), specifically the IPD-IMGT/HLA database. The IPD-IMGT/HLA database contains data for numerous human MHC sequences and forms part of the international ImMunoGeneTics project (IMGT). Numerous HLA allele hits were obtained for all the combinations obtained via Sanger sequencing and the highest number of hits helped in identifying the most common HLA allele present in a sample.

Chromatogram analysis was crucial due to the mis-calling of nucleotide bases caused by computer errors, but which could easily be spotted by humans. Ambiguous bases labelled as 'N' were sometimes easily identifiable by humans, or the opposite happened, where truly ambiguous peaks were wrongly identified as a base. These small errors go a long way in SNPs analysis, as one error can change the final HLA allele identified by adding or deleting SNPs. All the 28 ADR samples showed evenly-spaced, clear peaks with negligible baseline noise. Heterozygous peaks were identified by hand, where peaks were similar in height or with two different heights. All the different permutations and combinations obtained with the SNPs were used and inputted separately in NCBI and IMGT/HLA, in order to increase accuracy of HLA allele calls made. Once the DNA sequences were refined, we then moved on to HLA allele calling on NCBI and IMGT/HLA.

Table 4.10 shows the results for the 28 gout patients who showed ADRs/hypersensitivity reactions. These results again defied the expected, where samples with ADRs from gout patients taking allopurinol should be showing the presence of the genetic marker HLA-B\*58:01 only. Results from Table 4.11 showed an interesting number of HLA-B alleles,

where some of them have been identified as genetic markers for other disease-drug associations in the pharmacogenomics field. Twenty-four (85.7%) out of the twenty-eight ADR samples were all males and Malays. The remaining four samples had 3 female patients, with two Malays and one of mixed ethnic groups, and one last Chinese male patient.

IMGT/HLA results mostly yielded more general results with no decimal places to the HLA alleles, while more specific alleles were identified in NCBI, with up to 8 digits. There were numerous hits, as shown in Table 4.11, along with numerous SNPs identified (elaborated more in following sections). With those SNPs, different possible combinations of DNA sequences were entered into NCBI and IMGT/HLA for HLA allele calling (shown in Appendix G). One sample with 4 different rows, as shown for sample HLAG12, meant that there were 4 combinations of sanger sequencing data possible and all were inputted in the websites separately and the results collected. One sequence yielded around 100 results and the HLA allele with the most hits was identified as the final HLA allele in the sample.

Nine samples showed the presence of the HLA-B\*58:01 allele, namely HLAG12, HLAG26, HLAG29, HLAG56, HLAG66, HLAP6, HLAP25, HLAP46 and HLAP60. This totalled to 32.1% of the 28 ADR samples containing the HLA-B\*58:01 allele, as shown in Table 4.11. The HLA-B\*58:01 allele even presented itself up to 6 digits in all the 9 samples (HLA-B\*58:01:01). However, only around 10 hits of HLA-B\*58:01 alleles were found in the totality of the 100 hits. Table 4.12 shows that the HRM possible positives did not match the positives evaluated by Sanger sequencing. All the 9 samples which showed presence of the HLA-B\*58:01 allele as the first HLA-B allele and HLA-B\*35:01:01 as the second allele. Interestingly, these 9 samples were linked to range of varied ADRs namely, exfoliative dermatitis, DRESS syndrome, AHS, itchiness, SJS and generalized rashes, respectively. Ethnicity wise, all 9 samples were Malays and consisted of 8 males and 1 female patient. Other HLA-B\*58 alleles were identified in smaller numbers, namely HLA-B\*58:28, HLA-B\*58:34 and HLA-B\*58:65. Additionally, other HLA-B alleles were always found in combination in those 9 samples, such as HLA-B\*58:02 and HLA-B\*35:01:01:01, also present in 32.1% of the total ADR samples, as shown in Table 4.11. Some samples were heterozygous, such as HLA-B\*40 and HLA-B\*48 which were shown in the result hits as B\*40/B\*48. Other alleles present in significant numbers were the HLA-B\*15 and HLA-B\*15:02 alleles, making up 17.9% and 14.3% of the ADR samples. HLA-B\*44 and HLA-B\*44:03:01 were the next most abundant alleles at 10.7%

each. Two samples, HLAP3 and HLAP51, had ADRs to both allopurinol and febuxostat, but none showed the presence of the HLA-B\*58:01 allele. HLAP3 only had the HLA-B\*44:03:01 allele present, while HLAP51 showed the presence of two alleles, namely HLA-B\*15:02 and HLA-B\*14:02. Surprisingly, there were no HLA-A alleles found in all the 28 ADR samples. Furthermore, HLA-C alleles were also identified, with HLA-C\*04 and HLA-C\*04:01:01 having the highest presence at 14.3% each. Other HLA-C alleles identified were HLA-C\*08:01, HLA-C\*07/07:02:01/07:04, HLA-C\*12/12:02:01, HLA-C\*15 and HLA-C\*17.

Based on the HRM difference melt curve analysis in the previous section, the 8 possible positive samples with close melt curves were not found to contain the HLA-B\*58:01 allele by sanger sequencing, as shown in Table 4.12. Three samples contained the HLA-B\*15:02 allele, two contained the HLA-B\*44:03:01 allele, while others contained the HLA-C\*07:04 allele and the heterozygous HLA-B\*40/48 alleles. This showed that the matching of difference melt curves cannot be used as a definite HRM screening result, due to the presence of other HLA alleles. Further investigations will be necessary to elucidate the role of those HLA alleles and their link to the HLA-B\*58:01 allele and their accompanying ADRs in gout patients. Following sections will investigate, SNPs present, along with  $T_m$  changes and the GC content of those ADR samples. These results challenge the one HLA-B allele association with one disease-drug association and shows that multiple HLA-B alleles could be responsible for the varied range of ADRs found in gout patients.

Table 4-10 HLA allele calling with the Sanger Sequencing results of the 28 gout samples with ADRs from NCBI and IMGT/HLA website. The final 2 HLA-B alleles for each individual were shown in the fourth column, where the highest hit was called and linked to the ADRs observed in the samples. HLA-B alleles are simplified to B\* and HLA-C alleles to C\*. Some samples will have several rows due to the presence of numerous and corresponding SNPs.

Sample Code	NCBI BLAST results	IMGT/HLA BLAST results	Final 2 HLA-B alleles called	ADR observed	Gender	Ethnicity
HLAG12	B*35:01:01:01	B*58	B*58:01:01/B*35:01:01:01	Vasculitis rash	Male	Malay
	B*58:01:01					
	B*58:01					
	B*58:02					
HLAG13	C*08:01	C*08	C*08:01/B*40:02:01	Mild allergy- itchiness	Male	Malay
	C*08:01					
	B*40/48					
	B*40:02:01					
HLAG18	C*04:01:01	C*04	C*04:01:01	Mild allergy- rash	Male	Malay
	C*04:01:01					
HLAG19	C*07:04	C*07	C*07:04	Mild allergy- itchiness	Male	Malay
	C*07:04					
HLAG26	B*35:01:01:01	B*58	B*58:01:01/B*35:01:01:01	Exfoliative dermatitis	Male	Malay
	B*58:01:01					

	B*58:01 B*58:02					
HLAG28	B*44:03:01	B*44	B*44:03:01	Mild allergy- itchiness	Female	Malay
	B*44:03:01	B*44				
	B*44:03:01	B*44				
HLAG29	B*35:01:01:01	B*58	B*58:01:01/B*35:01:01:01	DRESS syndrome	Male	Chinese
	B*58:01:01					
	B*58:01					
	B*58:02					
HLAG38	B*15:02	B*15	B*15:02	SJS	Male	Malay
HLAG39	C*12	C*12	C*12:02	SJS	Female	Other
	C*12	C*12				
	C*12:02	C*12				
	C*12:02	C*12				
HLAG46	B*35:02	B*35 B*58 B*58:65 B*58:28	B*35:02/B*58	Allopurinol hypersensitivity syndrome- red eye + severe itchiness	Male	Malay
	B*35:02	B*35 B*58 B*58:65 B*58:28				

HLAG48	B*35:03:01	B*35 B*58:34	B*35:03:01/B*58:34	Tophi ruptured, allergy to allo	Male	Malay
HLAG53	B*35:03:01	B*35 B*58:34	B*35:03:01/B*58:34	Rashes	Male	Malay
HLAG56	B*35:01:01:01 B*58:01:01 B*58:01 B*58:02	B*58	B*58:01:01/B*35:01:01:01	DRESS- rash	Male	Malay
HLAG64	B*15:02	B*15	B*15:02	Rashes	Male	Malay
	B*15:02	B*15				
HLAG66	B*35:01:01:01 B*58:01:01 B*58:01 B*58:02	B*58	B*58:01:01/B*35:01:01:01	Itchiness	Male	Malay
HLAG68	B*15:02	B*15	B*15:02	Allergy to allo- Anaphylaxis reaction	Male	Malay
HLAG69	B*40:01 B*48	B*48	C*04:01:01:01/B*48	Allergy to allopurinol	Male	Malay
	C*04:01:01:01	C*04				
	C*04:01:01:01	C*04				
HLAP2	C*04:01:01:01	C*04	C*04:01:01:01	Generalised rash	Male	Malay
HLAP3	B*44:03:01	B*44	B*44:03:01	Allergy to allopurinol-	Male	Malay

				generalised rash + face swelling. Allergy to Febuxostat - rash		
HLAP6	B*35:01:01:01 B*58:01:01 B*58:01 B*58:02	B*58	B*58:01:01/B*35:01:01:01 (ambiguous)	SJS ( mouth and genitalia)	Male	Malay
	B*15 B*57	B*57				
	B*35:01:01:01 B*58:01:01 B*58:01 B*58:02	B*58				
HLAP25	B*35:01:01:01 B*58:01:01 B*58:01 B*58:02	B*58	B*58:01:01/B*35:01:01:01	Generalized skin rash	Male	Malay
HLAP37	B*56	B*56	C*04:01:01	Generalized rash	Male	Malay
	C*04:01:01	C*04				
	C*04:01:01	C*04				
HLAP46	B*35:01:01:01 B*58:01:01	B*58	B*58:01:01/B*35:01:01:01	rash on face	Female	Malay

	B*58:01 B*58:02					
HLAP51	B*15:02	B*15	B*15:02/B*14:02	Allopurinol- exfoliative dermatitis + febuxostat allergy	Male	Malay
	B*15:02	B*15				
	B*14:02	B*14				
HLAP52	C*12:02:01	C*12	C*12:02:01	Allergy to allopurinol-transaminitis- no rash seen. Another type of allergy?	Male	Malay
HLAP56	B*44:03:01	B*44	B*44:03:01	Generalized rash+ skin eruptions (raised and flat eruptions), macular popular rash	Male	Malay
	B*44:03:01	B*44				
	B*44:03:01	B*44				
HLAP60	B*35:01:01:01 B*58:01:01 B*58:01 B*58:02	B*58	B*58:01:01/B*35:01:01:01	Generalised rashes	Male	Chinese
HLAP64	B*40:01:02 B*48	B*48	HLA-B*40/HLA-B*48	Severe allergy to allopurinol- angioedema + eye swelling	Male	Malay
	C*07:02:01	C*15				
	B*40:02:01	C*17				

Table 4-11 Percentage of each HLA-B allele found in all the 28 ADR samples, as per Sanger sequencing.

HLA subgroup	HLA alleles identified	Number of samples containing the allele	Total percentage of samples containing the allele in 28 ADR samples
<b>HLA-B</b>	HLA-B*58:01	9	32.1%
	HLA-B*58:02	9	32.1%
	HLA-B*58:01:01	9	32.1%
	HLA-B*58	9	32.1%
	HLA-B*15	5	17.9%
	HLA-B*15:02	4	14.3%
	HLA-B*35:01:01:01	9	32.1%
	HLA-B*35:02	1	3.6%
	HLA-B*35:03:01	2	7.1%
	HLA-B*40/ HLA-B*48	1	3.6%
	HLA-B*40:01	2	7.1%
	HLA-B*40:02:01	2	7.1%
	HLA-B*44:03:01	3	10.7%
	HLA-B*44	3	10.7%
	HLA-B*48	2	7.1%
	HLA-B*57	1	3.6%
	HLA-B*56	1	3.6%
	HLA-B*14:02	1	3.6%
	HLA-B*14	1	3.6%
	B*58:65	1	3.6%
B*58:28	1	3.6%	
B*58:34	2	7.1%	

Table 4-12 Comparison of possible HRM HLA-B\*58:01 positive samples to HLA-B\*58:01 positive samples in Sanger sequencing, for the 28 ADR samples.

Sample code	HRM result	Sanger sequencing result
HLAG12	Negative	Positive
HLAG13	Negative	Negative
HLAG18	Negative	Negative
HLAG19	Positive	Negative
HLAG26	Negative	Positive
HLAG28	Positive	Negative
HLAG29	Negative	Positive
HLAG38	Positive	Negative
HLAG39	Negative	Negative
HLAG46	Negative	Negative
HLAG48	Negative	Negative
HLAG53	Negative	Negative
HLAG56	Negative	Positive
HLAG64	Positive	Negative
HLAG66	Negative	Positive
HLAG68	Positive	Negative
HLAG69	Positive	Negative
HLAP2	Negative	Negative
HLAP3	Negative	Negative
HLAP6	Negative	Positive
HLAP25	Negative	Positive
HLAP37	Negative	Negative
HLAP46	Negative	Positive
HLAP51	Negative	Negative
HLAP52	Negative	Negative
HLAP56	Positive	Negative
HLAP60	Negative	Positive
HLAP64	Positive	Negative

## 4.5.2 Further analysis done on HRM results

### 4.5.2.1 SNPs analysis with Melting point ( $T_m$ ) and GC content

Once the HRM difference melt curve was analysed, we moved on to try to find SNPs from the Sanger sequencing results and link it back to their base change and melting point ( $T_m$ ) shift. Table 4.13 below shows the standard method used during HRM screenings in order to find the SNPs present and base change which is in turn affecting the  $T_m$ . Each class of SNP was related to a base change, along with a specific increase in temperature and curve shift as shown in Table 4.13 below. Using this method, all the 28 ADR samples were linked to their possible SNPs by aligning their Sanger sequencing FASTA sequence to their previously identified HLA allele hits. This was shown in detail for a few samples in Appendix G, with several aligned data sets and the final results was inputted into Table 4.14 and 4.15. In many cases, due to the presence of numerous SNPs, the total base change was cancelled out.

Table 4-13 Standard SNP identification method by using base change and melting point of samples as variables.

SNP class	Base change	Typical $T_m$ curve shift	Frequency in the human genome
1	C/T and G/A	Large ( $>0.5^\circ\text{C}$ )	64%
2	C/A and G/T	Large ( $>0.5^\circ\text{C}$ )	20%
3	C/G	Small ( $0.2\text{-}0.5^\circ\text{C}$ )	9%
4	A/T	Small ( $<0.2^\circ$ )	7%

The difference in  $T_m$  is seen due to the DNA sequence variation (base change) in different samples compared to the reference DNA sequence used. The greatest change in fluorescence level will be seen at the  $T_m$  of the product, which is the point in a melt curve where 50% of DNA is in the dsDNA form and the other 50% is in the ssDNA form. Targets having higher GC content or a greater length will automatically have a higher  $T_m$  value. The difference in  $T_m$  was obtained from the melting point average recorded during the HRM run and by subtracting a sample's  $T_m$  from that of the positive control. This calculation was

shown in detail in Appendix G and the final results inputted in Table 4.14 and 4.15 on the next page.

From the first table for the 9 possible positive samples, the first two rows itself showed the same allele hit and base change getting different temperature shifts from their HRM melt curve. Moreover, these 2 samples (HLAG26 and HLAG29) showed a base change only for one HLA allele hit out of the four different hits previously identified by Sanger sequencing. It was difficult to identify a trend due to the erratic melting point change seen for similar base changes and samples with similar HLA allele hits. Table 4.15 showed the same erratic results, with the addition of 4 null and 5 negative values for the  $T_m$  change. Hence, these methods usually used failed to identify a specific trend for positive and negative samples due to erratic results.

Table 4.16 showed the compilation of all the possible methods used to define the HRM trend in one table for better comparison purposes, for the 9 possible positive samples. This was done to try to link the base changes noted from the SNPs analysis to the HRM difference melt curve. Samples HLAG26 and HLAG29 both had a possible base change of two cytosine molecules (CC), but showed different melt curve shapes. Similar base changes usually caused a similar change in the shape of the HRM difference melt curve, hence allowing the association of specific base changes to a specific peak or curve shift on the difference melt curve. Samples HLAG46 and HLAG56 show this same erratic result, with different  $T_m$  changes for the same possible base change (TAA) and the same HLA allele linked to the possible SNP. Their HRM difference melt curves showed completely opposite results with HLAG46 showing downward double-peaks, while HLAG56 showed upward peaks with different slopes and shape. Thus, the HRM difference melt curve shift cannot be quantified and linked to specific base changes, SNPs and  $T_m$  changes. Other variables must be present in all the aforementioned results, which were in turn changing the results obtained from the expected results.

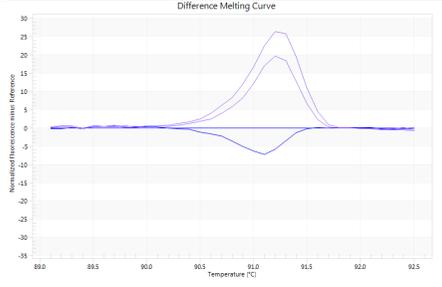
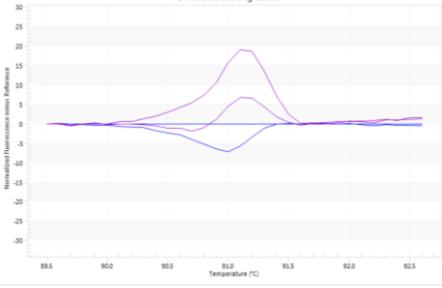
Table 4-14 Base changes seen in the 9 possible positive samples, along with their linked HLA alleles and T<sub>m</sub> change.

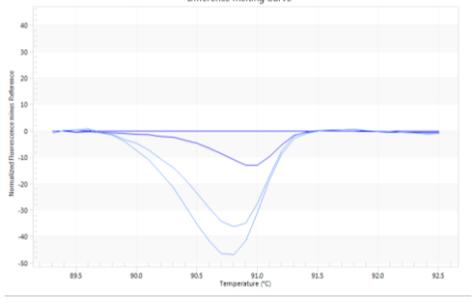
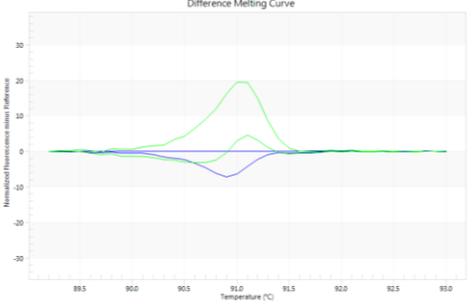
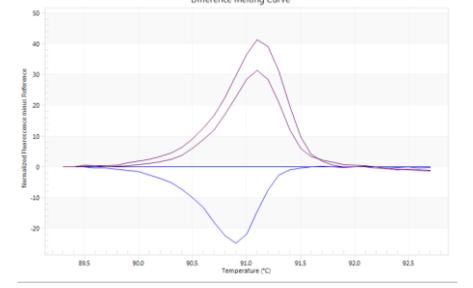
Sample	Base change noticed in Sanger sequencing	HLA-B allele discovered linked to base change	T <sub>m</sub> change (°C)
HLAG26	CC	B*58:02	2.7
HLAG29	CC	B*58:02	0.1
HLAG46	TAA	B*35	0.4
HLAG56	TAA	B*35	0.1
HLAG66	TAA or CC	B*35 or B*58:02	0.4
HLAP6	None	None	0.85
HLAP25	TAA or CC	B*35 or B*58:02	0.8
HLAP46	GGGG or TAA	B*15 or B*35	0.35
HLAP60	TAA	B*35	0.55

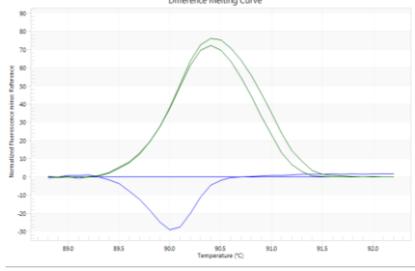
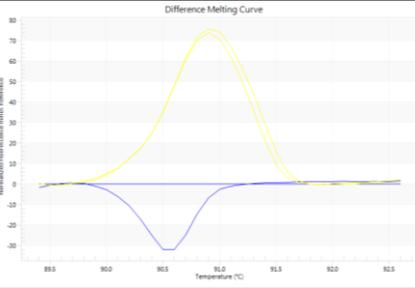
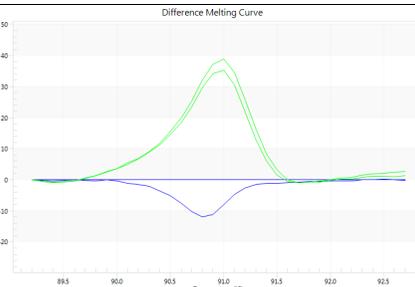
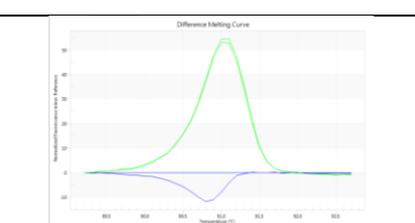
Table 4-15 Base changes seen in the 19 negative samples, along with their linked HLA alleles and T<sub>m</sub> change.

Sample	Base change noticed in Sanger sequencing	HLA-B allele discovered linked to base change	Temperature change (°C)
HLAG12	None	None	0.3
HLAG13	GA or AG	B*40 or B*48	0.1
HLAG18	G	C*04	0.2
HLAG19	CCCC	C*07	0
HLAG28	GGAA	B*44:02	0
HLAG38	GGGG or GGGA	B*15:01 or B*15:02	0.15
HLAG39	GC	B*12:02:02	0.25
HLAG48	T	B*35:03:01	0.2
HLAG53	T or AA	B*35:03:01 or B*35:02:02	-0.05
HLAG64	GGGG or GGGA	B*15:01 or B*15:02	0
HLAG68	GGGG or GGGA	B*15:01 or B*15:02	0
HLAG69	GA or AG	B*40 or B*48	-0.05
HLAP2	G	C*04	-0.6
HLAP3	GGA or GGAA	44:03:01 or B*44:02	-0.95
HLAP37	TTA	B*56	0.35
HLAP51	GGGA	B*15:02	0.3
HLAP52	GC	B*12:02:02	0.15
HLAP56	GGA	B*44:03:01	-0.05
HLAP64	GA or AG	B*40 or B*48	-0.1

Table 4-16 Compilation of all the different variables used to identify a trend in the HRM results for the 9 possible positive ADR samples. Sanger sequencing results from Table 4.11 were also added into 2 columns based on their databases. HRM difference melt curves and ADRs observed were shown in the last two columns.

Sample	Temperature change (degrees)	Possible base change + associated allele	NCBI BLAST	IMGT/HLA BLAST	HRM difference melt curve	ADRs manifested
HLAG26	2.7	CC (58:02)	B*53, B*35, B*58:01, B*58:02	All B*58	 <p>Difference Melting Curve for HLAG26. The y-axis is 'Normalized Fluorescence minus Reference' ranging from -35 to 30. The x-axis is 'Temperature (°C)' ranging from 89.0 to 92.5. The curve shows a positive peak of approximately 25 at 91.2°C and a negative peak of approximately -15 at 91.2°C.</p>	Exfoliative dermatitis
HLAG29	0.1	CC (58:02)	B*53, B*35, B*58:01, B*58:74, B*58:02	All B*58	 <p>Difference Melting Curve for HLAG29. The y-axis is 'Normalized Fluorescence minus Reference' ranging from -30 to 30. The x-axis is 'Temperature (°C)' ranging from 89.5 to 92.5. The curve shows a positive peak of approximately 20 at 91.2°C and a negative peak of approximately -10 at 91.2°C.</p>	DRESS Syndrome

HLAG46	0.4	TAA (B*35)	B*53, B*35, B*58:28, B*35:02,	Numerous B*35 , 58:65		Allopurinol hypersensitivity syndrome- red eye + severe itchiness
HLAG56	0.1	TAA (B*35)	B*53, B*35, B*58:62, B*58:74	All B*58		DRESS- rash
HLAG66	0.4	TAA (B*35) OR CC (58:02)	B*53, B*35, B*58:01, 58:02, B*58:74	All B*58		Itchiness

HLAP6	0.85	-	B*15, B*15:02, 58:06, 58:01, 58:02	B*15, B*44, B*53, 58:01, 58:91, 58:19		SJS ( mouth and genitalia)
HLAP25	0.8	TAA (B*35) OR CC (58:02)	B*35, B*53, B*15, , 58:74, 58:01, 58:02	All B*58		Generalized skin rash
HLAP46	0.35	B*15(GGGG) OR TAA (B*35)	B*35, B*53, B*15, , 58:01, 58:02, 57	All B*15, 58:97		rash on face
HLAP60	0.55	TAA (B*35)	B*35, B*53, B*15, , 58:01, 58:02, 58:74	All B*58		Generalized rashes

### 4.5.3 Combined HRM analysis and Sanger sequencing results for healthy volunteers

This section will use all the aforementioned HRM analysis methods in order to find a trend in the normal, healthy volunteer batch. Table 4.18 combined the HRM difference melt curves, Sanger sequencing results, gender and ethnic group, in order to allow proper analysis. 10% of the normal sample cohort (14 samples) were sent for Sanger sequencing and samples were selected randomly, in order to avoid bias.

In terms of HRM melt curve shape analysis and by using the aforementioned threshold (20 to -20), 51 samples were called as positive and 94 samples as negative for the HLA-B\*58:01 allele. This shows a greater percentage of samples being called as positive (35%) compared to the gout cohort (15%). In the 14 sequenced samples specifically, 6 samples had close HRM curves (threshold 20 to -20), namely, NS97, NS11, NS6, NS26, NS28, NS95.

Interestingly, Table 4.17 showed the presence of 4 samples with the HLA-B\*58:01:01 allele, along with the previously linked HLA-B\*35:01:01 allele. These samples were N54, N33, NS37 and NS95, with a distribution of gender and race more or less similar to the ADR samples. Sample NS95, showed the presence of HLA-B\*58:01 allele via sanger sequencing and a possible positive HRM curve. This shows that HLA-B\*58:01 is a significant allele seen, even in the absence of gout and accompanying ADRs. The most abundant alleles (in Table 4.17) were HLA-B\*15, HLA-B\*35, HLA-B\*53, followed by their variants HLA-B\*15:02, HLA-B\*15:21, HLA-B\*35:01:01, HLA-B\*53:01:01, as well as HLA-B\*58 and HLA-B\*58:01:01.

The sequencing results for the normal samples showed HLA-B\*15/35/53 as the highest hit, compared to HLA-B\*58/35, and their several variants, being the highest in the ADR samples. Moreover, more variants of the HLA-B\*58 allele were identified at higher percentages in gout/ADR samples, compared to the normal samples. This shows that HLA-B\*58 and its variants are associated and seen more in gout samples with ADRs, compared to the normal sample cohort. Other alleles, namely HLA-B\*57 and HLA-B\*40/48 were again not found in normal samples, but only in gout/ADR samples. More variants of the HLA-B\*15 allele were found in the normal samples, compared to the gout/ADR samples. The HLA-B\*35 allele had similar variants in both cohort, but were present at a higher percentage in normal samples. The HLA-B\*13 alleles were found only in normal samples, and not in gout/ADR samples. This

shows a trend, where HLA-B alleles previously linked closely to the HLA-B\*58:01 allele in gout/ADR samples were not present in normal samples.

No trend was seen again in HRM curve shape amongst the 4 positive samples identified, similar to ADR samples. NS6 was the only sample showing the presence of the HLA-B\*35:01 and HLA-B\*53:01 alleles, without any HLA-B\*58:01 hit like the 4 normal, positive samples. Samples identified with close HRM curves, namely NS11, NS26 and NS28, all had HLA-B\*15 alleles as the highest hit, thus showing that HRM curve shape was influenced by the presence of numerous alleles, and not just the desired, HLA-B\*58:01 allele. This was further validated by sample NS68, which had the HLA-B\*15 alleles as highest hit, but had a huge difference in the HRM curve shape.

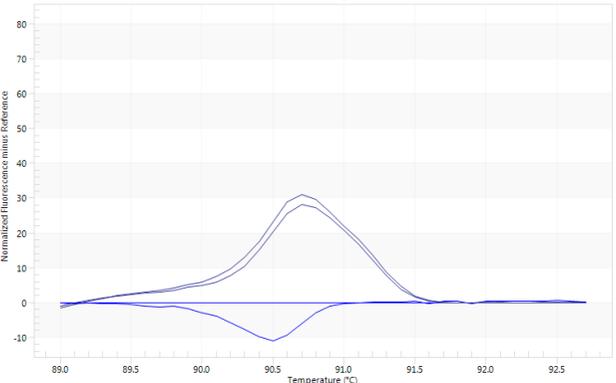
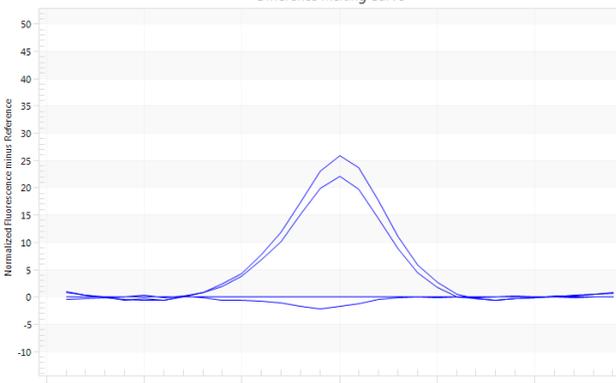
The sequencing results for the normal samples confirm what was previously seen in gout/ADR samples. Numerous HLA-B alleles were present even in normal, healthy volunteers and this distorted the HRM curve shapes and prevented a final screening result to be announced on HRM curve shape only. The HLA-B\*58:01 allele had a higher percentage in gout/ADR samples, but was still present in normal samples, showing its widespread presence in the Malaysian population. A set of HLA-B alleles were seen accompanying the HLA-B\*58:01 allele here too and other unseen HLA-B alleles were also added to that list in normal samples.

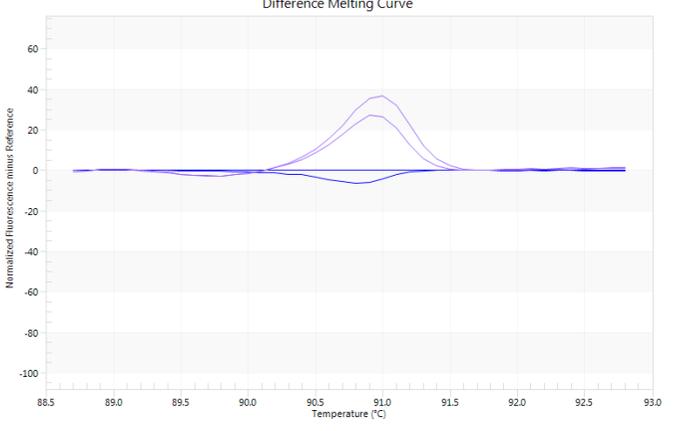
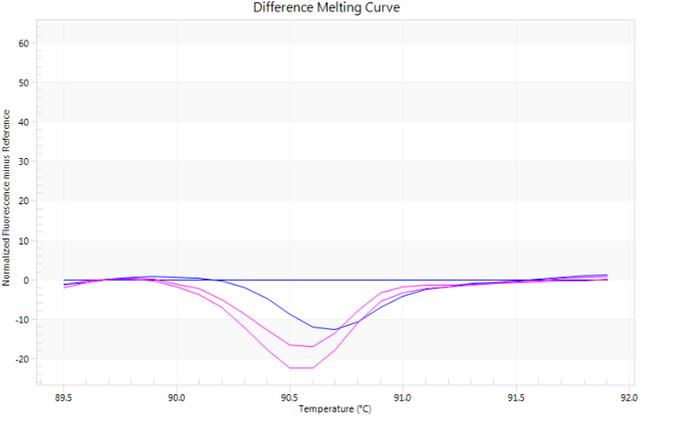
Table 4-17 Summary of HLA-B alleles present in the 14 sequenced normal samples.

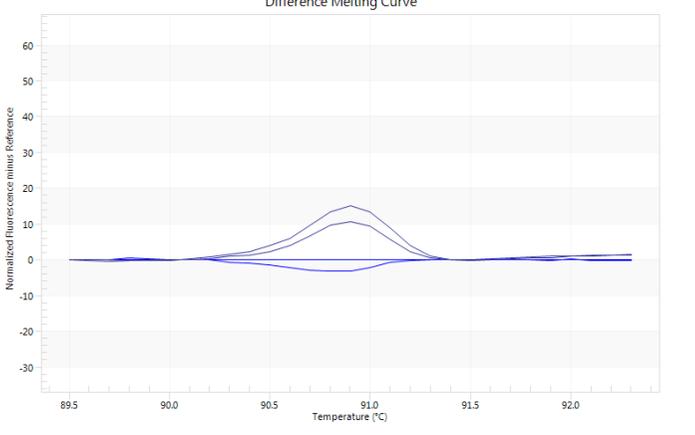
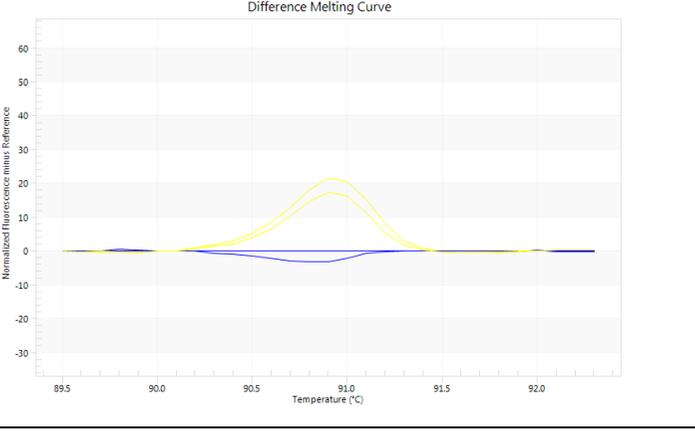
HLA subgroup	HLA alleles identified	Number of samples containing the allele	Total percentage of samples containing the specific allele
<b>HLA-B</b>	HLA-B*15	5	35.71
	HLA-B*35	5	35.71
	HLA-B*53	5	35.71
	HLA-B*58:01:01	4	28.57
	HLA-B*58	4	28.57
	HLA-B*15:21	4	28.57
	HLA-B*15:02	4	28.57
	HLA-B*35:01:01	4	28.57
	HLA-B*53:01:01	4	28.57
	HLA-B*44:03:01	2	14.29
	HLA-B*44	2	14.29
	HLA-B*15:02:01	1	7.14
	HLA-B*13	1	7.14
	HLA-B*13:02	1	7.14
	HLA-B*13:01:01:01	1	7.14
	HLA-B*35:03:01	1	7.14
HLA-B*35:02:02	1	7.14	

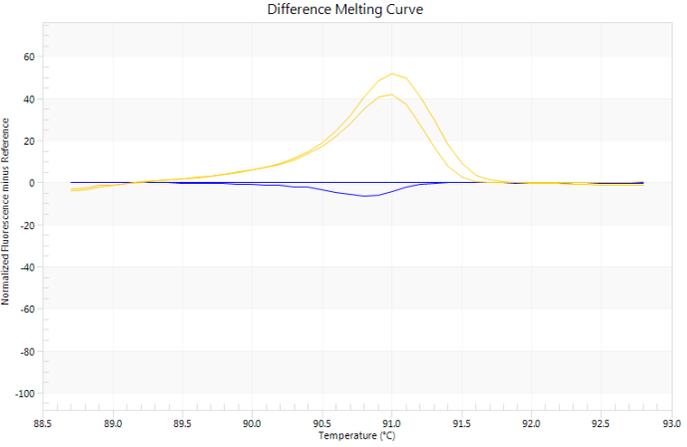
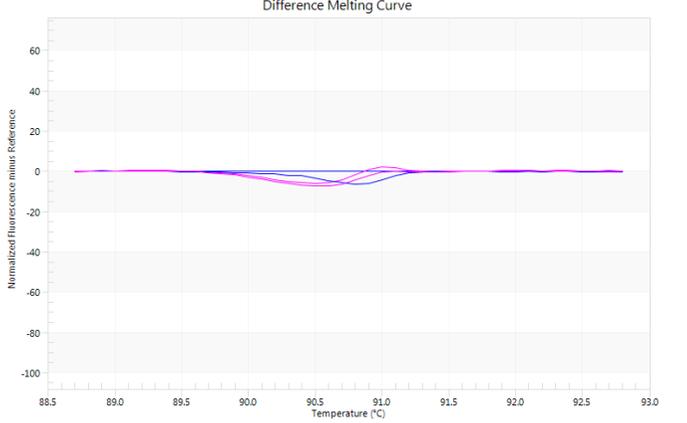
Table 4-18 14 random normal samples Sanger sequencing results for P3 primer and exon 3.

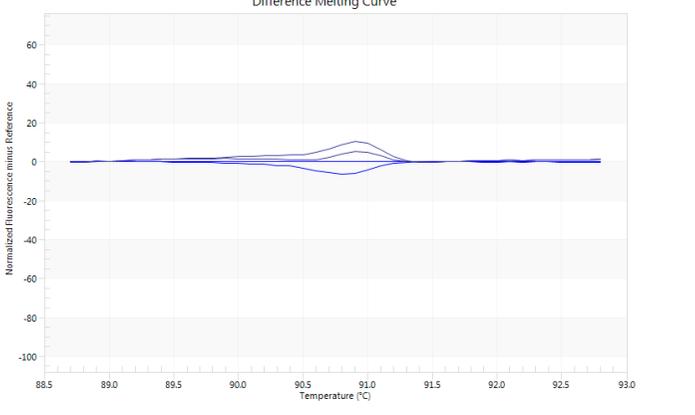
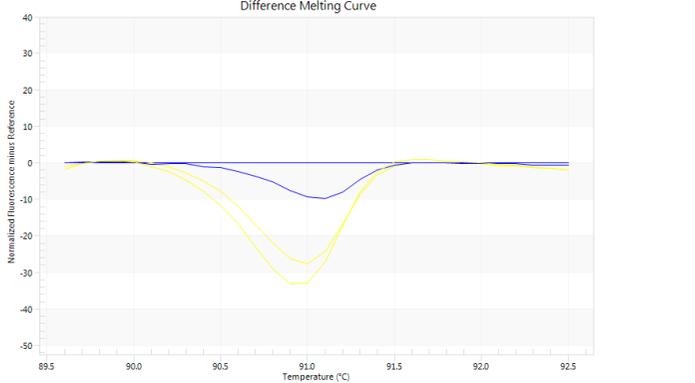
Sample code	2 HLA alleles identified by Sanger sequencing	Gender	Ethnic group	HRM screening result	HRM Difference melt curve
N54	HLA-B*35:01:01 HLA-B*58:01:01	Male	Chinese	Negative	
N33	HLA-B*35:01:01 HLA-B*58:01:01	Male	Malay	Negative	

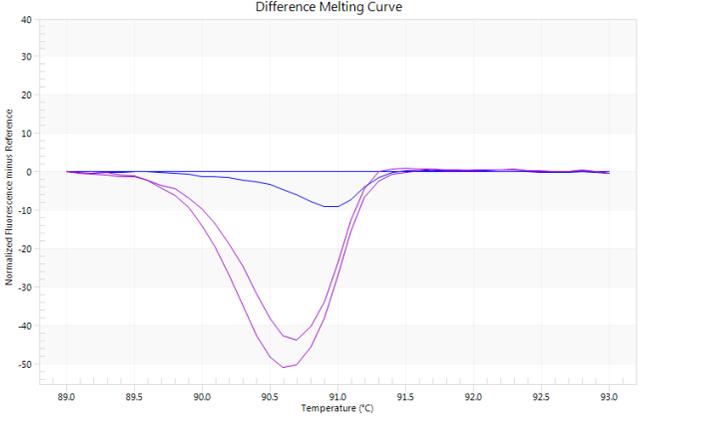
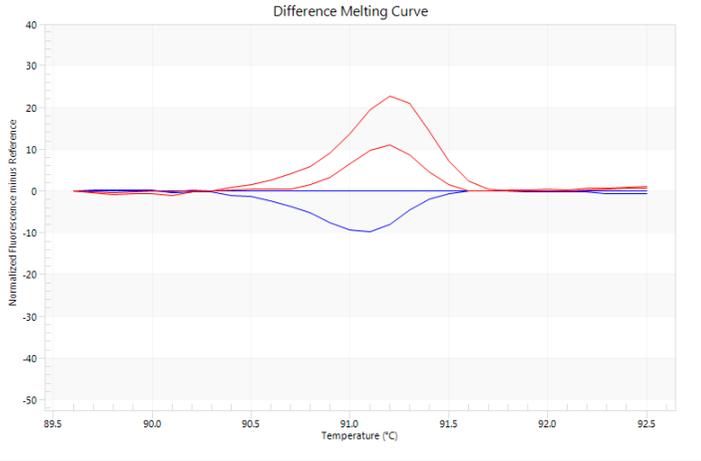
N38	HLA-C*07:02:01 HLA-C*15:05	Male	Malay	Negative	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus Reference</p> <p>Temperature (°C)</p> <p>The graph shows a positive peak centered at approximately 90.7°C, reaching a maximum value of about 30. A corresponding negative peak is visible at the same temperature, reaching a minimum of about -10. The x-axis ranges from 89.0 to 92.5°C, and the y-axis ranges from -10 to 80.</p>
N5	HLA-B*44:03:01 HLA-B*44	Female	Malay	Negative	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus Reference</p> <p>Temperature (°C)</p> <p>The graph shows a positive peak centered at approximately 91.0°C, reaching a maximum value of about 25. A corresponding negative peak is visible at the same temperature, reaching a minimum of about -5. The x-axis ranges from 89.5 to 92.5°C, and the y-axis ranges from -10 to 50.</p>

NS34	HLA-B*44:03:01 HLA-B*15	Female	Malay	Negative	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus Reference</p> <p>Temperature (°C)</p>
NS97	HLA-B*13:01:01:01 HLA-B*13:02	Male	Chinese	Possible positive	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus Reference</p> <p>Temperature (°C)</p>

NS11	HLA-B*15:02:01 HLA-B*15:21	Male	Malay	Possible positive	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus Reference</p> <p>Temperature (°C)</p>
NS6	HLA-B*35:03:01 HLA-B*53	Female	Malay	Possible positive	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus Reference</p> <p>Temperature (°C)</p>

NS37	HLA-B*35:01:01 HLA-B*58:01:01	Female	Malay	Negative	 <p>Difference Melting Curve</p> <p>Normalized Fluorescence minus Reference</p> <p>Temperature (°C)</p>
NS26	HLA-B*15:02 HLA-B*15:21	Female	Malay	Possible positive	 <p>Difference Melting Curve</p> <p>Normalized Fluorescence minus Reference</p> <p>Temperature (°C)</p>

NS28	HLA-B*15:02 HLA-B*15:21	Female	Malay	Possible positive	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus reference</p> <p>Temperature (°C)</p>
NS68	HLA-B*15:02 HLA-B*15:21	Female	Malay	Negative	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus reference</p> <p>Temperature (°C)</p>

NS296	HLA-C*15:02:02, HLA-C*15:10:02	Male	Malay	Negative	 <p>Difference Melting Curve</p> <p>Normalized Fluorescence minus Reference</p> <p>Temperature (°C)</p>
NS95	HLA-B*35:01:01 HLA-B*58:01:01	Male	Chinese	Possible positive	 <p>Difference Melting Curve</p> <p>Normalized Fluorescence minus Reference</p> <p>Temperature (°C)</p>

## 4.6 Evaluation of the HRM method

### 4.6.1 Sensitivity, specificity, positive predictive value and negative predictive value

The HRM screening method was evaluated here against the Sanger sequencing method to detect presence of the HLA-B\*58:01 allele. Sanger sequencing was done on 28 ADR/SCARs gout patients and yielded 9 HLA-B\*58:01 positive samples and 19 HLA-B\*58:01 negative samples. However, none of these 9 positive samples were identified as positive by the HRM curve analysis results. Sanger sequencing was also done on 14 healthy volunteers, yielding 4 HLA-B\*58:01 positive samples and 10 HLA-B\*58:01 negative samples. Out of these 10 negative results, 5 were also found negative by the HRM curve analysis results.

Values extracted from this study;

True positives are positive results in the HRM screening and in the sanger sequencing results. However, none match,  $a = 0$

False positives are positive results obtained by HRM screening but negative by sanger sequencing,  $b = 12$

False negative are samples with negative HRM screening results, but which are positive on sanger sequencing  $c = 12$

True negatives are samples with negative Sanger sequencing and HRM results,  $d = 16$

Therefore,

Sensitivity = 0

Specificity =  $0.57 = 57\%$

Positive predictive value = 0

Negative predictive value =  $0.57 = 57\%$

Hence, the HRM method has a sensitivity of 0, specificity of 57%, positive predictive value of 0 and negative predictive value of 57%.

## 4.6.2 Statistical Analysis

### 4.6.2.1 Fisher's exact test 1

Null hypothesis: There is no association in the HLA-B\*58:01 allele's distribution between the healthy volunteer cohort and the gout cohort.

Alternative hypothesis: There is an association in the HLA-B\*58:01 allele's distribution between the healthy volunteer cohort and the gout cohort.

The Fisher's exact test (2-sided), shows that there is no significant association, based on the p value of 2.496 ( $p > 0.05$ ) from the Pearson Chi-Square test. No Fisher's exact test value was obtained, owing to the larger sample size. Thus there is no significant association in the HLA-B\*58:01 allele's distribution between the healthy volunteer cohort and the gout cohort.

Table 4-19 Fisher's exact test for the association in the HLA-B\*58:01 allele's distribution between the healthy volunteer cohort and the gout cohort.

	Value	Df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)
Pearson Chi-Square	2.496	1	0.114	0.228
Likelihood Ratio	1.516	1	0.218	0.228
Fisher's Exact Test	-			0.228
N of valid cases	145			

#### 4.6.2.2 Fisher's exact test 2

Null hypothesis: There is no association between HLA-B\*58:01's presence and the different ethnic groups in the gout cohort.

Alternative hypothesis: There is an association between HLA-B\*58:01's presence and the different ethnic groups in the gout cohort.

The value of the Fisher's exact test (2-sided) statistic is 2.003, which results in a p value of 0.737( $p > 0.05$ ) as shown in Table 4.20 below. Thus there is no association between presence of HLA-B\*58:01 allele and the different ethnic groups in the gout cohort.

Table 4-20 Fisher's exact test for HLA-B\*58:01's association with the different ethnic groups in the gout cohort. Pearson Chi-squared test was also carried out in parallel.

	Value	Df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)
Pearson Chi-Square	0.268	3	0.966	1.000
Likelihood Ratio	0.451	3	0.930	1.000
Fisher's Exact Test	2.003			0.737
N of valid cases	145			

### 4.6.2.3 Fisher's exact test 3

Null hypothesis: There is no association between HLA-B\*58:01's presence and the different ethnic groups in the healthy volunteer cohort

Alternative hypothesis: There is an association between HLA-B\*58:01's presence and the different ethnic groups in the healthy volunteer cohort

The value of the Fisher's exact test (2-sided) statistic is 12.389, which results in a p value of 0.181 ( $p > 0.05$ ) as shown in Table 4.21 below. Thus there is no association between presence of HLA-B\*58:01 allele and the different ethnic groups in the volunteer cohort too.

Table 4-21 Fisher's exact test for HLA-B\*58:01's association with the different ethnic groups in the healthy volunteer cohort. Pearson Chi-squared test was also carried out in parallel.

	Value	Df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)
Pearson Chi-Square	7.788	6	0.254	0.181
Likelihood Ratio	6.858	6	0.334	0.128
Fisher's Exact Test	12.389			0.181
N of valid cases	145			

#### 4.6.2.4 One Way ANOVA

Null hypothesis: There is no link between HLA-B\*58:01's presence and the different types of allopurinol-induced hypersensitivity reactions seen.

Alternative hypothesis: There is a link between HLA-B\*58:01's presence and the different types of allopurinol-induced hypersensitivity reactions seen.

Before doing one way ANOVA, the homogeneity of variance of the data needs to be confirmed, as shown in Table 4.22 below. The p value is 0.491, greater than 0.05, thus data met the assumption of homogeneity of variance. Thus one way ANOVA can be conducted. Here, all allergies were grouped under hypersensitivity observed and HLA-B\*58:01 presence and absence was labelled as positive and negative. All of the data was made uniform for all allopurinol allergies and transformed to numerical values and the test was conducted. One way ANOVA showed  $p=0.905$  (as shown in Table 4.23 below.), hence showing no significant difference ( $p>0.05$ ) between different types of allergies and the presence or absence of the HLA-B\*58:01 allele. Therefore allele here is not strongly linked to diff types of allergic reactions.

Table 4-22 Test of Homogeneity of Variances for all the hypersensitivity observed.

Levene Statistic	df1	df2	Sig.
.489	1	26	0.491

Table 4-23 One Way ANOVA carried out for HLA-B\*58:01's presence in the different types of allopurinol-induced hypersensitivity reactions seen.

	Sum of Squares	Df	Mean Square	F	Sig.
<b>Between Groups</b>	0.201	1	0.201	0.014	0.905
<b>Within Groups</b>	361.906	26	13.919		
<b>Total</b>	362.107	27			

## 4.7 NGS Result compilation

### 4.7.1 Overall NGS Run results

NGS was done on the HLA-B region on the MiSeq system along with a team of experienced people for proper guidance in UMBI's laboratory and after the different layers of complicated analysis with the help of an experienced bioinformatician in UMBI the results were tabulated. This section will be divided into three parts, with the first one dealing with the sequencing quality control done by using the Sequence Analysis Viewer (SAV) from Illumina and the FastQC analysis reports for each sample. Once this step is done, downstream analysis will be performed, first with only the 6 gout samples and finally with of the 96 samples together.

### 4.7.2 Sequencing Quality Control

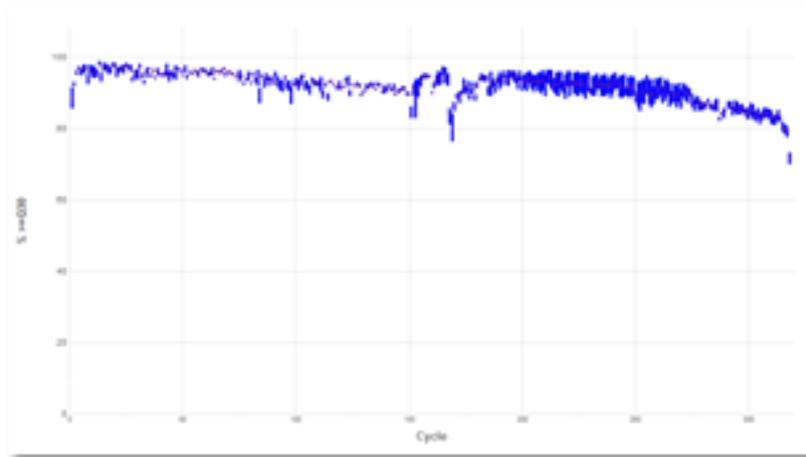
The first set of crucial results for the NGS run was the sequencing performance on the MiSeq system, that is the data quality and total data output and yield. The Sequence Analysis Viewer (SAV) was used to monitor the sequencing during the NGS run and to check the QC after the run on BaseSpace. The first result generated was the Run and Lane metrics table shown below, with the yield, error rate and % Q30. The Qscore or Q30 value was used as a quality control and shows if 99.9% of the bases are correct for the NGS done. Most Illumina runs will generate >70-80% Q30 data and this NGS run had a value of 91.79%, which was considered a very good Q30 value, along with very good quality of data generated (shown in Table 4.24). Graph A and B in Figure 4.7 both showed a good clustering of data without any sudden drop or erratic data. Graph C showed the optimal density range with the density box plots close together and a good cluster %PF value of 83.41%  $\pm$  5.17. Appendix H shows the good Fast Q scores of 38-40 for all samples and high percentages mapped to the human genome. Trimming was done with parameters adjusted to a score of 20, showing that 75% of the reads were left for alignment and analysis. Mapping was done to the full Human genome. A Kaviar p value less than 0.05 was obtained, along with low error rate of 0.54% shown in Table 4.30, showing accuracy of results. The few non-HLA-B

hits identified in a few samples were seen to be mapped to 89.61% of the human genome, with a still high percentage enough to show no deviation.

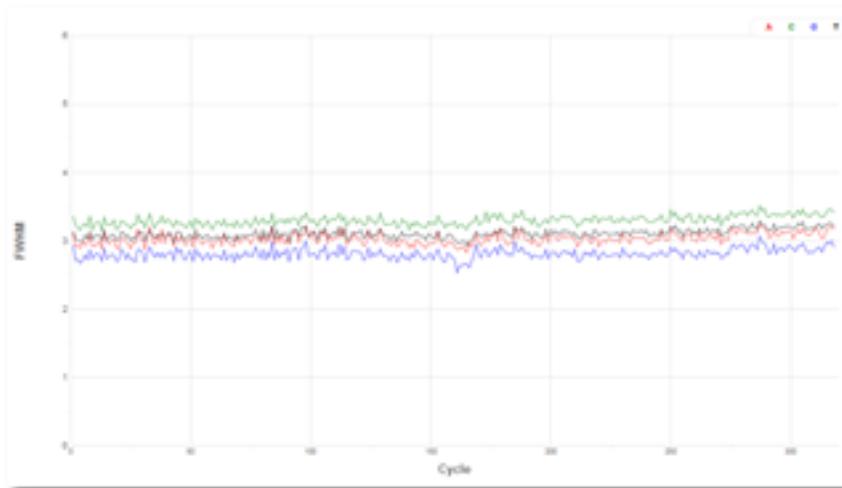
Table 4-24 Per Read metrics extracted from BaseSpace, showing the yield, error rate and %≥Q30.

	Cycles	Yield (Mbp)	Projected yield	Aligned (%)	Error rate (%)	Intensity cycle 1	% ≥ Q30
<b>Read 1</b>	151	136.56	136.56	1.37	0.50	131	93.91
<b>Read 2 (I)</b>	8	6.37	6.37	0.00	0.00	532	92.39
<b>Read 3 (I)</b>	8	6.37	6.37	0.00	0.00	554	93.48
<b>Read 4</b>	151	136.56	136.56	1.36	0.57	114	89.57
<b>Non-Index Reads total</b>	302	273.12	273.12	1.37	0.54	122	91.74
<b>Totals</b>	318	285.87	285.87	1.37	0.54	333	91.79

**A- Data by Cycle against %  $\geq$ Q30**



**B- Data by Cycle against FWHM**



**C- Data by Lane against Density**

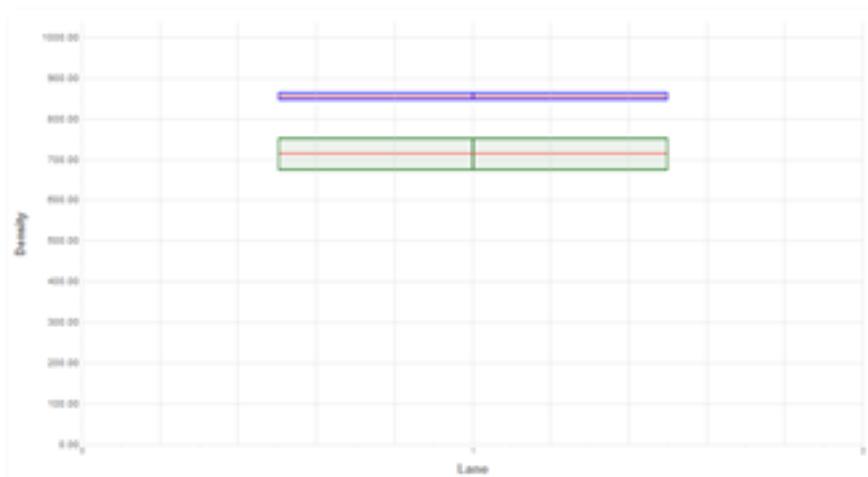


Figure 4-7 Compilation of 3 graphs for QC check, where graphs A and B showed Data by cycle against %  $\geq$ Q30 and FWHM respectively, followed by graph C showing Data by Lane against density.

### 4.7.3 Standard method of NGS analysis for all SNPs

Each sample's result were detailed with their identified SNPs and corresponding HLA allele called. The SNP number was first entered in dbSNP in NCBI in order to get a specific DNA sequence (FASTA format), along with the different base change possible. The obtained FASTA sequence was copied into the Nucleotide Similarity Search from NCBI and 2 databases were selected from IMGT's dropdown list, namely IMGT/HLA (cds) and IMGT/HLA (genomic). All the possible combinations were entered and the results were compiled with the HLA alleles with the highest presence in each sample, hence explaining the numerous HLA hits shown in the following Tables. Here, the hit B\*48, for example, indicates a wide range of B\*48 alleles (B\*48:01- B\*48:98) identified with one hit each, hence one allele cannot be called specifically. One complete and detailed example of SNPs and HLA allele hit analysis was shown in Table 4.25 on the following page. All the 96 NGS samples were analysed in the exact same way for all following NGS sections.

Table 4.25 shows the list of SNPs identified first, followed by the target area identified from NGS done, that is, the HLA-B region. In the first two lanes, even in the presence of 2 different base changes, no HLA alleles or other hits were identified, showing that is a novel SNP found and is classified as such in the all the NGS results. Other rows from Table 4.25, showed a minimum one 1 base change possible to a maximum of 4 base change possible and all the hits are combined to see which HLA alleles are most prominent. One base change yields a total of 100 HLA allele hits, with the highest similarity from the NCBI and IMGT/HLA database. One SNP identified can have a maximum of 4 different possible base change, which in turn yielded a total of 400 possible HLA allele hits. This shows that the NGS results obtained were aligned to a large number of HLA alleles in order to see the most prominent alleles. In some cases, the different base changes for one SNP identified numerous, different HLA alleles, while in other cases, all base changes identified only one HLA allele. The latter was seen in the second half of the Table 4.25, specially for HLA-B\*56 hits.

Table 4-25 Full details of how SNPs and HLA allele hits were analysed for sample HLAG12.

SNPs Identified by NGS	Target area from NGS	Variant from IMGT/HLA BLAST				
		Final HLA allele call	1st base + hits	2nd base + hits	3rd base + hits	4th base + hits
rs2395474	HLA-B	no hits - novel SNP	A- no hits	G- no hits	-	-
rs7762909	HLA-B	no hits - novel SNP	A- no hits	G-no hits	-	-
rs2308655	HLA-B	B*58:01/B*57:01	C-B*58:01	G-B*57:01	-	-
rs3926873	HLA-B	no hits - novel SNP	A- no hits	C- no hits	T-no hits	-
rs2442717	HLA-B	B*40/ B*58:01	A- B*58:01	C- B*58:01	G- B*40	-
rs2596497	HLA-B	B*58:01/B*40	C-58:01	T -B*40	-	-
rs17193040	HLA-B	no hits - novel SNP	G- no hits	T- no hits	-	-
rs1130992	HLA-B	B*48/B*48:01/39/44/58	A- B*39	C- B*48 (48:01)	G- B*48; B*44 (B*58 once)	-
rs1050823	HLA-B	B*58 (58:66-58:97)/39	C-B*39	T - B*58:97-58:52	-	-
rs709052	HLA-B	B*58/59/57/78/67	A- B*58/B*59	C- B*57:03/B*57	G- B*78/B*67/B*59/B *58	T-B*58:50- 58:97
rs1050747	HLA-B	E*01:01:01/E*01:03 , C*17, B*40	C- E*01:01:01/E*01:03 ; C*17	T- B*40	-	-
rs375356947	HLA-B	C*17:03-17:41	only one possibility	-	-	-
rs3180380	HLA-B	B*67/B*78/B*81/B*82	B*58:01 once	-	-	-

rs576010607	HLA-B	B*56/B*67	only one possibility	-	-	-
rs540530530	HLA-B	B*56/B*67/B*78/B*81	only one possibility	-	-	-
rs3190923	HLA-B	B*56	C- B*56 numerous	G- B*56 numerous	T- B*56 numerous	-
rs1050388	HLA-B	B*56	A- B*56 numerous	G- B*56 numerous	-	-
rs1131215	HLA-B	B*56	A- B*56 numerous	C- B*56 numerous	G- B*56 numerous	T- B*56 numerous
rs750527298	HLA-B	B*56	C- - B*56 numerous	C-- B*56 numerous	-	-
rs749557905	HLA-B	B*56	GA- - B*56 numerous	TT- B*56 numerous	-	-
rs1131204	HLA-B	B*56	A- B*56 numerous	C- B*56 numerous	G- B*56 numerous	-
rs1071816	HLA-B	B*56	A- B*56 numerous	C- B*56 numerous	G- B*56 numerous	T- B*56 numerous
rs1050564	HLA-B	B*56; B*59	only one possibility	-	-	-
rs1050517	HLA-B	B*81/B*67	B*58:01 once	-	-	-
rs147324178	HLA-B	A*24	only one possibility	-	-	-

## 4.7.4 Analysis of NGS done on 6 gout samples

### 4.7.4.1 Detailed NGS analysis for the 6 gout samples

The 6 gout samples sent for NGS in 2017, were the only ones available with the more severe ADRs/SCARs, hence went in the NGS run in 2017. The NGS results for the 6 gout samples show similar results to the sanger sequencing results, where the SNPs identified on samples showed numerous HLA alleles hits and not just one specific allele. This is due to the broad target of the HLA-B primer used, which amplifies the whole 4000bp of this region. The 6 tables compiled on the following pages show the detailed SNPs list along with their respective final HLA hits for the 6 gout samples. Samples HLAG12, HLAG19, HLAG26 and HLAG28 had more HLA-related SNPs and HLA allele hits compared to the other three samples. HLAG12 showed a wide range of HLA hits in Table 4.26, ranging from HLA-A, HLA-B, HLA-C, HLA-E and even MICB gene. The HLA-B\*56 had the highest hit number from 10 different SNPs. HLA-B\*58 (2 hits) and HLA-B\*58:01 (3 hits) were seen, along with single hits of the following alleles; HLA-B\*57:01, HLA-B\*48/67/78/81 and HLA-A\*24. Only one HLA-A\*24 and HLA-C\*17 hit were seen in the whole table. The HLA-B\*58:01 allele was seen linked to the HLA-B\*57:01 and HLA-B\*40 alleles here and present once in 2 other SNPs. Labelling of a hit as HLA-B\*58 means that too many HLA-B\*58 alleles were found and the main allele only was written.

Sample HLAG19, in Table 4.27, showed only 3 hits, but identified common HLA-B alleles. HLA-B\*58:01, HLA-B\*57:01 and HLA-B\*40 were seen in 2 SNPs each. The HLA-B\*58:01 allele was seen linked to the HLA-B\*57:01 and HLA-B\*40 alleles here and this was also seen in sample HLAG12 previously. HLAG26 had numerous novel SNPs and only HLA-B\*15:01:01:02 was identified twice, as seen in Table 4.28. Some of the novel SNPs had other results such as intronic variants (LOC105377891 & LOC105377899) and transcription factors (JARID2). HLAG28 showed triple hits for the HLA-B\*58:01 allele and double hits for the HLA-B\*58 and HLA-B\*40 alleles in Table 4.29. The HLA-B\*58:01 allele was seen linked to the HLA-B\*57:01 and HLA-B\*40 alleles again, as previously seen in samples HLAG12 and HLAG19. Only HLA-

B\*48 and HLA-A\*24 were seen as single hits, in one SNP each. Other HLA alleles seen were all seen once and linked to other alleles, along with a total of 3 different novel SNPs.

The positive control and sample HALG29 both had only one hit for HLA-A\*31:37, along with several novel SNPs, as shown in Tables 4.30-4.31. Only samples HLAG12, HLAG19 and HLAG28 showed presence of the HLA-B\*58:01 allele amongst the 5 samples and the positive control sent for the NGS. Sample HLAG12 had the HLA-B\*58:01 allele detected by sanger sequencing and sample HLAG28 was a possible positive detected by the HRM method. Samples HLAG12 and HLAG28 showed presence of numerous HLA-B\*58 alleles, along with other common alleles such as HLA-B\*57:01, HLA-B\*48/56/67/81 and HLA-A\*24. However, these results were not completely decoded due to a high number of novel SNPs identified for each sample, as no match were found in the global NCBI dbSNP database or IMGT/HLA database online. More gout samples with SCARs need to be sent for NGS in the future for better and more in-depth analysis.

Table 4-26 SNPs identified from NGS results accompanied by HLA allele calling for sample HLAG12.

SNPs identified for HLAG12	HLA allele calling
rs2395474	no hits – novel SNPs
rs7762909	no hits – novel SNPs
rs2308655	B*58:01/B*57:01
rs3926873	no hits- novel SNPs
rs2442717	B*40/58:01
rs2596497	B*58:01/B*40
rs17193040	no hits – novel SNPs
rs1130992	B*48/B*48:01/39/44/58
rs1050823	B*58 (58:66-58:97)/39
rs709052	B*58/59/57/78/67
rs1050747	E*01:01:01, E*01:03 , C*17, B*40
rs375356947	C*17 (C*17:03-17:41)
rs3180380	B*67, B*78, B*81, B*82, B*58:01 once
rs576010607	B*56, B*67
rs540530530	B*56, B*67, B*78, B*81
rs3190923	B*56

rs1050388	B*56
rs1131215	B*56
rs750527298	B*56
rs749557905	B*56
rs1131204	B*56
rs1071816	B*56
rs1050564	B*56, B*59
rs1050517	B*81, B*67, B*58:01 once
rs147324178	A*24

Table 4-27 SNPs identified from NGS results accompanied by HLA allele calling for sample HLAG19.

SNPs identified for HLAG19	HLA allele calling
rs41561016	58:01/57:01
rs41543314	B*40/B*48/ 57:01
rs2596497	B*58:01/B*40

Table 4-28 SNPs identified from NGS results accompanied by HLA allele calling for sample HLAG26.

SNPs identified for HLAG26	HLA allele calling
rs9266193	B*15:01:01:02
rs9266194	B*15:01:01:02
rs9464782	JARID2
rs9464782	no hits – novel SNPs
rs9632507	no hits – novel SNPs
rs9632508	no hits – novel SNPs
rs60367535	LOC105377891
rs1906960	LOC105377899

Table 4-29 SNPs identified from NGS results accompanied by HLA allele calling for sample HLAG28.

SNPs identified for HLAG28	HLA allele calling
rs2395474	no hits – novel SNPs
rs2308655	B*57:01/B*58:01
rs3926873	no hits – novel SNPs
rs2442717	B*58:01/B*40
rs2596497	B*58:01/B*40
rs17193040	no hits – novel SNPs
rs1130992	B*48
rs1050823	B*58 numerous (B*58:66-58:97)
rs709052	B*58 numerous (B*58:50-58:97)
rs1050564	B*56; B*59
rs1050517	B*67, B*81, B*48, B*51, B*58
rs147324178	A*24

Table 4-30 SNPs identified from NGS results accompanied by HLA allele calling for sample HLAG29.

SNPs identified for HLAG29	HLA allele calling
rs116039287	no hits – novel SNPs
rs7774984	no hits – novel SNPs
rs74948045	no hits – novel SNPs
rs6911090	LOC105377975 - intron variant
rs1856859	A*31:37, CCDC170- intron variant
rs10644644	no hits – novel SNPs
rs67830808	no hits – novel SNPs
rs9364495	no hits – novel SNPs

Table 4-31 SNPs identified from NGS results accompanied by HLA allele calling for the positive control.

SNPs identified for the positive control	HLA allele calling
rs116039287	no hits – novel SNPs
rs7774984	no hits – novel SNPs
rs74948045	no hits – novel SNPs
rs6911090	LOC105377975 - intron variant
rs1856859	A*31:37, CCDC170- intron variant
rs10644644	IPCEF1- intron variant
rs67830808	IPCEF1- intron variant
rs9364495	SYTL3 (94120)

#### 4.7.4.2 Comparing Sanger sequencing results to NGS results for 6 gout samples

The Sanger sequencing results for those 6 samples were compared to the NGS results in order to see if any HLA allele hits were similar in Table 4.36. Only sample HLAG12 matched its Sanger sequencing and NGS results, with numerous HLA-B\*58 alleles and the HLA-B\*58:01 allele in common. Interestingly, this similarity of NGS and Sanger sequencing results happened in a sample with vasculitis rash only, and no severe SCARs like SJS/TEN/DRESS. The positive control (PC5801) and sample HLAG29 showed similar NGS results, with only HLA-A\*31:37 identified in both and their Sanger sequencing results both showed presence of HLA-B\*58:01. These two samples had SJS and DRESS syndrome respectively, considered amongst the most severe type of SCARs. Here, Sanger sequencing identified the desired allele and not NGS, which was intriguing when based on the level of complexity of both of these sequencing methods.

Sanger sequencing identified the presence of HLA-B\*58:01 in 4 samples (PC58:01, HLAG12, HLAG26, HLAG29), while NGS identified this allele in 3 samples (HLAG12, HLAG19, HLAG28). When observing both methods of sequencing across all those 6 samples, the HLA-B\*58:01 allele has been identified in at least one method for all samples. When taking into consideration the ADRs and SCARs first, all the various types seen in Table 4.36 were seen to be associated with the HLA-B\*58:01 allele at least once. Samples HLAG19 and HLAG28 were

the only samples without the HLA-B\*58:01 allele identified in Sanger sequencing, but this was challenged by the presence of SNPs carrying the HLA-B\*58:01 allele in the NGS method. Furthermore, both samples had itchiness as ADRs and they both had the HLA-B\*57:01, HLA-B\*40 and HLA-B\*48 alleles in common. Sanger sequencing previously showed presence of the HLA-B\*35:01:01 and HLA-B\*58:02 alleles always accompanying the HLA-B\*58:01 allele, but this wasn't seen in the NGS results for the gout samples as shown in Table 4.36. The HLA-B\*44 allele was found by Sanger sequencing in the HLAG28 sample and by NGS in the HLAG12 allele.

The two samples with the highest number of total HLA hits for both sequencing methods were HLAG12 and HLAG28. Both samples had milder forms of ADRs, with vasculitis rash and itchiness, respectively. These two samples differed greatly in their sanger sequencing results, but had numerous alleles common in their NGS results, namely, HLA-B\*58:01, HLA-B\*57:01, HLA-B\*40, HLA-B\*48, HLA-B\*67, HLA-B\*81, HLA-B\*56, HLA-B\*59 and HLA-A\*24. The HLA-B\*58:01 and HLA-B\*57:01 alleles were both linked to those samples and have both been identified previously as genetic markers in the pharmacogenomics field. This shows the possible involvement of numerous alleles, even in milder ADR cases. Some samples contained too many novel SNPs in the NGS results to be matched properly to the Sanger sequencing results. Further research must be carried out, with a greater number of gout samples, in order to validate the aforementioned results/claims in a bigger sample size.

Table 4-32 Comparison of the Sanger sequencing and NGS results for the 6 gout samples. Common alleles observed and the SCARs identified are also included for analysis purposes.

Sample code	Sanger sequencing results	HLA hits in NGS results	Common alleles	HLA ADRs and SCARs observed
PC5801	B*58:01	A*31:37	None	SJS
HLAG12	B*35:01:01, B*58:01:01, B*58, B*58:01, B*58:02	B*58:01, B*57:01, B*40, B*48:01, B*39, B*44, B*58, B*57, B*78, B*67, B*81, B*82, B*56, B*59, C*17, A*24	B*58, B*58:01	Vasculitis rash
HLAG19	C*07:04, C*07	B*58:01, B*57:01, B*40, B*48	None	Mild allergy-itchiness
HLAG26	B*35:01:01, B*58:01:01, B*58, B*58:01, B*58:02	B*15:01:01:02	None	Exfoliative dermatitis
HLAG28	B*44, B*44:03:01	B*58:01, B*57:01, B*40, B*48, B*58, B*56; B*59, B*67, B*81, B*51, A*24	None	Mild allergy-itchiness
HLAG29	B*35:01:01, B*58:01:01, B*58, B*58:01, B*58:02	A*31:37	None	DRESS syndrome

## 4.7.5 Result compilation and analysis the 96 NGS samples

### 4.7.5.1 Comparison of HLA alleles' percentage throughout research

The common alleles in all the 3 cohorts were gathered, namely the NGS cohort, the gout-ADR cohort and the healthy volunteers sent for Sanger sequencing. The sample size for all those cohorts were very different, as shown in Table 4.37 below. However, compiling all their HLA alleles should give us a better idea of their percentages and their trend in the Malaysian population. Four common HLA-B alleles were found. The HLA-B\*58 and HLA-B\*58:01:01 alleles were found at a high percentage, with around a third of all cohorts containing those alleles. This proves that those alleles are present in a third of the Malaysian population, irrespective of disease (gout) or healthy individuals. HLA-B\*15 was found more in the healthy volunteers, compared to the NGS and gout-ADR cohorts. Interestingly the HLA-B\*44 allele was also found in high percentages across the cohorts, and in the same pattern as the HLA-B\*15 allele. Compiling the HLA percentages across the 3 cohorts showed that the main allele being investigated, that is the HLA-B\*58:01 allele, was significantly present, in higher percentages compared to initial predictions.

Table 4-33 Compilation of the 3 cohorts' common HLA alleles and their percentage.

HLA alleles discovered in NGS	Percentage (%) of samples with allele (n=6)	Percentage (%) in gout-ADR samples (n=28)	Percentage (%) in healthy volunteers (n=14)
B*58	50.0	32.10	28.57
B*58:01:01	50.0	32.10	28.57
B*15	16.70	17.90	35.71
B*44	16.70	10.70	14.29

# CHAPTER 5: DISCUSSION

## 5.1 Building the backbone of this study

### 5.1.1 Demographic data

Demographic data collected for both cohorts (Table 4.1) show a difference in the distribution of males and females, with more male gout patients (92.4%) than male healthy volunteers (46.9%). The opposite was seen for females, with 7.6% in the gout cohort and 53.1% in the healthy volunteer cohort. This difference arose mainly due to limited number of healthy volunteers in UMBI and limited time available to complete the study due to delayed sample collection. Thus, ethnicity was given priority over gender for sample matching as pharmacogenomics studies proved that it's the main factor linked to susceptible HLA-B allele presence and frequency (Nebert & Menon, 2001). The three specific ethnic groups had similar distributions in the gout and healthy volunteer cohort. The gout cohort had 79.3% Malays, 18.6% Chinese, 1.3% Indian and 0.7% of other ethnic groups. The age mean calculated for both cohorts were only four years apart (gout: 55 years & volunteer: 51 years), thus showing an acceptable balance. Ageing is known as the major risk factor in developing gout for both genders. Therefore, this attributes to multifactorial causality which then leads to hyperuricemia hence, predispose to crystal deposition (Roddy et al., 2007; De Leonardis et al., 2007).

Gout has always been known as a disease of elderly men, with an average onset age of 60 years or more (Zhang et al., 2016). Hence, the average age of 55 years old in the gout cohort adheres to this. However, recent studies in Taiwan found a trend of earlier gout onset in men, where they fall in between 30 to 39 years old (Yu & Luo, 2003; Kuo et al., 2015; Chen & Shen, 2007). This was also observed in this research, with 19.3% of the gout cohort aged less than 40 years old, showing proof of early onset of gout. 21 out of the 28 patients presenting with early onset gout were aged 30 to 39 years old, as aforementioned. This earlier peak of incidence is mostly attributed to overweight patients, with a family history of gout as proven by researchers in Taiwan (Chen et al., 2003). All the female gout patients were above 55 years old, except for one (37 years old), showing a majority of postmenopausal women. Worldwide studies have shown that gout affects predominantly male patients (Roddy & Doherty, 2010). However, the small percentage of female gout

patients are usually postmenopausal women, due to a rise in urate levels after a certain age. High level of oestrogen in women is known to increase the uric acid excretion, hence protecting them from gout (Anton et al., 1986). An increase in gout risk of around 26% was recorded with menopause and a proven decrease in gout risk with postmenopausal hormone use (Hak et al., 2010). More Malay patients were recorded in Malaysian gout studies, due to genetic factors in uric acid metabolism in Melanesian people and the shared genetic profiles with the Maori and Pacific Islanders known for their higher gout prevalence and incidence (Klemp et al., 1997; Robinson et al., 2013). The highest worldwide gout prevalence has been reported in Oceanian countries, particularly within ethnic groups such as the Maori people, where estimates can go more than 10% (Klemp et al., 1997). Since an official, specific number is not available for Malaysia yet, the data for Singapore is used as a reference due to the high similarity in the two populations. The Singapore Chinese Health Study recruited 52,322 people age 45 to 74 years and identified 2,117 (4.1%) people with gout in the period between 1999 and 2004 (Teng et al., 2012). Chinese gout patients are amongst the second highest most prominent group, thus showing that Malays will have an even greater prevalence than 4.1% once proper nation-wide studies will be completed. Moreover, considering that the Singaporean study aforementioned stopped in 2004, the gout prevalence nowadays would definitely have increased more.

The demographic distribution seen for this gout cohort is supported by numerous studies done in Malaysia, showing the absence of any bias. Among the three major ethnic groups in Malaysia, the incidence of gout is higher in Malay men. However, ageing is a major risk factor in gout development for both genders, where the SUA level increases above 40 years old. A rheumatology clinic in Ipoh had a majority of Malay men having gout, followed by Chinese and Indians with a mean age of 53 years and an age range of 26 to 80 years old. Higher prevalence of gout was seen in men, compared to women from all age groups with ratio ranges from 2:1 to 4:1, respectively (Sulaiman et al., 2019). A general hospital in Sarawak had 88.9% males, 11.1% females, a higher gout prevalence in Malays (39.7%), the Iban (31.7%), Chinese (17.5%) and Bidayuh (11.1%). The native Malaysian formed more than 50% of this study, with 28.6% of patients having a positive family history of gout. A mean age of 60 years old was recorded, along with all the female patients being postmenopausal (Teh et al., 2014).

Two hospitals in Seremban, conducted a male only study and showed a mean age of 52 years old, along with a total of 79.5% Malay, 15.1% Chinese and 5.5% Indian patients (Das Gupta et al., 2018). The authors also investigated the familial risk of gout, specifically polymorphisms in SLC2A9/GLUT9 rs1172228. At least 28 genetic loci have been established by GWAS (Kottgen et al., 2013) and SNP analyses at these loci have identified polymorphic alleles that may potentially contribute to the familial risk of gout (Merriman et al., 2014). Authors concluded that the CC genotype in gout had significant associations with higher levels of SUA, renal calculi and a positive family history of gout in Malaysian men (Das Gupta et al., 2018). Family history of gout was unfortunately rarely mentioned in the medical records of both hospitals and this should be improved. Mohd and colleagues showed the same pattern in a Seremban hospital, with 89% males, 11% females, 55.5% patients with gout onset at 30 to 40 years old, a majority of Malays (72%), followed by Indians (20%) and Chinese (8%) (Mohd et al., 2006; Mohd et al., 2011). A Malay only study done in Kelantan, had 92% of men in their study, with an age mean of 44 years (17-71 years) (Wan Rohani et al., 2018).

### 5.1.2 Severe Cutaneous Adverse Drug Reactions (SCARs)

A total of 28 ADRs/SCARs cases were recorded in this study over 3 years, showing an incidence rate of 19.3% from the total of 145 gout samples. Specific SCARs recorded had an incidence rate of 4.14% (6 cases), with three cases of SJS, two DRESS cases and one AHS case. Hence, this study recorded a lower number of 2 SCARs cases per year as compared to six other Malaysian SCAR studies. Two Johor hospitals recorded 36.2 cases/year (Choon & Lai, 2012) and 35.1 cases/year (Ding et al., 2010), followed by 10.7 cases/year in the University Malaya Medical Centre (UMMC) in Selangor (Tan et al., 2017). The smallest number of yearly SCAR cases in literature were; 6 cases/year in a Sarawak hospital (Yap et al., 2008), 6.6 cases/year in another Selangor hospital (Tee & Ng, 2014) and 4.73 cases/year in a Penang hospital (Loo et al., 2018). All these aforementioned studies gathered data from a SCAR database in their hospital and had proper SCAR classification, thus making their SCAR cases reporting more accurate. Moreover, all their patients were directed to the dermatology department for treatment on appearance of SCARs, making it possible to gather all the data in one department. The lack of proper SCAR classification methods and absence of a link with a dermatology department caused the low SCAR case number in this study. 46.4% of ADR/SCAR cases were non-specifically labelled as mild allergy (rashes/itchiness), 10.7% as simple allopurinol allergy and another 21.6% was labelled with six different non-specific terms. This makes a total of 78.7% of ADR/SCAR cases which were non-specifically labelled and thus contributes to the underestimation of specific SCAR cases in this study.

Classification of SCARs/ADRs was not done according to the SCORETEN scale or any other fixed criteria or guidelines, as observed in both hospitals. Allergic reactions to allopurinol or febuxostat were labelled as what they were observed as and there were no standard to compare them to. On appearance of rashes or itchiness, allopurinol was stopped immediately and febuxostat was started. In rarer cases of febuxostat allergy, febuxostat was stopped immediately and treatment focused on other drugs. The specific onset time of allergic reactions was also not recorded by doctors. It is imperative to classify and distinguish all the specific clinical features for SJS, SJS/TEN overlap and TEN, to properly administer required treatment and care. SCARs can be distinguished by their primary

lesions, distribution on the body, mucosal involvement, systemic symptoms and their percentage of body surface area detached (shown in Table 2.3 in Literature review chapter). This table, along with the SCORETEN scale allows proper classification of a patient's allergic reaction, their possible mortality rate and hence allows for prevention with early detection (Bastuji-Garin et al., 2000). The lack of proper SCAR classification in both hospitals is a limitation in this study, where possible SCARs cases were merely labelled as allopurinol allergy/rashes/itchiness and other terms. Thus, the number of SCAR cases in this study was underestimated, along with all its other linked results.

A wide range of ADRs/SCARs were observed, including vasculitis rash, mild allergy (itchiness/rash), maculopapular rash with skin eruptions, exfoliative dermatitis, DRESS syndrome, SJS, allopurinol hypersensitivity syndrome (AHS), allergy with anaphylaxis, allergy with transaminitis, severe allergy with eye angioedema and allopurinol allergy. A total of 12 different terms were used in both hospitals in order to describe the different types of allergies to allopurinol by a group of around 10 to 15 doctors. A greater number of ADRs/SCARs patients were obtained from HUKM and they were more scattered around all these aforementioned terms (8 total terms used as seen in Table 4.3 in Results chapter). Hospital Putrajaya had slightly less variations in the classification of ADRs/SCARs, mostly due to their smaller number of ADRs/SCARs patients and interconnected computerized system. Sample HLAP56 with maculopapular rash and skin eruptions is considered a nonspecific term in clinical diagnosis and might have been a SJS, TEN or DRESS patient. Similarly, patients HLAG26 and HLAP51 with exfoliative dermatitis, may fall within the SJS/TEN/DRESS categories. Sample HLAP64 with severe allergy and eye angioedema again shows enough symptoms to be classified as SJS/TEN/DRESS. Both samples HLAG68 and HLAP52 experienced allopurinol allergy but with symptoms such as anaphylaxis and transaminitis which suggest a more severe allergic case. The use of 'allopurinol allergy' only in doctors' notes further proves the lack of guidelines for proper ADRs/SCARs classification and undermines their effect in this study.

The most frequent types of ADRs/SCARs were mild allergy (rashes/itchiness) (46.4%), followed by SJS (10.7%) and exfoliative dermatitis/DRESS syndrome/allopurinol allergy (7.10%). In the total gout sample cohort of 145 samples, those percentages still remain significant, with 8.97% for mild allergies, followed by 2.07% for SJS and 1.38% for

DRESS/exfoliative dermatitis/allopurinol allergy. The percentage recorded for SJS cases (10.7%) in this study was lower than the other five Malaysian studies from literature. The highest SJS percentage was recorded in Hospital Tengku Ampuan Rahimah (HTAR) in Selangor, with 75.8% SJS cases on its total SCARs cases (Tee & Ng, 2014). A Penang hospital showed a high percentage of 55.0% of SJS cases (Loo et al., 2018), followed by 54.2% in a Sarawak hospital (Yap et al., 2008). Two Johor hospitals had SJS percentages of 24.3% (Choon & Lai, 2012) and 28.1% (Ding et al., 2010) respectively. The lower percentage in this study was due to the normal collection of gout samples in a hospital setting, while the other five studies gathered only SCARs patients over numerous years (5 to 10 years). This can also be attributed to referral bias and different prescribing habits in different hospitals around Malaysia. These aforementioned studies specifically collected only SCARs cases over numerous years and hence may be a reason for their higher percentages of SJS cases. Thus, 10.7% of SJS cases in this study can be used as a closer estimate of the true percentage of Malaysian SJS patients in a gout cohort. DRESS was responsible for 7.10% of ADRs/SCARs cases in this study and matched the percentages seen in four other Malaysian hospitals; 9.4% (Choon & Lai, 2012), 6.8% (Ding et al., 2010), 12.7% (Loo et al., 2018) and 15.1% (Tee & Ng, 2014). SJS and DRESS are in the top 3 ADRs/SCARs in this study and this was similarly found in all the aforementioned SCARs Malaysian studies.

A mean age of around 57 years old was calculated for the 28 ADRs/SCARs samples, with an age range of 32 to 81 years old. The mean age in this study was higher compared to three other Malaysian SCARs studies; 39.6 years (Choon & Lai, 2012), 39 years old (Ding et al., 2010), 42.8 years (Tee & Ng, 2014) and 45 years (Loo et al., 2018). Moreover, the age range was shifted to the right in this study, as two studies showed an age range of 20 and 59 years old for around 67.1% patients (Choon & Lai, 2012; Ding et al., 2010). The male to female ratio is 25:3 in this study, showing a very high male preponderance. Most of the Malaysian SCARs studies showed a male preponderance, but the ratio of males was eight time less. Four studies had male:female ratios of 1.14:1 (Choon & Lai, 2012), 1.19:1 (Ding et al., 2010), 1:1 (Voo et al., 2017), 1.36:1 (Tee & Ng, 2014) and only one study showed higher female preponderance (female:male ratio= 1.2:1.0)(Loo et al., 2018). The ethnicity ratio for Malay:Chinese:Indian:others is 25:2:0:1, with a clear majority being Malay (89.3%) and the minority being Chinese (7.1%) and of other ethnic group (3.6%). Two studies showed a

smaller Malay and a higher Chinese percentage; 64% Malay and 26% Chinese (Ding et al., 2010) and 49% Malays and 38% Chinese (Loo et al., 2018). Moreover, this study did not have any Indian ADRs/SCARs patients, whereas other studies showed a presence of 7.5% (Ding et al., 2010) and 11% Indian patients (Loo et al., 2018). The lack of Indian patients and lower number of Chinese patients in this study may be linked to the small sample size of 28 ADRs/SCARs patients and the lack of proper SCARs classification too. The higher percentage of Malay patients with ADRs/SCARs suggests a stronger link in this ethnic group for this study. However, the range of ADRs/SCARs for this group of Malay patients is wider compared to the other Malaysian studies, due to lack of proper SCARs classification.

Mild allergy was the most frequent form of ADRs, be it rashes or itchiness, and emphasizes on its importance in association studies. Maculopapular rashes/eruptions are known to be the most common SCAR in several Malaysian SCAR studies; 42.3% (Choon & Lai, 2012), 39.5% (Ding et al., 2010), 38.1% (Voo et al., 2017). However, only one patient was found to have maculopapular rash in this study. The high number of rashes and itchiness cases recorded in the two hospitals might fall in the maculopapular rash/eruption group, but this was not possible due to lack of details and classification of SCARs. Pharmacogenomics studies have been focusing on the severe SCARs and leaving the mild allergies out, which may have led to other HLA-B allele associations being missed or underestimated. The full range of different allergies observed to allopurinol must be investigated for possible HLA-B allele links. This will be shown in the following sections in the discussion chapter, where all the 28 ADRs/SCARs will be linked to their HLA-B alleles.

Interestingly, the presence of both allopurinol and febuxostat allergies was recorded in two patients, namely HLAP3 (generalised rash) and HLAP51 (exfoliative dermatitis). The onset time of both allergic reactions, along with their different presentations were not recorded in their hospital records. Only little information was available on their allergic reactions to both drugs and this again undermines their severity and link to SCARs. Allopurinol is a purine analog and non-specific competitive inhibitor of XO, whereas febuxostat is a non-purine inhibitor of XO. These two drugs are structurally distinct and were anticipated to not share the same side effects (Chohan, 2011). It is likely that this cross-reaction is due to a non-immunological action, linked to the common pharmacological action of xanthine oxidase inhibition (Lien & Logan, 2017).

Other factors complicating the classification and identification of ADRs/SCARs is the presence of numerous comorbidities and their respective medications taken in conjunction with their gout therapy. Without proper ADRs/SCARs classification system in Malaysian hospitals, the association of HLA-B alleles to gout will be erratic at best through studies. A proper system needs to be established for SCARs identification and recording across all Malaysian hospitals, in order to cement the HLA-B alleles' role in allopurinol-induced ADRs/SCARs.

### 5.1.3 Optimization of DNA extraction

All the blood samples collected in HUKM were processed in UMBI, the laboratory next to the hospital. However, the samples collected from HP had to be taken back to UNMC for DNA extraction and required a few hours of travel. Blood samples were transported in a cooler box filled with ice and no change was observed in the samples' appearance. DNA concentration and quality were also not affected as shown by the results in Appendix B. Concentration had a range of 1000-3000 ng/ $\mu$ L for all samples, showing that the method was optimized. A pure extracted DNA sample is expected to have OD ratio of 1.7-2.0, as shown in literature (Jorgez et al., 2006). All of the samples extracted fell in this range, as shown Appendix B, showing that pure DNA samples were extracted. However, a total of seven samples had a lower concentration and six of them were older patients (60 to 82 years old), who had less than 10 mL of blood collected. All these samples were collected from HUKM and no samples from HP had a lower volume of blood or DNA concentration. The main reason was the faster blood taking procedure by the trained phlebotomist nurses in a designated blood-taking room in Hospital Putrajaya. The nurses usually hit the correct vein and withdrew blood on the first try, with limited number of second venipuncture observed. However, blood taking was done by the rheumatologist in HUKM or medical students present for training, both showing signs of difficulty in finding patients' veins numerous times. The designated blood taking room in HUKM did not entertain side projects and all blood taking for research were the doctor's responsibility. Sample HLAG30 was the only young patient of 33 years old to have a low DNA concentration, but who had 10 mL of blood collected. This might be due to other anthropometric, epidemiological and technical factors (Caboux et al., 2012).

Numerous factors affect the DNA yield, namely age, gender, body mass index (BMI) and tobacco consumption. Lower DNA yield is observed in men compared to women, due to the lower comparative lymphocyte, platelets and neutrophil count (Bain, 1996). A progressive decrease in DNA yield is also observed in older people with 0.11 mg loss per year and might be caused by the lower white blood cell count in the peripheral circulation (Erkeller-Yuksel et al., 1992; Richardson et al., 2006). BMI and tobacco consumption on the other hand cause an increase the DNA yield. Another possible source of DNA yield variation are

technical ones, but these were minimized by the several rounds of optimization performed (Caboux et al., 2012).

Difficulty in drawing blood in elderly patients can be attributed to several factors, namely, the aging process, thin skin, nerve damage and weaker blood vessels. Numerous failed attempts at drawing blood, as experienced several times in the two hospitals, leads to pain, bruising, hematomas, smaller veins moving and the risk of veins collapsing. Several methods were used to improve the blood taking procedure, namely, use of tourniquets, smaller needles, butterfly needles, changing veins/arms/location and careful palpation of the arm or hand. Even experienced nurses, doctors and phlebotomists run into major challenges in drawing blood from the elderly, even with all these tips to better the process (Klosinski, 1997). A few patients were lost during blood sample collection as enough blood could not be drawn. Gout patients in the study were mostly elderly patients and care needs to be taken to make their participation painless and smooth. Blood was collected as a DNA source due to its easier handling in a hospital setting. However, elderly patients were proven to have a lower DNA yield, as the latter decreases with age (Caboux et al., 2012). All the seven patients with lower DNA yield were men, with a mean age of 64 years old, higher than the total gout cohort age mean (55 years old). This proves that gender is also a factor to consider when collecting blood samples, further supported by 92.4% of the total gout patients being males (Bain, 1996; Caboux et al., 2012). Hence, in order to improve further studies, other DNA sources should be considered.

Other methods such as collection of whole saliva via the Oragene kit (OG-250) (DNA Genotek, Canada) could be implemented in such situations, without compromising the DNA quality and quantity (Nunes et al., 2012; Rylander-Rudqvist et al., 2006). The kit showed a range of 62 µg to 158 µg of DNA amount extracted, with a median of 110 µg (DNA Genotek, Canada). The Oragene kit had higher median DNA yield when compared to other methods, such as using a cotton swab (1.9 µg) (Cozier et al., 2004), Guthrie cards (2.3 µg) (Harty et al., 2000), Cytobrush (6.8 µg) (Garcia-Closas et al., 2001) and mouthwash (35.1 µg) (Le Marchand et al., 2001). Quantitative and qualitative tests were successfully performed on the kit, namely, spectrophotometry and genotyping. Quality of DNA was evaluated, at 98% success rate, by TaqMan, High Resolution Melt (HRM) and restriction fragment length polymorphism-PCR (RFLP-PCR). DNA extracted could also be stored for 8 months at room

temperature, without any effect on the DNA quantity and quality (Nunes et al., 2012). Saliva was also compared to urine and gave the highest yield and DNA concentration (Nauwelaerts et al., 2020). Several genetic and genotyping studies used this kit and obtained good results (Keag et al., 2020; Rylander-Rudqvist et al., 2006). The initial stage of primer testing of this study was performed using the Oragene kits in UNMC and no difference in amplification or HRM effectiveness was observed. This non-invasive method made prospective participants more willing to register for the study and should be considered for future use in hospital settings.

## 5.1.4 Optimization of Polymerase Chain reaction (PCR) and Cloning

### 5.1.4.1 Polymerase Chain Reaction (PCR)

PCR reactions with specific primers was the first step in building the HRM screening method. Sensitivity, specificity and accuracy of the HRM method started with a proper optimized PCR for the specific primers designed. Because the HLA-B\*58:01 sequence differs in many point mutations from HLA-B alleles, this raises more challenge and difficulties in technical method design and development process (Kang et al., 2016). Specific PCR primers were designed for exon 2 and exon 3 of the HLA-B\*58:01 allele, as these were the regions which codes for the peptide binding groove of the allele and as it has the most polymorphisms present there (Zhang et al., 2015).

MHC Class I proteins are expressed, to varying degrees, on the surface of all nucleated cells and are comprised of one transmembrane heavy chain with three extracellular domains ( $\alpha 1$ ,  $\alpha 2$ , and  $\alpha 3$ ) and a  $\beta 2$ -microglobulin light chain that anchors the heavy chain to the cytoplasmic membrane. The  $\alpha 1$  and  $\alpha 2$  segments, coded for by exons 2 and 3, form a peptide binding groove, which presents peptide antigens to CD8+ T lymphocytes (Mosaad, 2015). HLA class 1 alleles are also determined by the sequence of exon 2 and 3 (Shiina et al., 2012). Hence, these two exons were targeted as they determine the sequence of the peptide binding region and ultimately the specific HLA-drug immune response elicited (Madden & Chabot-Richards, 2019).

Numerous studies, using various methods, have amplified the HLA-B\*58:01 allele before by targeting its exon 2 and 3. Kang et al. used a real-time PCR assay to do so (Kang et al., 2016), followed by Kwok & Kwong using loop-mediated isothermal amplification (LAMP) (Kwok & Kwong, 2013) and Zhang et al. using a TaqMan assay (Zhang et al., 2015). Nguyen et al. used real-time PCR to target the area between exon 2 and 3, saying that polymorphisms located there can discriminate between HLA-B\*58:01 and the rest of HLA alleles. All these studies attest to the importance of targeting these 2 exons properly and this used in this study.

The 2 sets of primers targeting exon 2 and 3 were designed by paying attention to a set of criteria for production of a specific primer; the primer length should not be more than 20bp, with a GC content of around 40-60%, close melting temperature of the primer pair around

50-60°C and a low self- annealing value around 1-2 (Chaung et al., 2013). Primers P2 (for exon 2) and P3 (for exon 3) proved their specificity and stability by showing bright amplified bands (as shown in **Figure 4.1**) throughout the 4 years of research. The primers were optimised stepwise, starting with changing the annealing temperature, followed by the primer and DNA concentration. Sanger sequencing showed (**Table 4.6**) that P2 primers covered 86% of its target (sample HLAG26), with 97.2% being identical to exon 2 of the HLA-B\*58:01 allele. P3 primers showed better results, with 97% covered on exon 3 for the same sample and 100% being identical to exon 3. Hence, P3 showed better specificity and coverage capability and HRM results were focused more on this set of primer. Exon 3 showed the presence of more SNPs compared to exon 2 of the HLA-B\*58:01 allele, as shown in Appendix C and may thus be the reason for a stronger and more specific binding and amplification.

Stability of primers over 4 years was monitored by sending a number of samples for Sanger sequencing after amplification. However, the smaller coverage percentage can be attributed to the small gap left in the middle of exon 2 and exon 3 sequencing by Sanger, as shown in Appendix C. This was linked to the smaller resolving power of Sanger sequencing when faced with secondary and tertiary structures in the highly polymorphic exons. Both primer sets targeted the middle of their exons as shown in Appendix C, the areas where secondary and tertiary structures are most likely to happen.

Moreover, their Sanger sequencing results reflected this by having constant good coverage of the polymorphic exons for 28 samples, at around 95.15% and 98% for exon 2 and 3 (shown in Appendix C). P3 however had better specificity for the 28 samples, with an average of around 93% covered by Sanger sequencing results, along with an average of 98% of the P3 PCR products being similar to the exon 3 of HLA-B\*58:01. A small gap was seen in the middle of the exons, due to the presence of ambiguous DNA sequence and absence of any alignment score. The alignment score in NCBI BLAST shows how well the aligned sequences match, with the maximum score being 200, as shown in Appendix C. Hence, the gap seen, usually in the middle of the exons amplified, showed an area where the PCR product doesn't match the original exon and could possibly be one or several SNPs. One major problem in Sanger sequencing is the compression of bands during electrophoresis of the DNA fragments, which can lead to misreading of the sequence. Compressions usually

occur when a DNA fragment migrates faster than the others, leading to an overlap with an adjacent fragment and hence causes problems in the interpretation of results. These compressions are due to strong secondary structures in the DNA fragment. Other sequence motifs, such as palindromic sequences form stable hairpin structures which causes abnormal mobility during gel electrophoresis (Ronaghi et al., 1999). Sanger sequencing is known to have disadvantages when faced with the complex HLA-B alleles, such as incomplete sequencing, limitations in resolution, ambiguous HLA alleles called by typing and haplotype phase issues (Erlich, 2012; Shiina et al., 2012). The start and end of Sanger sequencing results were also made of several bases labelled as 'N' as specific bases couldn't be called as shown in Appendix C. This still could not be completely resolved by checking the chromatogram manually and making more accurate base calls. Genomic sequence variations in HLA alleles are caused by SNPs, indels (insertions and deletions) of one or more bases and repeat length polymorphisms and rearrangements. SNPs and indels are thus closely associated and found in HLA alleles, further making the Sanger sequencing process difficult (Longman-Jacobsen et al., 2003). Hence, the highly polymorphic exon 2 and 3 couldn't be fully sequenced and read by Sanger sequencing. This is why NGS was used later to resolve all these ambiguities and to increase the accuracy of HLA calling.

Internal Control (IC) primers were used in multiplex PCR reactions with P2 and P3 primers separately to confirm the correct amplification and optimum PCR conditions for both primer sets. The sequence of the IC primers were obtained from previous studies done on  $\beta$ -globin gene, located on chromosome 11 (Talmaci et al., 2004). Two other HLA studies used this  $\beta$ -globin gene as an IC for HLA-B\*27 genotyping (Bon et al, 2000; Faner et al., 2004). However, IC PCR products were very faint compared to P2 and P3 products as shown in **Figure 4.2**, even in the presence of DMSO as a stabiliser. DMSO was added as additives stabilize multiplex PCR reactions, preventing secondary structure formation and reducing the melting temperature of GC-rich sequences. Several rounds of optimisation were carried out (shown in Appendix D) for these two multiplex reactions, by changing the annealing temperature, primer concentration, DNA concentration and DMSO concentration. However, IC products always had fainter bands compared to P2 and P3 primers, showing stronger primer specificity and binding power in the latter. Research shows that preferential amplification of one target sequence, in this case exons 2 and 3 amplification by P2 and P3 primers, can

result in weaker amplification of the second target in multiplex reactions (Elnifro et al., 2000). Thus, proper amplification of the human HLA-B region was verified in this step, which acted as a checkpoint.

Another approach tested for the HLA-B\*58:01 allele amplification was the splitting of the full gene into halves and amplifying the two regions of around 1500bp (shown in Appendix E). The first set of primers (WGH1) amplified exons 1-3, while the other primer set (WGH2) amplified exons 4-7 of the HLA-B\*58:01 allele. This approach was tested to try to amplify the whole gene, thus helping to decipher and discover more of its role. However, due to the high number of SNPs throughout the full gene, only around 75% of the gene was amplified, split into two by the two sets of primers. WGH2 primers, targeting exons 4-7, showed less specificity, where only 56% of the exons were amplified, showing a possible higher number of polymorphisms there and a subsequent need to investigate the whole gene. Overlap PCR was also tested on WGH1 and WGH2 primers, to try to amplify the full HLA-B\*58:01 gene, to no avail.

#### **5.1.4.2 Positive control cloning**

Positive control cloning was a determining step for this whole study, as the latter needed to be isolated and produced in large amounts to be used for all future HRM screenings. Only a limited amount of positive control DNA was obtained from a patient in a previous study and cloning needed to be executed precisely and fast. However, only the presence of HLA-B\*58:01 was determined by the previous study in the positive control. The second HLA-B allele was unknown, hence created more ambiguity in the HRM results.

The Qiagen PCR Cloning plus kit (Qiagen, Germany) was used, as it offered an already optimised and accurate cloning method, leaving no chance of wastage of the positive control DNA. Cloning was successfully and quickly carried out and verified by double restriction digestion and Sanger sequencing. Due to the significantly bigger vector size (~4000bp), P2 and P3 products were seen as faint bands. However, Sanger sequencing confirmed presence of the two exons in the extracted plasmids, with 96.6% of exon 2 present and 100% of exon 3 present. P3 primers again showed better performance in cloning, compared to P2 primers. Hence, P3 primer results will be emphasized more in the following sections. The WGH1 and WGH2 primers were also tested out to clone the positive

control, in order to try and clone a bigger part of the HLA-B\*58:01 allele, to no avail. Double restriction digestion results, along with Sanger sequencing results were too erratic to be used in any way. In order to check stability and presence of the cloned positive control in the plasmids, a few cloned samples were always sent for Sanger sequencing from every single batch of produced and extracted clones. This was especially important as clones were frozen in -80°C in glycerol and countless sub-cultures and plasmid extraction cycles were done to produce enough reference DNA for all the HRM screening. Every single batch of positive control produced was checked for presence of the HLA-B\*58:01 gene by Sanger sequencing, before any HRM screening was performed. Appendix F shows an example of a cloned plasmid's Sanger sequencing results, with all the hits being the HLA-B\*58 allele and the majority being HLA-B\*58:01. There was a risk of mutation and alteration of the cloned positive control DNA which needed to be monitored and avoided at all costs. The cloned positive control showed proper resilience over the full four years of research, with similar Sanger sequencing results and presence of exon 2 and exon 3 targets in all plasmids.

### **5.1.5 Conclusion**

Hence, to conclude this first section of discussion, all the aforementioned factors were monitored and optimized in order to build the HRM screening method for the HLA-B\*58:01 allele. Due to the unavailability of any other positive controls, this one had to be used as no other HLA-B\*58:01 studies were being done in 2014 to 2015, when this research was started.

## 5.2 Core findings: Evaluation of the HRM method

### 5.2.1 Decoding the HRM results

The HRM screening was performed on the gout cohort of 145 patients and the volunteer cohort of 145 individuals, for both exon 2 and exon 3 of the HLA-B\*58:01 allele, separately. Hence, a total of 580 unique HRM screening results were generated for analysis purposes. Since this was a novel study in Malaysia in 2014 to 2015, it was important to look at all the results generated one by one and establish a certain trend for positive and negative results. This study used the theory of seeing a change in the HRM melt curves' shape for a negative HLA-B\*58:01 sample when compared to the positive control and a matching of HRM melt curves for a HLA-B\*58:01 positive sample. This theory stemmed from the ability of the HRM method to detect the slightest change in one DNA base, which will then be reflected in a change of the melt curve shape. This theory is called the reference curve-based targeted genotyping, and it has been used in the few HLA HRM studies done before (Rani et al., 2018; Imperiali et al., 2015; Cui et al., 2013; Lundgren et al., 2012; Zhou et al., 2004).

This research was based off a method previously developed in HUKM to screen for the HLA-B\*15:02 allele in Malaysian epileptic patients in order to avoid carbamazepine-induced SCARs. Authors used the HRM method's normalized and difference melt curves for their samples and compared it to their positive control and negative control. Hence, a significant deviation or closeness of HRM melt curves from either controls showed the absence or presence of the HLA-B\*15:02 allele. They validated their results against Sanger sequencing and showed a 100% match, along with 100% specificity and sensitivity of their method. Rani et al. also compared their HRM method to a multiplex-PCR HLA-B\*15:02 screening method and showed that the latter generated a false positive result and was less specific. Moreover, after their HRM method was implemented in HUKM before treatment, no cases of carbamazepine-induced SJS was recorded. This study provided a good foundation and proof of concept for this new research in a Malaysian hospital setting (Rani et al., 2018). However, the authors did not provide a detailed method by which HRM analysis was performed and only mentioned the use of normalized and difference melt curves. All the available HRM studies done in the past were gathered and their different methods of HRM analysis were applied in this research.

### 5.2.2 The HRM method: reference curve-based targeted genotyping

Each sample was screened in duplicates, along with its NTC for proper monitoring of reproducibility, accuracy and contamination. Results in this section will be focused on the P3 primers for exon 3 amplification of the HLA-B\*58:01 allele, due to its higher specificity and accuracy. The Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines were used to guide the real-time PCR part of the HRM method (Bustin et al., 2009). Once manual normalisation of signal intensity was performed by toggling the pre- and post-melt region (Wu et al., 2008), the results table was checked to make sure the HRM run was of good quality. The  $C_q$  value, also known as quantification cycle, was the first result checked as it shows specific amplification ( $C_q < 30$ ) and reproducibility of amplification (Bustin et al., 2009). Table 4.8 shows the results generated for the first gout sample, HLAG1, along with its positive control. The two NTC wells showed no  $C_q$  values, showing absence of any contaminants. This was the first step checked for all of the 290 samples and no NTC well showed any  $C_q$  value. None of the five HRM studies done on HLA alleles mention checking of  $C_q$  values and go directly to the specific HRM curve analysis (Rani et al., 2018; Imperiali et al., 2015; Cui et al., 2013; Lundgren et al., 2012; Zhou et al., 2004). Since HRM is an extension of real-time PCR, it is important to check the standard real-time PCR results too and make sure it is constant throughout the whole batch of 290 samples. MIQE guidelines stated that repeatability and reproducibility of results here refers to the precision and robustness of the assay with the same samples repeatedly analyzed in the same assay. This was assessed via the Standard deviation (SD) for the  $C_q$  value. Table 4.8 shows the small SD for the  $C_q$  values for both HLAG1 sample (0.05) and the positive control (0.07), hence proving good repeatability. This was checked for all the 290 samples and they all showed small SD for the  $C_q$  values (shown in Appendix G). The melting points shown in Table 4.9, also show consistency and minimum variations for both the HLAG1 sample and the positive control and this was repeatedly checked for all the 290 samples. Once all of these values fell within the required range, the normalized and difference melt curves were analysed.

Figure 4.5 and 4.6 however show an unexpected result, where the positive control's melt curves were seen to deviate from each other. This shows that this positive control is most

probably a heterozygous sample with one allele having a slight change in its DNA sequence. Unfortunately, only one positive control was obtained from HUKM's previous study and no other positive samples were identified. Hence this complicates the HRM analysis, but doesn't make it impossible. The heterozygosity of the positive control is probably the factor which contributed to the HRM method's 0% sensitivity. Adding to that the highly polymorphic HLA-B\*58:01 region and its numerous SNPs further complicated the HRM analysis. The change in a base was not observed properly due to these factors coming into play. The initial theory of an unknown, HLA-B\*58:01 positive sample completely matching the positive control's horizontal reference curve could not be used. However, the closeness of the HRM melt curves for unknown samples could be assessed, as the closer the curve and its shape, the closer the DNA sequence of the unknown sample is to the positive control. A standard for the reference was set, where one horizontal line with one downward peak with a melting point ( $T_m$ ) of 90.8°C was always seen for all the HRM curve analysis. The matching or closeness of the 290 samples' melt curves will further show more details for the HRM analysis.

The thermal stability of a PCR product is determined by its base sequence. When the PCR product sequence is altered, duplex stability is changed, leading to different melting behavior. When the change is homozygous, a shift in melting temperature is usually observed to the right. When the change is heterozygous, four duplexes are formed following PCR: two heteroduplexes and two homoduplexes. Each duplex will have differing stabilities, the sum of which can be observed by high-resolution melting analysis. Instability of heteroduplexes results in a lower  $T_m$  than homoduplexes. As heterozygotes contain both heteroduplex and homoduplex species, the heterozygote melting curve is a combination of their melting profiles (Montgomery et al., 2007). Furthermore, high-resolution melting curve analysis produces different shapes for different heterozygotes, often allowing identification of the specific sequence variant. Similarly, different homozygotes can often be distinguished, including homozygotes that differ by only a single base (Gundry et al., 2003).

Hence, both samples HLAG1 and the positive control are different heterozygotes, as shown by their double normalized melt curves (Figure 4.5), with similar melting points of 90.8°C. The sample HLAG1 (in green in Figure 4.5) has its melt curves in between the positive control's melt curves. The difference melt curve in Figure 4.6 further confirms this deduction

by showing two blue lines with different shapes (one horizontal and one downwards peak) for the positive control. The sample HLAG1 shows two close green lines with slightly different shapes, which are again within the two positive control's melt curves. HLAG1 sample can thus be deduced to be HLA-B\*58:01 negative as none of its duplicates matched any positive control's HRM melt curve.

This aforementioned method of HRM melt curve analysis was applied for all the samples, but this section will focus on the 28 ADRs/SCARs samples identified in the gout cohort. This increases the chances of seeing a trend for positive samples due to the strong association of the HLA-B\*58:01 allele to allourinol-induced SCARs samples. While analysing all the samples, a working threshold of 20 to -20 for the fluorescence level was seen, this distinguishes samples with close HRM curves and clearly distant curves. Interestingly, not a single sample from the 290 samples matched the positive control's curves completely. The melt curves shapes were all very different and did not even match the downward peak of the positive control. This proves that other HLA-B alleles and factors are at play here and need further investigation by sequencing later. By analysing the 28 ADRs/SCARs samples in Table 4.9, 8 samples (HLAG19, HLAG28, HLAG38, HLAG64, HLAG68, HLAG69, HLAP56 and HLAP64) were found to have the closest difference melt curves when compared to the positive control. HLAG19 and HLAG68 had the closest normalized and difference melt curves when compared to the positive control. Both samples showed two almost identical normalized curve shapes, with slight variations in shapes. Those variations became clearer in their difference melt curves of different shapes which were within the reference's two melt curves. However, the huge difference in all the ADRs/SCARs normalized and difference melt curves shown in Table 4.9 throws off the possible heterozygous and homozygous sample identification as aforementioned. Thus, this set of ambiguous results again proves that more than a few factors are at play here to divert the melt curves so much.

### 5.2.3 Validation and verification of the HRM method via Sanger sequencing on the gout cohort

In order to decipher the different factors affecting the HRM results, Sanger sequencing was done on all the 28 ADRs/SCARs samples as shown in Table 4.10. Interestingly, a few HLA-B alleles previously identified as pharmacogenetic markers were identified in this set of 28 samples. Nine samples (HLAG12, HLAG26, HLAG29, HLAG56, HLAG66, HLAP6, HLAP25, HLAP46 and HLAP60) were confirmed to have the HLA-B\*58:01:01 allele via Sanger sequencing and none of them were listed as possible positive samples in the previous close HRM curve analysis. This further validates that HRM melt curves are currently too ambiguous to be used as a definite screening method. Sanger sequencing results showed the presence of HLA-B\*58:01:01, up to 6 digits of resolution, in all the 9 positive samples. By using only those 9 positive samples, a frequency of 6.21% of the HLA-B\*58:01:01 is estimated to be present in the gout cohort (n=145) and a frequency of 32.1% in only the ADRs/SCARs cohort. However, this is an underestimation of the frequency of HLA-B\*58:01 present as not all the 145 gout samples were sent for Sanger sequencing due to limited grant funds. This frequency of 6.21% still falls in the wide range of Malaysian HLA-B\*58:01 frequencies identified in database mining done in the literature review (Table 2.9) (Ghattaoraya et al., 2016).

Interestingly, these 9 positive samples were linked to range of cADRs namely, exfoliative dermatitis, DRESS syndrome, AHS, itchiness, SJS and generalized rashes. Positive samples HLAG12 (vasculitis rash), HLAG66 (itchiness), HLAP25 (generalized skin rash), HLAP46 (rash on face), HLAP60 (generalized skin rash) all show presence of mild cADRs. HLAG29 and HLAG56 show presence of DRESS syndrome, followed by HLAP6 with SJS and HLAG26 with exfoliative dermatitis. This shows that HLA-B\*58:01 is not only involved with SCARs, but also with milder forms of allopurinol-hypersensitivity. A study done on Han Chinese patients, found HLA-B\*58:01 in 100% of all allopurinol-induced maculopapular eruption (MPE) patients, showing its involvement in more than just SCARs (Cao et al., 2012). Interestingly, one patient with allopurinol-induced MPE developed TEN four days after re-exposure to allopurinol. Several similar cases have also been reported and this proves that milder forms of allopurinol hypersensitivity shouldn't be ignored in studies (Carstens et al., 1994). However, one recent Australian report found that none of the 12 patients with the milder

allopurinol-induced MPE reaction carried the HLA-B\*58:01 allele, including one South East Asian origin patient (Lee et al., 2012). This discrepancy might be explained by the different ethnic backgrounds. Significant differences between DRESS and SJS/TEN in the pattern of immune responses have been found, causing a marked deviation in the pathological phenotype of severe drug eruptions (Pichler et al., 2010). The lymphocyte transformation test, a reflection of drug-specific T-cell responses, has been reported to be positive at the same acute stage in MPE and SJS/TEN, while positive reactions were only obtained at the recovery stage in DRESS, which suggested differences in the immunological reactions among three types of cADRs (Kano et al., 2007). Genetic predisposition and other factors leading to the diversity of the immunological reactions could be involved in the complex determinism of the various spectrums of ADRs caused by allopurinol. Hung et al. suggested that HLA-B\*58:01 was necessary but not sufficient for the occurrence of allopurinol-induced SCARs (Hung et al., 2005). Samples HLAG38 and HLAG39 both have SJS but are negative for HLA-B\*58:01. HLAG38 shows presence of HLA-B\*15:02 and HLAG39 shows presence of HLA-C\*12:02, thus proving that even in the absence of HLA-B\*58:01, SJS can manifest in gout patients. Hence, further investigation is needed to find the other factors responsible for SJS in these two patients. Cao et al. showed that HLA-B\*58:01 is thus a strong risk factor for not only SCARs, but also MPE. However, it is likely that other factors would also be implicated in the pathogenic process, and that may be one of the reasons why some HLA-B\*58:01 carriers did not present cADRs after exposure to allopurinol (Cao et al., 2012).

Another anomaly seen in the NCBI BLAST results are the presence of only 10% of the total 100 hits being the HLA-B\*58:01:01 allele. The rest of the 90% hits are distributed amongst other HLA-B alleles, with some of them being genetic markers for other drug-disease associations. This proves that the primers designed are not amplifying only the HLA-B\*58:01 region and that the latter has too high polymorphism seen. Hence, the amplification of exons for HLA-B\*58:01 detection is not feasible.

All the different HLA-B alleles seen in the ADRs/SCARs cohort were found in other pharmacogenetic association studies for HLA-B\*15:02 and HLA-B\*57:01, which will be detailed in this discussion section. The causes of this multiple HLA-B alleles marker concept will also be discussed to show how they are all interconnected in numerous diseases and not restrained to one HLA-drug-disease association. However, literature review revealed

only HLA-B\*58:01 focused studies and no studies evaluating the possibility of other HLA-B alleles' involvement. This may be caused by the timing of the discovery and validation of these pharmacogenetic markers, HLA-B\*57:01 being the first one, followed by HLA-B\*15:02 and finally HLA-B\*58:01. Literature shows that the usefulness, validity, accuracy and implementation of HLA-B\*58:01 as a pharmacogenetic marker is still being evaluated worldwide.

HLA-B\*35:01:01 and HLA-B\*58:02 were identified in all the 9 positive samples, along with HLA-B\*58:01:01, suggesting a possible strong link here. These 3 alleles always happened together in all the HLA-B\*58:01:01 positive samples and can thus be said to be other possible genetic markers for allopurinol-induced cADRs in Malaysian gout patients. However, more investigations are needed to validate this. Pompeu et al. stated that allopurinol hypersensitivity responses can be caused by HLA-B\*58:01 allele only and not the closely related HLA-B\*58:02 allele. There are six residues that distinguish HLA-B\*58:01 from HLA-B\*58:02, three of which are located within the antigen-binding cleft. Since the designed primers target the peptide-binding groove (exon 2 and exon 3 of the HLA-B\*58:01 allele) it can be deduced that these 3 residues are not enough to distinguish between HLA-B\*58:01 and HLA-B\*58:02 alleles. Hence, this encourages the amplification of the whole HLA-B\*58:01 allele and not just the exon 2 and 3 for allopurinol-induced ADRs/SCARs studies. Allopurinol may bind in the cleft contacting polymorphic HLA residues and altering the repertoire of bound peptides, or it may bind outside the antigen-binding cleft in a manner that influences T-cell recognition. An alternative possibility is that allopurinol binds to multiple sites on the HLA molecule, including polymorphic sites and perhaps conserved sites. For example, the F pocket has chemical and geometric features that are compatible with allopurinol binding. Allopurinol binding the F pocket in HLA-B\*58:01 individuals would be expected to alter HLA-binding motif at the ninth position (due to F pocket steric effects) and the P5 position (due to HLA-B\*58:01 specific residues). Small molecule drugs such as allopurinol may bind a diverse set of sites on HLA molecules and trigger adverse effects through alternative unknown mechanisms that including direct TCR binding. Hence, with the still ambiguous altered-peptide repertoire mechanism, along with allopurinol binding to multiple sites on HLA molecules, the drug-HLA-SCARs association still need deeper

deciphering (Pompeu et al., 2012). Pharmacogenetic research thus provides the opportunity to explore other possible mechanism of actions for drug-HLA-SCARs interactions.

Another cluster of HLA-B alleles, namely HLA-B\*57, HLA-B\*58:01, HLA-B\*58:02 was found in sample HLAP6 with SJS. The HLA-B\*15 allele was also found in HLAP6, showing a total of three possible genetic markers (HLA-B\*57, HLA-B\*15, HLA-B\*58:01) in SJS. In sub-Saharan Africa, HLA-B\*57, HLA-B\*58:01, and HLA-B\*81:01 are protective against human immunodeficiency virus type 1 (HIV-1) infection progression, while HLA-B\*18:01 and HLA-B\*58:02 are associated with rapid progression. One contributory factor is the role of HLA molecules in presenting particular epitopes to induce CD8+ T-cell responses against HIV-1. Gag-specific CD8 + T-cell responses are associated with lower viremia levels and are dominant in people with protective HLA alleles, such as HLA-B\*57:03. In contrast, disease-susceptible HLA alleles, such as HLA-B\*58:02, typically present relatively ineffective non-Gag responses associated with high viral set points (Leitman et al., 2016). HIV-1 infected Africans were found to have protective HLA-B\*57:03 and HLA-B\*58:01, then were susceptible to HLA-B58:02. In addition, HLA-B35:01 is strongly associated with high HIV-1 viral load in infected cohorts in Japan (Mori et al., 2014). HLA-B\*57:01 was also discovered as a functional determinant of flucloxacillin-induced liver injury (Monshi et al., 2013). HLA-B\*57:01 and B\*58:01 are part of the same HLA-B17 serotype; they differ in structure by only five amino acids and have a significant overlap in their antigenic peptide repertoire. Abacavir has recently been shown to interact with amino acid residues located deep within the HLA-B\*57:01 (but not HLA-B\*58:01) binding groove and alter the repertoire of self peptides that are presented to CD8+ T cells. In contrast, flucloxacillin does not alter peptide binding to HLA-B\*57:01, which shows that functional flucloxacillin antigens derive from naturally processed drug-protein conjugates. Interestingly, activation of CD8+ clones was detected with flucloxacillin-pulsed antigen-presenting cells from volunteers expressing both HLA-B\*57:01 and B\*58:01 in this study (Monshi et al., 2013). He et al. used several methods for HLA typing and recognized a few HLA alleles in a Leukemia patient, namely, HLA-B\*15:01, B\*15:20 and B\*58:01. However, this turned out to be a case of HLA-B allele dropout and intronic polymorphism, which generated a novel allele, HLA-B\*58:01:01:02. The later had some nucleotides difference at 509 C>T, 521 T>G and CCC insertion in position 503 of intron 2, compared to HLA-B\*58:01:01:01 (He et al., 2016). This may have happened

in the results too, as several HLA-B alleles were linked together (58:01:01/58:02/35:01:01:01) in the positive samples, thus indicating the possible presence of a novel allele in Malaysian gout patients.

The presence of HLA-B\*35:01:01:01 (in 9 positive samples), HLA-B\*35:02 (1 sample) and HLA-B\*35:03:01 (2 samples) in the HLA-B\*58:01 positive samples, shows another possible association. HLA-B\*35 alleles including HLA-B\*35:05 are known to be linked to nevirapine hypersensitivity. The most significant difference between HLA-B\*35:05 and HLA-B\*35:01 is at position 121 in the floor of the antigen-binding cleft in a site where intermolecular contact with peptide at P5 may be altered (Pompeu et al., 2012). The HLA-B\*35:01:01 allele has also been previously associated, in its haplotype HLA-A\*02:01:01/B\*35:01:01/C\*04:01:01, with with lamotrigine-induced maculopapular exanthema (MPE) in Mexican Mestizo patients. This association hence links HLA-B\*35:01:01 to HLA-B\*15:02, which is also responsible for lamotrigine-induced cADRs (Fricke-Galindo et al., 2014). HLA-C\*04:01:01 was identified in 4 gout samples, but none of them showed presence of HLA-B\*35:01:01 too. HLAG18 showed mild allergy as rashes and had only HLA-C\*04:01:01 identified via Sanger sequencing. HLAP2 had only HLA-C\*04:01:01 present too and showed generalised rashes. HLAG69 showed an ambiguous allergy to allopurinol and showed presence of HLA-C\*04:01:01 and HLA-B\*40:01/48. HLAP37 showed presence of HLA-C\*04:01:01 and HLA-B\*56 and had generalized rashes.

A Thai study found that HLA-B\*35:05 acts as a protective allele, along with HLA-B\*57:01 in HIV-1 infection. HLA-B\*27 has also been identified as a protective allele in white populations and HLA-B\*58:02 as a disease susceptible allele in African populations. This Thai study also shows a big set of HLA alleles identified in HIV-1 patients. 111 alleles (27 HLA-A, 58 HLA-B, and 26 HLA-C alleles) were detected, with the most prevalent HLA-B alleles being HLA-B\*46:01 (26.6%), followed by HLA-B\*40:01 (20.1%) and HLA-B\*13:01 (17.8%). HLA-B\*57:01 was detected in only 1.4% of this study. Of the HLA-C alleles, HLAC\*01:02 was the most prevalent (31.6%), followed by HLA-C\*07:02 (30.3%) and HLA-C\*08:01 (18.3%) (Mori et al., 2014). No HLA-A alleles were identified in the ADRs/SCARs cohort, but we identified HLA-B\*40:01 (7.1%), HLA-B\*57 (3.6%), HLA-C\*08:01 (3.6%), HLA-C\*07:02:01 (3.6%), similar to the aforementioned Thai study. This again proves that a cluster of HLA alleles are common in several populations and act as risk factors for ADRs/SCARs, albeit at different frequencies

and for multiple diseases and infections. Mori et al. also showed that HLA molecules with a more restricted peptide-binding repertoire, such as HLA-B\*27, HLA-B\*57, and HLA-B\*35:05 are more likely to be protective alleles, showing the importance of further studies on the peptide-binding repertoire again. HLA-B\*58:01 was the fourth most prevalent HLA-B allele in this Thai study, 17.6% by population frequency. Moreover, HLA-B\*58:01 was not protective in this HIV-1 cohort, similar to an African-American cohort, but contrasting the results for Native Africans and Whites. HLA-B\*58:01-positive patients possessed the T242N escape mutation in Gag p24 and other compensatory mutations, namely, H219Q, M228I, or G248A. There was a high frequency of compensatory mutations in the HLAB58:01-positive individuals in this study, suggesting that it was easier for the HIV-1 virus to adapt to immune pressure exerted HLA-B\*58:01 (Mori et al., 2014). Hence, mutation in HLA-B alleles is another branch that needs to be emphasized on, as their roles can change completely according to their mutational effects.

Other HLA-B alleles were identified in the ADRs/SCARs cohort at high frequencies, namely, HLA-B\*15 (17.9%), HLA-B\*15:02 (14.3%), HLA-B\*44:03:01 (10.7%), HLA-B\*44(10.7%), HLA-B\*48 (7.1%) and HLA-B\*35:03:01 (10.7%). HLA-B\*15:02 was found in 14.3% of ADRs/SCARs samples. HLA-B\*15:02 is a well-known genetic marker for carbamazepine, lamotrigine and phenytoin-induced cADRs, which are structurally related through an aromatic ring (Fricke-Galindo et al., 2014). There are numerous other genetic markers discovered in literature which corresponds to the results. Chang et al. did a thorough review of pharmacogenetic testing available worldwide and found a large number of HLA alleles linked to CBZ-induced SCARs. HLA-B\*15:02 and HLA-B\*15:11 were potential risk factors for CBZ-induced SJS/TEN in Japanese and Korean patients. HLA-A\*31:01 was reported as a marker for CBZ hypersensitivity in Europeans. HLA-A\*31:01 was strongly associated with CBZ-induced DRESS but not SJS/TEN in European and Han Chinese populations. Occurring at a low frequency in Japanese, Korean, and European populations, HLA-A\*31:01 has been shown to closely correlate with CBZ-induced DRESS by meta-analyses. A prospective screening of HLA-A\*31:01 in new Japanese CBZ users showed effective prevention of hypersensitivity reactions induced by CBZ. In addition, a recent study from RegiSCAR group showed HLA-B\*57:01 was strongly associated with patients of CBZ-induced SJS/TEN in Europeans. HLA-B\*59:01 was also reported to be a risk gene for CBZ-induced SJS/TEN (Chang et al., 2020).

HLA-B\*14:01 and HLA-B\*35:01 were also found to be linked to trimethoprim-sulfamethoxazole induced liver injury (Li et al., 2020). HLA-B\*40:02 and DRB1\*04:03 are risk factors for oxcarbazepine-induced maculopapular eruption (Moon et al., 2016).

Rani et al. showed that 41.9% of patients treated with CBZ developed rashes on the skin and they were positive for other types of HLA-B alleles; HLA-B\*40:01,\*35,\*48:31,\*48:12,\*51:01,\*58:11,\*41:02,\*44 (Rani et al., 2018). Two Thai studies also showed a similar set of HLA-B alleles identified in their cohorts, when compared to this study. Puangpetch et al. previously reported HLA-B polymorphisms from 986 Thai individuals. The top five of such HLA-B alleles consisted of HLA-B\*46:01 (11.51%), HLA-B\*58:01 (8.62%), HLA-B\*40:01 (8.22%), HLA-B\*15:02 (8.16%), and HLA-B\*13:01 (6.95%) (Puangpetch et al., 2015). Another Thai population also showed multiple HLA-B alleles and their frequencies; HLA-B\*46:01 (14.04%), -B\*15:02 (7.66%), -B\*40:01 (6.60%), -B\*58:01 (6.38%), -B\*13:01 (5.96%), -B\*44:03 (4.47%), and -B\*38:02 (4.26%) (Satapornpong et al., 2020). Huh et al. also showed the presence of multiple HLA-B alleles, with smaller percentages, in a Korean cohort; 07:02 (0.1%), 13:02 (0.08%), 15:01 (0.08%), 27:05 (0.07%), 35:01 (0.08%), 40:01 (0.07%), 40:02 (0.13%), 40:06 (0.13%), 44:03 (0.14%), 46:01 (0.1%), 48:01 (0.11%),51:01 (0.13%), 54:01 (0.07%), 57:01 (0.20%), 59:01 (0.06%) (Huh et al., 2018).

Other HLA-B\*58 alleles were also identified in smaller percentage, namely HLA-B\*58:28 (3.6%), HLA-B\*58:34 (7.1%) and HLA-B\*58:65 (3.6%). Huh et al. showed similar results in a Korean study, where their primers could not discriminate between HLA-B\*58:01 and HLA-B\*58:02, HLA-B\*58:31N, HLA-B\*58:76. Interestingly, HLA-B\*58:02 wasn't found in any Asian population before and HLA-B\*58:31N and HLA-B\*58:76 are rare alleles (Huh et al., 2018). Two other studies also showed similar results, where they couldn't distinguish between these specific HLA-B\*58 alleles (Kang et al., 2016; Nguyen et al., 2017). Moreover, their primers also bound to several HLA-B\*57 alleles; HLA-B\*57:01, HLA-B\*57:03, HLA-B\*57:05, HLA-B\*57:06, HLA-B\*57:11, HLA-B\*57:29, HLA-B\*57:79N, HLA-B\*57:82 and HLA-B\*57:83 (Huh et al., 2018). Primers P2 and P3 in this study showed better specificity at distinguishing between HLA-B\*57 and HLA-B\*58 alleles, with only HLA-B\*57 detected in one sample. Two samples, HLAP3 and HLAP51, had ADRs to both allopurinol and febuxostat, but none showed the presence of the HLA-B\*58:01 allele. HLAP3 only had the HLA-B\*44:03:01 allele present, while HLAP51 showed the presence of two alleles, namely HLA-B\*15:02 and HLA-

B\*14:02. Only one study checked HLA allele presence in a patient with allopurinol and febuxostat allergy. The patient had the HLA-B\*58:01 allele, but there was a drug-drug interaction of allopurinol with bupropion, hence making the result ambiguous (Tana & Vadas, 2019). More studies are needed to check which HLA alleles are involved in double allergy to allopurinol and febuxostat. Heterozygous samples were also identified, HLA-B\*40/HLA-B\*48 in one sample and HLA-C\*08/-B\*40 in another sample.

Sanger sequencing performed on the ADRs/SCARs patients opened up a whole new perspective on screening of HLA alleles. The focus has always been on the strength of single HLA-drug-disease associations in different populations and ethnicities. However, the HLA alleles are a complex, polymorphic and tightly connected set of genes which work together to modulate our immune response. This concept is reflected in these findings, where a cluster of HLA-B alleles are responsible for a wide range of ADRs/SCARs in the gout patients. This was also reflected in numerous studies, as aforementioned, with the same HLA alleles found in different populations. Moreover, the previously validated pharmacogenetic markers, such as HLA-B\*57:01, HLA-B\*15:02, HLA-B\*58:01, were found to have roles in all their diseases. Other factors which further complicate HLA-omics are the ambiguity of HLA's altered-peptide repertoire mechanism, multiple binding sites of drugs on HLA molecules, presence of protective HLA alleles and ongoing mutations in HLA molecules. The future of pharmacogenomics will thus rely on screening for a set of pharmacogenetic markers, rather than one HLA allele, and can be applied for numerous diseases, infections and ADRs.

#### 5.2.4 Combining Sanger sequencing and HRM results to find a specific trend in results

Once the HRM melt curve analysis proved too ambiguous to be used for HLA calling, the Sanger sequencing results were combined with HRM results to try to find a specific trend. The sequencing results were aligned to their specific HLA allele hits and SNPs present were analysed together with the  $T_m$  changes when compared to the positive control and the resulting change in HRM difference melt curve (Table 4.13-4.15). Similar base changes usually caused a similar change in the shape of the HRM difference melt curve, hence allowing the association of specific base changes to a specific peak or curve shift on the difference melt curve. After PCR amplification of heterozygotes, four duplexes are formed: two homoduplexes and two heteroduplexes. Each duplex has a characteristic melting temperature, and the sum of all transitions can be observed by melting curve analysis. The presence of heteroduplexes changes the melting curve shape of amplified heterozygotes compared with homozygotes. The changes are small but can be reliably detected with high-resolution melting analysis. This technique has also been applied to single-nucleotide polymorphism (SNP) typing (Reed et al., 2004).

Reed et al. showed a sensitivity of 98% for their HRM analysis of SNPs for a data set that included PCR products up to 1000 bp with the SNP centered. However, this was not possible in this study as the primers were not designed specifically for SNPs targeting and also due to the presence of too many SNPs and HLA alleles amplified. Reed et al. also proved that as the product length increases, the difference between wild-type and heterozygote curves became smaller, and the calls seemed to become harder to make (Reed et al., 2004). Even if the designed primers targeted a smaller region of 149-249 bp, the  $T_m$  changes were too erratic and Sanger sequencing results had gaps. It was difficult to identify a trend due to the erratic  $T_m$  change seen for similar base changes and samples with similar HLA allele hits called. SNPs typing is usually limited to only one of two bases at a given position (Gundry et al., 2003) and this was not the case in this study. Complete genotyping of SNPs by HRM was possible in 90% of cases with short PCR products, but only in the presence of a heterozygous, homozygous and wild-type DNA control (Palais et al., 2005). Heterozygous and homozygous samples could not be differentiated due to the lack of more controls and as the positive control itself was heterozygous. Moreover, a lot of SNPs were seen in these

samples (Table 4.14-4.16), and they cancelled each other out, due to similar bases and base changes. This was shown by Palais et al. too, due to neighbor thermodynamic symmetry, where bases adjacent to the SNP are identical on both DNA strands and the SNP consist of an interchange between complimentary bases (Palais et al., 2005). No trend was seen in SNPs,  $T_m$ , GC % of HLA-B\*58:01 positive samples and thus could not be linked to a change in melt curve shape. Hence, this again shows that other factors are at play and are causing the erratic results.

Literature showed four studies using HRM specifically for HLA typing, all using different methods of analysis and HLA typing. Zhou et al. established HRM typing for HLA-A alleles by targeting its exon 2 and 3 too, but combined nested PCR to improve sensitivity and specificity, followed by HRM analysis for transplantation purposes. Moreover, they had eight different samples of known genotypes to use as controls for their HRM analysis (Zhou et al., 2004). Imperiali et al. designed a HRM assay to detect all subtypes of HLA-B\*51, the strongest known genetic risk factor for Behçet disease. Authors used another method, with the presence of double peaks in the melting curve for positive samples and internal control presence, while negative samples showed only one peak for the internal control. However, this method couldn't be applied to current results as the internal control wasn't included in HRM screening as it was dominated by P2 and P3 primers. Moreover, authors amplified all HLA-B\*51 subtypes here, instead of one specific four digit HLA-B allele (Imperiali et al., 2015). Cui et al. developed a HRM method for the genotyping of HLA-DQA1 allele by targeting two specific SNPs and in the presence of several DNA references. Hence, the two SNPs were easily identified by the altered shape of the melting curves or the homozygote (AA or GG) was identified by a change of  $0.4^{\circ}\text{C}$  in melting temperature (Cui et al., 2013).

Rani et al. used the HRM method to screen for HLA-B\*15:02 epileptic patients with SCARs presentation, by targeting the exon 2 and 3 on a Rotor-Gene 6000 HRM System (Corbett Life Science, USA) (Rani et al., 2018). This method was used as a foundation for this research, but the completely different results raises the possibility of a problem with the instrument's resolution. Instruments specifically designed for HRM displayed the least variation, suggesting better scanning sensitivity and specificity compared to instruments with wider applications. (Herrmann et al., 2006). He et al. indicated indicated that multiplatform should be used to improve the HLA typing accuracy when a conclusive HLA genotype cannot be

determined (He et al., 2016). However, only the Eco-Real Time Illumina machine was capable of performing HRM in the University of Nottingham and it has a wide application, is not HRM specific and did not go through regular maintenance. The power of DNA melting analysis depends directly on the resolution of the melting instrument. Melting resolution depends on temperature control, temperature measurement, and fluorescence measurement. Temperature control determines both intra- and intersample variations. Melting rate, sample volume, and sample geometry affect temperature variations within a sample. Fixed, 96-well instruments showed the greatest sample variation, whereas the single-sample instrument had the least. This machine had 48 wells and thus was prone to greater sample variation. The resolution of temperature and fluorescence measurement depends on sensor and electronic noise, signal intensity, integration time, and the bit depth of analog-to-digital conversion. Overall measurement resolution is also determined by melting rate, extent of data averaging, and the resulting data density (Herrmann et al., 2006). Hence, many factors were identified as variables which contributed to the erratic HRM results seen in this study.

### 5.2.5 Validation and verification of the HRM method via Sanger sequencing on the volunteer cohort

All the aforementioned methods used for HRM and Sanger sequencing analysis was repeated for the healthy volunteer cohort to try to find a trend for HLA-B\*58:01 positive and negative samples. All the volunteers did not have gout and did not take allopurinol. In terms of HRM melt curve shape analysis, no specific trend was seen again, as shown in Table 4.17. By using the same temporary fluorescence threshold of 20 to -20 on the difference melt curve, 51 samples were called as positive and 94 samples as negative for the HLA-B\*58:01 allele. This was however impossible as it yields a higher percentage of positives (35%) compared to the gout cohort (15%). This high percentage by far exceeds and contradicts the Malaysian HLA-B\*58:01 frequencies identified in database mining done in the literature review (Ghattaoraya et al., 2016). The lack of HRM melt curve trend here, further supports the results found in the gout cohort, which points towards HRM result ambiguity due to the presence of other variables. No trend was seen again in HRM curve shape amongst the 4 positive samples identified, similar to ADR samples. NS6 was the only sample showing the presence of the HLA-B\*35:01 and HLA-B\*53:01 alleles, without any HLA-B\*58:01 hit like the 4 normal, positive samples. Samples identified with close HRM curves, namely NS11, NS26 and NS28, all had HLA-B\*15 alleles as the highest hit, thus showing that HRM curve shape was maybe influenced by the presence of numerous alleles, and not just the desired, HLA-B\*58:01 allele. This was further validated by sample NS68, which had the HLA-B\*15 alleles as highest hit, but had a hugely difference in the HRM curve shape.

The sequencing results for the normal samples showed HLA-B\*15/35/53 as the highest hit, compared to HLA-B\*58/35 being the highest in the ADR samples. More variants of the HLA-B\*15 allele were found in the normal samples, compared to the gout/ADR samples. The HLA-B\*35 allele had similar variants in both cohort, but were present at a higher percentage in normal samples. The most abundant alleles in the volunteer cohort were HLA-B\*15 (35.7%), HLA-B\*35 (35.7%), HLA-B\*53 (35.7%), followed by their variants HLA-B\*15:02 (28.6%), HLA-B\*15:21 (28.6%), HLA-B\*35:01:01 (28.6%), HLA-B\*53:01:01(28.6%), as well as HLA-B\*58 (28.6%) and HLA-B\*58:01:01(28.6%). The volunteer cohort thus shows the presence of a similar set of HLA-B alleles, previously identified in the ADRs/SCARs gout

cohort. This further confirms the initial theory of the presence of multiple pharmacogenetic markers in Malaysians, regardless of the presence of a disease or not. Sanger sequencing identified the presence of four HLA-B\*58:01:01 positive samples, along with the previously linked HLA-B\*35:01:01 allele, in the random 14 volunteer samples. HLA-B\*58:01:01 allele comes in as the fourth most abundant allele at 28.6%. A significant percentage of the volunteer cohort thus showed similar percentages of HLA-B\*58:01:01 (28.6%), HLA-B\*58 (28.6%) and HLA-B\*35:01:01 (28.6%). The ADRs/SCARs cohort (n=28) previously analysed showed 32.1% of these three HLA-B alleles. This shows that these 2 alleles have a strong link to Malaysians, even in the absence of gout and accompanying ADRs/SCARs. However, HLA-B\*58:02 was not found linked to these two alleles in the volunteer cohort, compared to the gout cohort.

This proposes a possible link of HLA-B\*58:02, as a susceptible allele, in gout patients with ADRs/SCARs, but not in the general healthy population. There are a few explanations for this found in literature, mainly, the mutation of HLA-B alleles, the binding of allopurinol to more than one HLA site causing ADRs through alternative unknown mechanisms, mutations in germline cells, the apparition of new alleles with selective advantages/disadvantages and micro variation within serotypes (Williams, 2001; Pompeu et al., 2012). One example is the allele *DRB1*\*0802, which was present in high frequency in Bering Strait Immigrants who moved to America and the resulting subtle variants of *DRB1*\*0802 such as \*0807, \*0811 and \*0802 found in present day native Americans. These variants are thought to have arisen in individuals with *DRB1*\*0802 through mutations in germline cells (Williams et al., 1994; Mack & Erlich, 1998). New alleles appear to emerge at a fairly high rate and become fixed in populations, in theory, if they provide a selective advantage in presenting peptides from infectious organisms. New alleles are generated via point mutation, recombination, and gene conversion-like events. This was shown in AIDS too, where HLA alleles may confer resistance or susceptibility to disease progression, along with dosage effects. Recent studies have suggested that homozygosity at HLA loci and/or the presence of *A*\*29, *B*\*35-*Cw*\*04, *B*\*54, \*55, and \*56, and *DRB1*\*11 are associated with more rapid progression, while HLA locus heterozygosity and/or the presence of *B*\*14 and *Cw*\*08 are associated with longer latency periods (Carrington et al., 1999; Hendel et al., 1999). HLA-B\*58:02 was also identified as the principal risk allele in HIV patients in African cohorts, while HLA-B\*51:01

appears both protective and disease-susceptible, depending on the geographic and evolutionary context (Goulder & Walker,2012). HLA-B alleles can thus be both protective and disease-susceptible, proving the constant mutations happening and the need for further investigation.

The HLA-B\*15:21 allele was found only in the volunteer cohort, the sixth most abundant allele here (28.6%). Jaruthamsophon et al showed a significant association between CBZ-SJS and HLA-B\*15:21 and HLA-B\*15:11. HLA-B\*15:21 is known to be common in Southeast Asian and neighbouring populations whose HLA-B\*15:02 frequency is also high. Hence, it is not surprising that association studies focused on the globally verified HLA-B\*15:02 and not on HLA-B\*15:21. Although it is a significant new finding, the positive association between CBZ and HLA-B\*15:21 is not an unexpected finding, since the HLA-B\*15:21 marker is a member of the HLA-B75 serotype group, similar to *HLA-B\*15:02* and *B\*15:11*. These HLA-B75 serotype members, share the 63<sup>rd</sup> arginine residue which is critical for the formation of the drug-HLA molecule complex by a direct pharmacological interaction. The molecular docking study revealed that HLA-B\*15:11, HLA-B\*15:21, and HLA-B\*15:08 are also capable of binding to the CBZ molecule directly. Moreover, authors also showed that CBZ was bound to HLA-B\*15:01, but at a groove not participating in the antigen-presenting domain. Thus, patients with CBZ-SJS might have negative screening for HLA-B\*15:02 but positive screening for other HLA-B\*15 alleles (Jaruthamsophon et al., 2017). Filipino patients were also found to have the HLA-B\*15:21 link to CBZ-SJS/TEN (Capule et al., 2020). A recent worldwide review (Then & Raymond, 2019). showed the presence of 15 different HLA-B alleles (*HLA-B\*15:02*, *\*31:01*, *\*58:01*, *\*15:21*, *\*13:01*, *\*38:02*, *\*51:01*, *\*56:02*, *\*56:04*, *\*15:13*, *\*38:01*, *\*24:02*, *\*13:02*, *\*40:02*, *3\*5:01*), from 15 different countries, in drug-induced cADR in epileptic patients. Malaysian studies showed the involvement of HLA-B\*15:02 and HLA-B\*15:13 alleles only (Then & Raymond, 2019). Hence, this study fully supports the findings of numerous pharmacogenetic markers presence and contradicts the one HLA-drug-disease association. However, literature search revealed no such HLA-B\*58:01 association studies with the presence of numerous pharmacogenetic markers.

Other HLA-B alleles were also found exclusively in the volunteer cohort, namely, HLA-B\*53, HLA-B\*53:01:01, HLA-B\*13, HLA-B\*13:02 and HLA-B\*13:01. The prevalent African allele HLA-B\*53:01 was linked to raltegravir-induced DRESS syndrome in HIV patients. Virtual

modeling suggests that raltegravir may bind within the antigen binding cleft of the HLA-B\*53:01 molecule, but not within the closely related HLA-B\*35:01 molecule (Thomas et al., 2017). Only one Malaysian study was found with DNA typing of the HLA-B gene in the peninsular Malaysian aborigines by using the SSOP method and comparison done to conventional serological typing (Hirayama et al., 1996). 56 unrelated individuals were investigated for their DNA polymorphism of the HLA-B gene, and 13 different HLA-B alleles were found; HLA-B\*15:13, \*35, \*58, \*40, \*62, \*38:02, \*57, \*18, \*13, \*40, \*51, \*7 and \*18. Interestingly, a rare allele, HLA-B\*15:13 was found to be prevalent in the different groups of Malaysian aborigines and shared the C-terminal peptide binding pocket (pocket F) with an African resistant type against severe malaria, HLA-B\*5301. Amino acid sequence of this major allele, HLA-B\*15:13, in the aborigines especially around the peptide binding groove (B and F pockets), was compared with that of African B\*5301 that had been suggested to confer resistance to malaria infection in Africa. The amino acid residues composing of the F pocket were completely identical in B\*1513 and B\*5301. HLA-B\*5301 and B\*3501 had a difference located in the F pocket, but only HLA-B\*53:01 was seen linked to malarial protection due to its high affinity to some major antigenic peptide produced in the host hepatocyte. Hence the shape of the C-terminal pocket was critical for the development of immunity against malaria. These observations suggest that a common environmental factor, the malaria infection, might have independently enhanced the evolution and selection of functional change in the polymorphic portion of HLA-B gene in Africa and in South-East Asia. Thus, the fact that B\*1513 and B\*1502 differ only in Bw4/6 within the a1 domain implies that the same kind of genetic events such as gene conversion may have happened during the course of evolution of the HLA-B gene in Malaysian aborigines. This study raised the possibility that the specific shape of the F pocket shared by B\*53:01 and B\*15:13 had some advantage to be maintained in the endemic areas for malaria (Hirayama et al., 1996).

A recent study showed that the HLA-B\*13:01 allele was significantly associated with AEDs-induced cross-reactivity, in Chinese and Thai (Min et al., 2019). The rate of cross-reactivity caused by two aromatic anti-epileptic drugs (AEDs) was higher than that caused by one aromatic drug and one non-aromatic drug. It was reported that cross-sensitivity induced by aromatic AEDs (CBZ, LTG, OXC, PHT, and PB) occurred in 40–58% of patients because of their similar chemical structures and intermediary metabolism molecules (Handoko et al., 2008).

Previous studies demonstrated that HLA-B\*13:01 was associated with PHT-SJS/TEN in Han Chinese and with severe cADRs induced by PHT in Thai people (Hung et al., 2010; Yampayon et al., 2017). Thus studies revealed revealed that HLA-B\*13:01 may be a potential risk factor for cADRs induced by anti-epileptic drugs (AEDs). Additionally, other studies also reported that HLA-B\*13:01 was associated with cADRs induced by dapsone and salazosulfapyridine (Yang et al., 2014; Tempark et al., 2017). Another study showed that HLA-B\*13:01 is strongly associated with dapsone DRESS in Taiwanese and Malaysian populations (Chen et al., 2018). The HLA-B\*13:01 allele is present as a medium common allele in Asian populations. Based on these data, HLA-B\*13:01 may be used as a common genetic marker for predicting cADRs in Asian populations. A few HLA-C alleles were found exclusively in the volunteer cohort, namely, HLA-C\*15 and all its variants (HLA-C\*15:10:02, HLA-C\*15:02 and HLA-C\*15:05). Very little is know about HLA-C's role in disease associations. HLA-C is a major determinant for NK cell activity and has high importance in pregnancy and placentation. HLA-C has also been associated with many diseases, including viral infections, cancer, autoimmune diseases and transplant failure. More studies are needed to properly valid HLA-C alleles' disease association (Papúchová et al., 2019).

The erratic HRM results worked in our favor, as we focused on finding another reason/cause for our results, instead of the typical one HLA-drug-SCAR interaction. This pushed us to forget about the normal expected theory and look at this from another perspective, as one would look at an anomaly. This perspective was also brought about by the fact that this was a novel research in Malaysia in 2014-2016 and no data was available to show a certain standard or trend we should adhere to. We took a discovery approach and matched the results to literature to show that it differs completely.

### 5.2.6 HRM screening method's performance evaluation

The HRM method has been proven to have a 90-100% sensitivity and specificity in numerous studies, with varied positive predictive (PPV) and negative predictive values (NPV) (Reed et al., 2004; Cui et al., 2013; Pavlov et al., 2015; Imperiali et al., 2015; Rani et al., 2018; Saskit et al., 2017). Reed et al. showed a sensitivity of 98% for their HRM analysis of SNPs (Reed et al., 2004), while Cui et al. showed a sensitivity mean of 98.8% for the HRM screening of 2 different HLA SNPs on 2 different instruments (Cui et al., 2013). The HRM method also showed 100% sensitivity, 100% specificity, 0% false positive and negative rate for HLA-B\*51 detection (Imperiali et al., 2015). The HRM method was used to detect the HLA-B\*58:01 allele with a high sensitivity and specificity (93.75% and 88.08%) with a PPV and NPV of 79.65% and 96.59% (Saksit et al., 2017). The HRM method developed for HLA-B\*15:02 screening in Malaysia showed 100% sensitivity and specificity (Rani et al., 2018).

However, the HRM method in this study showed a sensitivity of 0, specificity of 57%, positive predictive value of 0 and negative predictive value of 57%. Sensitivity measures the proportion of positives that are correctly identified, while specificity measures the proportion of negatives that are correctly identified in a method (Reed et al., 2004). The latter reflects the overall results as the positive samples identified by HRM were not the exact same samples identified as HLA-B\*58:01 positive by the Sanger sequencing method, giving a sensitivity of 0. However, both the HRM method and Sanger sequencing identified five common HLA-B\*58:01 samples, contributing to the 57% specificity. Hence, the PPV value is 0 too, followed by a NPV value of 57%. All these values could be improved by standardizing all limiting factors seen in this study. There are numerous factors contributing to these low results seen. The biggest factors contributing to the 0% sensitivity and PPV value are the heterozygosity of the positive control and the high polymorphism of the HLA-B\*58:01 allele. There was a small number of ADRs/SCARs samples (n=28) collected over 3 years due to slow sample collection. Moreover, there was no standard ADRs/SCARs classification method in the two hospitals, causing an underestimation of the true ADRs/SCARs. The biggest contributing factor here was the primer specificity and the HLA-B\*58:01 high polymorphism seen. This renders exon amplification not feasible in this research. Future studies should focus on SNPs amplification. The lower resolution power of

the Eco Real-Time Illumina machine used was also a limiting factor. The low resolution was further explained when comparing Rani et al's study to this study, as the latter was used as a reference in this research. Rani et al. targeted exons 2 and 3, used the EvaGreen dye in HRM and the Rotor-Gene 6000 HRM System for their method (Corbett Life Science, USA) (Rani et al., 2018). All these factors were similar, except for the primers and type of instrument used. Hence, their specific HRM system displayed better scanning sensitivity and specificity, compared to the Illumina machine which has a wider range of application (Herrmann et al., 2006). Moreover, the Illumina machine experienced several breakdowns over the years and no proper maintenance. Statistical analysis was further used to assess the overall HRM results when compared to the Sanger sequencing results obtained. The Sanger sequencing results were used to assess if there was an association between the HLA-B\*58:01 allele's distribution in the healthy volunteer cohort (n=145) and the gout cohort (n=145), via Fisher's exact test. There was no significant association found between these 2 cohorts in terms of HLA-B\*58:01's distribution. The full 290 cohort was used here, taking confirmed Sanger sequencing results only, hence 9 positive for gout cohort and 4 positive for volunteer cohort. This was used instead of HRM results as the latter is too ambiguous.

The association between HLA-B\*58:01's presence and the different ethnic groups in the gout cohort were also assessed via Fisher's exact test and showed a p value of 0.737( $p>0.05$ ). Thus there was no association between presence of HLA-B\*58:01 allele and the different ethnic groups in the gout cohort. This may have been contributed by the small number of positive HLA-B\*58:01 samples (9 positive), obtained from the Sanger sequencing results done on only 28 samples out of the gout cohort. Moreover, there was a big difference for the number of patients in the different ethnic groups in the gout cohort. The association between HLA-B\*58:01's presence and the different ethnic groups in the volunteer cohort were also assessed via Fisher's exact test and showed a p value of 0.181 ( $p>0.05$ ). Hence, there is again no association, due to the small number of positive HLA-B\*58:01 samples (4 positive), obtained from the Sanger sequencing results done on only 14 samples out of the volunteer cohort. In the future, the whole volunteer and gout cohorts must be sent for Sanger sequencing in order to find the true strength of association between ethnicities in a general Malaysian population and the HLA-B\*58:01 allele presence. The link between HLA-B\*58:01's presence and the different types of allopurinol-induced

hypersensitivity reactions seen were assessed via One way ANOVA and showed no significant link with a p value of 0.905. This was probably caused by the poor classification of ADRs/SCARs in the two hospitals, with many different defining terms and the presence of only 9 HLA-B\*58:01 positive samples.

In terms of cost-effectiveness, a proper cost for the HRM method was quite ambiguous in our laboratory, as only the cost of reagents and consumables used could be calculated and came to around USD \$1.50 per sample. Other studies also support the low cost of HRM of around USD \$1.50 per 20 $\mu$ L sample, similar to this study (Reed et al., 2007; Bai et al., 2012). However, a more accurate cost would be that of Rani et al. calculated in a hospital setting and its affiliated laboratory, at <\$16.40 USD/test (Rani et al., 2018). Hence, this cost can be applied in the future to this HRM method, if it was to be implemented in the HUKM hospital as all the factors are similar to Rani et al's method. Moreover, this HRM method cost was less compared to USD \$25.15 and USD \$25.70 for multiplex-PCR and DNA sequencing respectively (Then et al., 2013; Rani et al., 2018). Rapid turnaround time (TAT) is crucial for a diagnostic laboratory which offers the HLA genotyping as a molecular genetic testing. The HRM screening preparation takes about 30-45mins depending on the handler's experience and 1 hour 30 minutes for the run and generation of results. The TAT for HRM method is faster compared to multiplex-PCR and DNA sequencing (Then et al., 2013; Rani et al., 2018). 15 samples can be screened at once and results generated in 2 hours, showing that it is rapid and better than any current HLA screening methods (Fortini et al., 2007; Polakova et al., 2008). It is also cost-effective, less labour intensive and error-prone compared to other methods (Lundgren et al., 2012).

Several countries show reluctance to implement HLA-B\*58:01 screening due to a negative cost-effectiveness analysis, namely, in Malaysia (Chong et al., 2018), Singapore (Dong et al., 2015), Korea (Park et al., 2015), United Kingdom (Plumpton et al., 2017) and United States (Jutkowitz et al., 2017). All these aforementioned studies, however, stated the same limitations, showing us how to tackle PGx implementation in the future. High genotyping cost, limited availability choices of drugs for chronic gout treatment and lower positive predictive value for SJS/TEN were the factors that reduced the cost-effectiveness of HLA-B\*58:01 genotyping (Dong et al., 2015; Park et al., 2015; Plumpton et al., 2017; Jutkowitz et al., 2017; Chong et al., 2018). Moreover, methods of HLA-B\*58:01 screening were all time-

consuming and were not available in routine clinical practice (Chong et al., 2018). However, taking Rani et al's study example as an example in Malaysia, all these limitations were overcome, as it has been already implemented in HUKM hospital and its affiliated laboratory UMBI. Rani et al. have been screening for the HLA-B\*15:02 allele for around two years in epileptic patients and have shown a significant reduction in the number of SCARs patients recorded in HUKM (Rani et al., 2018). This shows that the implementation of pharmacogenetic screening is hospital-specific, based on the availability of laboratories and their willingness to participate in the PGx field. Thus, HLA-B\*58:01 detection will soon be a reality in HUKM and Hospital Putrajaya, contributed by the willingness of doctors there to provide better treatments for their patients.

### **5.2.7 Conclusion**

Hence, to conclude the second section of discussion, the HRM method was developed based on literature, but did not show a good sensitivity, specificity, PPV and NPV due to the presence of numerous limiting factors. The primers designed were not specific enough, the positive control turned out to be a heterozygous sample, and the high level of polymorphism recorded in the HLA-B\*58:01 region only exacerbated this. Ambiguity in HLA typing by HRM and Sanger sequencing is linked to the high polymorphism observed in the HLA-B\*58:01 allele. Exon amplification was not specific enough and several HLA alleles were seen after sanger sequencing. The HLA-B\*58:01 allele was still linked to gout patients to a certain level and the theory of multiple pharmacogenetic markers was brought up by Sanger sequencing. The HRM method also showed no significant difference when compared to Sanger sequencing via statistical analysis.

## 5.3 Core Findings: Validation of multiple pharmacogenetic markers presence by NGS

### 5.3.1 NGS foundation

The human leukocyte antigen (HLA) region on the chromosome 6p21 comprising six classical polymorphic HLA genes and at least 132 protein coding genes plays important roles in regulation of immune system as well as fundamental molecular and cellular processes (Shiina et al., 2009). The completion of a continuous 3.6 Mb of HLA genomic sequence together with annotation of 224 genes, was first reported by The MHC Sequencing Consortium in 1999 (The MHC Sequencing Consortium, 1999). In addition, the MHC Haplotype Project conducted by the Sanger Institute provided genomic sequences and gene annotation of eight different HLA haplotypes, which were registered in the UCSC hg19 and NCBI GRCh37 reference assembly (Stewart et al., 2004; Traherne et al., 2006). This 3.6 Mb segment occupies only 0.13% of the human genome but is associated with more than 100 different diseases, mostly autoimmune diseases such as type I diabetes, rheumatoid arthritis, psoriasis, and atopic asthma. Recently, HLA-B genes attracted special attentions, because specific alleles of HLA genes are strongly associated with drug hypersensitivity induced by specific drugs. For better understanding of disease causality and drug hypersensitivity, NGS has been used to completely decode these HLA-B genes (Hosomichi et al., 2013).

The NGS method was initially included in this study to further validate the HRM method as the Sanger sequencing method leaves gaps, which cannot be resolved in the results. However, the slow sample collection has greatly interfered with the NGS timeline and not enough allopurinol-induced ADRs/SCARs samples were collected on time for the NGS run in 2017. Only 5 samples with SCARs/ADRs were available for the NGS run, along with the positive control used for HRM screening methods. The initial plan was to include 48 gout samples with ADRs/SCARs, along with 48 healthy volunteers, matching in age, gender and ethnicity. This would enable us to delve deeper into the link between the HLA-B\*58:01 allele in Malaysia and the different ethnicities. The extent of HLA-B\*58:01's involvement in a wide range of different ADRs/SCARs would be studied in depth. NGS would thus allow us to confirm if the HLA-B\*58:01 allele has a significant role in the multi-ethnic Malaysian

population. The multiple pharmacogenetic markers theory would also be explored via NGS as it covers the full HLA-B region of around 4000bp. The association of any other strong HLA-B alleles will thus be seen and compared to the HLA-B\*58:01 allele. However, the reagents had to be used due to their close expiry date and all the machines and equipment were booked in UMBI laboratory in HUKM hospital.

Furthermore, as shown by the Sanger sequencing raw data (Appendix H), the full 4000bp wasn't sequenced properly due to high polymorphism and presence of secondary and tertiary structures which makes the sequence difficult to decipher by Sanger sequencing (Ronaghi et al., 1999). Sanger sequencing is known to have disadvantages when faced with the complex HLA-B alleles, such as incomplete sequencing, limitations in resolution, ambiguous HLA alleles called by typing and haplotype phase issues (Erich, 2012; Shiina et al., 2012). The presence of SNPs and indels which are closely associated and found in HLA-B alleles, further made the Sanger sequencing process difficult (Longman-Jacobsen et al., 2003). However, the HLA-B primers used were used and validated, hence showing its accuracy and specificity (Hosomichi et al., 2013). Hosomichi et al. used long-range PCR to amplify six HLA genes (*HLA-A*, *-C*, *-B*, *DRB1*, *-DQB1*, and *-DPB1*) followed by indexed multiplex NGS with the Illumina MiSeq sequencer. Authors, assigned up to 8-digit HLA allele numbers to their results, regardless of whether the alleles are rare or novel, resulting in high resolution HLA typing and new allele detection (Hosomichi et al., 2013). Hence, by using the same primers and platform used for NGS in this study, the variables and errors are minimized. The search for NGS papers in Malaysia yielded no specific NGS papers, but other HLA related studies with different typing method. Hence, this is the first NGS study, to our knowledge, done on a Malaysian cohort to further investigate the role of HLA-B alleles.

### **5.3.2 NGS method evaluation**

The first set of crucial results for the NGS run was the sequencing performance on the MiSeq system, that is the data quality and total data output and yield. Numerous factors affect these results, such as proper preparation, quantification, denaturation and dilution of libraries, along with clonal clusters' density, loading concentration and sequencing run parameters (Ravi et al., 2018). All those values were checked and validated to be of good quality by an NGS analysis expert from UMBI and compared to a full set of successful results

published by a similar NGS run (Ravi et al., 2018). The Qscore or Q30 value was used as a quality control and shows if 99.9% of the bases are correct for the NGS done. Most Illumina runs will generate >70-80% Q30 data and this NGS run had a value of 91.79% (shown in Table 4.24), which was considered a very good Q30 value, along with very good quality of data generated. Appendix H shows the very good Fast Q scores of 38-40 for all samples and high percentages mapped to the human genome. A total of 623 HLA-B related SNPs hits, all mapped to chromosome 6, were found from this NGS run on the 96 samples. Figure 4.7 shows results of 3 graphs for QC check, where graphs A and B both showed a good clustering of data without any sudden drop or erratic data. Graph C showed the optimal density range with the density box plots close together and a good cluster %PF value of  $83.41\% \pm 5.17$ . These three graphs further match the set of successful MiSeq sequencing run results, along with the similar graphs in the Sequencing Analysis Viewer (SAV) software used (Ravi et al., 2018). Hence all the factors that could possibly affect the results were minimized and a good quality NGS library was obtained, as validated by literature and a bioinformatics/NGS expert in UMBI (Ravi et al., 2018). The MiSeq run took only 12 hours to complete for the 96 samples NGS, compared to the 5 day turnaround of typing for five clinical samples in a clinical setting by another study using the same Illumina MiSeq platform (Wang et al., 2012). Moreover these NGS results were analysed by the specific SNP numbers obtained for each sample in dbSNP in NCBI in order to get a specific DNA sequence (FASTA format), along with the different base change possible. The obtained FASTA sequence was copied into the Nucleotide Similarity Search from NCBI and compared to 2 databases; IMGT/HLA (cds) and IMGT/HLA (genomic). All the possible HLA-B hit combinations were entered and the results were compiled with the HLA alleles with the highest presence in each sample (shown in Table 4.25), hence explaining the numerous HLA hits shown. All the SNPs numbers obtained were compared to the dbSNP database throughout 2017 to 2019, so that new SNPs entered in the database will be counted and compared too. Wang et al. compared their results to IMGT/HLA database in 2011, where of 6,398 cDNA reference sequences for HLA-A, -B, -C, and -DRB1 genes in the IMGT-HLA database, only 375 (5.8%) of them had genomic sequences (Wang et al., 2012). We had the advantage of having more HLA alleles to compare these results to, with a total of 18,500 Class 1 alleles and 7000 Class 2 alleles upto the year 2019 (IMGT/HLA, EMBL-EBI). Other NGS studies focused on exons' amplifications, mostly exons 2-4, and showed a limited number of HLA-B alleles identified

(Bentley et al., 2009; Wang et al., 2012). Although an exon based strategy will resolve many ambiguities, by way of the clonal sequence obtained, it will still produce ambiguities due to the incomplete HLA gene sequencing (Lind et al., 2010). Non-coding regions are also known to have an impact on gene regulation, mRNA splicing (Hosomichi et al., 2013). Hence, results will not portray the full complexity of the introns and exons working together to make functional HLA-B alleles. In this study, we amplified the full HLA-B gene (~4000bp) and had HLA-B introns and HLA-B exons identified, providing a wider range of HLA-B hits.

### 5.3.3 NGS on allopurinol-induced SCARs samples from gout cohort.

The 6 gout samples sent for NGS in 2017, were the only ones available with the more severe ADRs/SCARs. The NGS results for the 6 gout samples show similar results to the sanger sequencing results, where the SNPs identified on samples showed numerous genetic markers and not just the expected HLA-B\*58:01. Only a few South East Asian NGS studies were found, namely, on the Polynesian people (Ton et al., 2018) and Vietnamese people (Do et al., 2020), while a Chinese Mongolian population (Wang et al., 2016) was investigated by polymerase chain reaction–sequence-based typing (PCR-SBT) and cloning. These studies all show ancestral links and the presence of common ethnic groups compared to the Malaysian population and was thus used as comparison.

The HLA-B\*56 allele, with its numerous variants, had the highest hit number from 10 different SNPs in the gout patients. Presence of numerous HLA-B\*56 alleles, such as HLA-B\*56:01:01 and HLA-B\*56:02:01, was shown in another study done on Polynesian people (Ton et al., 2018). This is further validated by the fact that here is a clear ancestral link proven between Malaysians and Polynesians, hence showing the appearance of a possible new pharmacogenetic marker, that is, HLA-B\*56:01 (Chambers & Edinur, 2013). Half of the samples (HLAG12, HLAG19, HLAG28) showed presence of the HLA-B\*58:01 alleles and multiple other HLA-B\*58 alleles, including HLA-B\*58:02. Three samples (HLAG12, HLAG19, HLAG28) had HLA-B\*58:01, HLA-B\*57:01 and HLA-B\*40 alleles, showing presence of two previously verified pharmacogenetic markers together. Heterozygous alleles identified were B\*58:01/B\*57:01, B\*58:01/B\*40, B\*48/B\*48:01 and B\*58/39. The Polynesian study showed HLA-B\*40:01:01 (28.95%) as the most common allele, supporting these results. HLA-B\*57:01 was also shown in Polynesian people, but interestingly, there were no HLA-B\*58:01 observed (Ton et al., 2018). The Chinese Mongolian study showed presence of similar HLA-B alleles, namely, B\*40:01, B\*40:02, B\*40:06, B\*56:01, B\*56:03, B\*57:01 and B\*58:01 (Wang et al., 2016). The Vietnamese study showed presence of HLA-B\*58:01:01 and HLA-B\*40:01:02 only (Do et al., 2020). Thus, even with a small samples size of 6 gout samples, the HLA-B\*58:01's link to gout in Malaysians was further cemented, along with the presence of other common alleles seen in South East Asian populations too (HLA-B\*57:01, HLA-B\*56 alleles and HLA-B\*40 alleles). Unfortunately, the positive control showed few

SNPs hits compared to the other gout samples, which was attributed to its minimal amount of DNA, stored over more than 5 years. The patient could not be tracked and contacted as the latter formed part the previous study too long ago. Moreover, results were not completely decoded due to a high number of novel SNPs identified for each sample and no match were found in the global NCBI dbSNP database or IMGT/HLA database. Thus, further studies are needed to link these novel SNPs to HLA-B alleles.

When this set of 6 NGS results were compared with their respective Sanger sequencing results, only one sample, HLAG12, showed the presence of common alleles in both methods. The HLA-B\*58:01 allele and other variants of the HLA-B\*58 alleles, like HLA-B\*58:02 were commonly present in the sample. Even more intriguing, is the fact that sample HLAG12 had only allopurinol-induced vasculitis rash as ADRs, confirming the link of the HLA-B\*58:01 allele and milder cADRs in Malaysian patients. Combining the Sanger sequencing results and NGS results, further shows the presence of HLA-B\*58:01 confirmed by one of these methods, in the positive control and all the other 5 samples with different allopurinol-induced ADRs/SCARs. Furthermore, sample HLAG28 showed the presence of the HLA-B\*58:01 allele both by NGS and HRM, showing the HRM's potential as a good screening method. However, the HRM method should focus on SNPs detection in order to accurately perform HLA typing. The various SNPs found from NGS should be used in further studies.

### 5.3.4 Single Nucleotide Polymorphisms' role in HLA-B alleles

HLA-B\*58:01 was found in three single SNPs (rs11423052, rs151341211 and rs9279154) and no a literature search revealed no other studies with these SNPs. Thus, these SNPs might be Malaysian specific SNPs linked to the HLA-B\*58:01 allele, which can be used in the future for studies. Saksit et al., 2017 showed the importance and association between three SNPs in chromosome 6 and allopurinol-induced SCARs in a Thai population. They showed that three SNPs (rs9263726, rs2734583, and rs3099844) were significantly associated with allopurinol-induced SCARs but with a lower degree of association when compared with the HLA-B\*58:01 allele. The sensitivity, specificity, PPV, and NPV of these SNPs were comparable to those of the HLA-B\*58:01 allele (Saskit et al., 2017). These three SNPs were not found in this NGS run, again showing the proof of Malaysian specific SNPs found.

### 5.3.5 Conclusion

Hence, to conclude the third section of discussion, the NGS method could not validate the HRM method and sanger sequencing results, due to limited sample size. NGS should be run again on a proper cohort with a set number of allopurinol-induced SCARs patients. The target region could also be decreased for future studies, from the whole HLA-B region to a few exons, in order to obtain more specific results.

## 5.4 Limitations of study

Numerous unexpected limitations and hurdles were experienced over the three years of this research. The most significant limitations experienced were the slow sample collection, lack of proper SCARs classification in hospitals, primer specificity, HLA-B\*58:01 high polymorphism, low resolution of the Illumina platform used, time constraints and ending of the grant funds 2 years after the start of the study.

Slow and unexpected sample collection in HUKM was observed during the first year of collection and prompted us to immediately apply for sample collection in a second hospital in Putrajaya. The ethics approval took another year and collection started in Hospital Putrajaya in 2017. Patients were also recycled and came to the hospital every 3-6 months, again limiting samples collection. Numerous obstacles were experienced in HUKM, starting with the manual screening and reading of thick paper-based medical record files for each patients. Patients had to be screened and selected manually daily, followed by their transfer to the doctor affiliated with our study. There was a limit for the number of patients the designated doctor would take and other doctors in the department refused to help with the collection due to their busy schedules. The doctor in charge encountered several problems in blood taking as he has not been practicing blood taking for a long time. This led to the loss of several patients, as no blood could be withdrawn even after three jabs. The designated blood taking place in the rheumatology department was not willing to help and participate in our study. HUKM had a more reclusive and strict work environment compared to Hospital Putrajaya (HP). Sample collection in HP was smoother and faster as their whole system was computerized and the whole department took part in blood sample collection. Medical records of gout patients were already categorized and easily accessed and screened for allopurinol use daily. No patients were lost in HP as the nurses were experienced and collected blood swiftly on the first jab itself. The whole collection process was more organized in HP and around 65 blood samples were collected in 6 months, compared to 80 patients from HUKM collected throughout 3 years. Sample collection was an unexpected variable, completely out of our control, but we managed to collect enough samples due to timely the addition of a second hospital.

The overall slow samples collection limited the access to more allopurinol-induced ADRs/SCARs cases and only 28 samples could be collected over 3 years. Even though the original calculated sample size of 145 gout patients and 145 healthy volunteers were reached, it was not significant enough in the end due to the novelty of this research in Malaysia. Moreover, the original plan of including all the allopurinol-induced ADRs/SCARs cases in the NGS run couldn't be done as only 5 ADRs/SCARs cases were available in 2017, in time for the NGS run. This would have allowed a stronger and more significant research on HLA-B\*58:01's role in gout patients. Moreover, more HLA-B alleles could have been identified to further validate the multiple pharmacogenetic markers concept.

The second significant limitation was the lack of proper SCARs classification seen in both hospitals. This led to a lot of cases being labelled haphazardly by numerous doctors and led to an underestimation of ADRs/SCARs impact in this whole study (Bastuji-Garin et al., 2000). Recording other data such as allergic reactions, other medications taken by patients and comorbidities was again very tedious in HUKM. A form needs to be filled and submitted to the archive section, followed by around 2 weeks for record withdrawal. A maximum of 20 records can be requested in one go and all reading must be done in the archive section according to HUKM's rules. Multiple medical records had to be requested more than three times due to the archive officers not finding them. Reading all the medical records comprising of around 30 pages to more than 100 pages took quite a while and the poor readability of the several doctors' handwriting didn't help. Accessing medical records was faster in HP as everything was computerized in one location and took only one day.

Another significant limitation was the ending of the grant fund 2 years after the start of the research. All the NGS reagents had to be bought and ordered quicker and thus we couldn't wait for more allopurinol-induced ADRs/SCARs cases and had to run it immediately in the second year of the research itself. More Sanger sequencing couldn't be performed on a bigger sample size due to limited time and limited funds. The first batch of 28 ADRs/SCARs samples couldn't be matched with healthy volunteers due to the time constraint and only 14 samples were sent for Sanger sequencing in the end. This prevents proper matching of both cohorts in terms of age, gender and ethnicity. The fourth limitation was that the only one positive control was available due to the novelty of this study in Malaysia in 2014 to 2015. Only one HLA-B allele was known as the HLA-B\*58:01 allele for the positive control,

with the other allele remaining unknown. Moreover, the patient couldn't be contacted for more blood, in order to get more DNA for analysis. We had to clone the positive control and use this single control throughout the study. The heterozygous nature of the positive control was discovered only after a year when the HRM method started. Thus, we had no choice but to continue using this positive control for the full HRM screening as slow sample collection also prevented us from collecting a sample with SCARs for a long time. Hence, the positive control's heterozygous nature further increased variations in the HRM screening method we developed (Gundry et al., 2003; Montgomery et al., 2007).

Primer specificity and high polymorphism in HLA-B\*58:01 were significant limitations which were discovered only in the middle to end of this research. By using the common method of exon 2 and 3 amplification in literature, this research could not attain the same results as other studies due to several limitations. The novelty of the HLA-B\*58:01 study in 2014 to 2015 when this research started also acted as a limiting factor, where benchmarks were not available. Both exon 2 and 3 were sequenced, however gaps and ambiguity were still seen in sanger sequencing, due to the high polymorphism in the HLA-B\*58:01 allele.

The fifth limitation was the lower resolution of the Illumina platform used for the HRM method (Herrmann et al., 2006). This was only noticed halfway into the research, when no samples were matching the positive control's HRM melt curve at all and when Sanger sequencing was performed when enough samples were collected. Moreover, this was the only real-time PCR machine with the HRM function in the University of Nottingham and another platform couldn't be accessed to improve the results' resolution. Hence, all the aforementioned factors contributed to the low sensitivity and specificity of the HRM method (Herrmann et al., 2006). The method's full potential couldn't be revealed when compared to literature.

# CHAPTER 6: FINAL CONCLUSION AND FUTURE WORK

## 6.1 Final conclusion

This whole research focused on the novel study of HLA-B\*58:01 allele's role in allopurinol-induced SCARs in Malaysia, by developing a specific HRM method to detect the allele in gout patients and healthy volunteers. The Sanger sequencing and NGS methods were used to validate the HRM method and further delve into the complexity of the HLA-B alleles' role in Malaysians. The newer theory of the presence of numerous HLA-B alleles as pharmacogenetic markers compared to the use of a single allele in HLA-drug-SCAR associations was investigated in this research. Therefore, this thesis and research as a whole tackled all the important principles of the pharmacogenomics field in a Malaysian setting.

A hypothesis of this research was to show that the HRM method is accurate, sensitive and specific enough to detect the the HLA-B\*58:01 allele. This was shown only to a small extent, due to the presence of significant limiting factors. The foundation of the HRM method, that is the primer design, PCR and positive control cloning, was developed to the best of our abilities. However, the specificity of the designed primers were still lacking, due to the high level of polymorphisms seen in the HLA-B gene. Specific primers for exon 2 and 3 of the HLA-B\*58:01 allele were developed and checked by Sanger sequencing. Sanger sequencing showed gaps in the results, hence giving incomplete results. The lacking primer specificity was again proven by the large number of HLA-A, HLA-B and HLA-C alleles seen in sanger sequencing results. The HRM method was also fully developed based on literature, but the method's strength could not be fully exploited. The limiting factors affecting the latter were the lack in primer specificity, highly polymorphic HLA-B\*58:01 region, low resolution of the Eco Real-Time Illumina platform, heterozygous nature of the positive control, slow sample collection and the subsequent low number of allopurinol-induced ADRs/SCARs cases. The HRM method showed a sensitivity of 0%, specificity of 57%, PPV of 0% and NPV of 57%. One main contributing factor to the 0% sensitivity and PPV was the heterozygous nature of the positive control. By using only close curves within the -20 to 20 threshold, HRM singled out 8 possible positive results, compared to the 9 HLA-B\*58:01 positive samples detected by Sanger sequencing. Hence, the HRM method shows great potential and can be further improved in the future, with the significant findings mentioned below.

The first hypothesis of this research stated that the HLA-B\*58:01 allele is strongly linked to allopurinol-induced SCARs in Malaysian gout patients. HLA-B\*58:01 was seen as a strong risk factor and genetic marker in a wide range of allopurinol-induced hypersensitivities, not limiting it to SCARs, thus validating the first and third hypothesis. HLA-B\*58:01:01 allele presence was confirmed in allopurinol-induced SJS, itchininess, vasculitis rash, exfoliative dermatitis and DRESS syndrome. Sanger sequencing showed the presence of HLA-B\*58:02 and HLA-B\*35:01:01, both with a frequency of 32.1%, in 100% of all the HLA-B\*58:01 positive ADRs/SCARs samples. Moreover, healthy volunteer samples showed high frequencies of 28.5% for both the HLA-B\*58:01:01 and HLA-B\*35:01:01 alleles. The B\*58:01:01 and HLA-B\*35:01:01 alleles are thus identified as possible pharmacogenetic markers, linked to allopurinol-induced ADRs/SCARs in gout patients and in healthy volunteers in Malaysia. Moreover, HLA-B\*58:02 shows the possibility of being a susceptible allele too, which in turn increased the chances of allopurinol-induced SCARs. Due to the novelty of this HLA-B\*58:01 PGx study in Malaysia, a greater sample size will be needed in the future to further validate this. Moreover, healthy volunteer samples showed high frequencies of 28.5% for HLA-B\*15:02, HLA-B\*15:21, and HLA-B\*53:01:01 alleles via Sanger sequencing, showing the presence of more pharmacogenetic markers in Malaysians.

The HLA-B\*58:01:01 allele, along with other HLA-B\*58 variants, had frequencies of 50.0% each, in the NGS cohort, showing a greater frequency and involvement than originally expected. However, the small sample size needs to be significantly increased to prove this theory. This was further shown in the Sanger sequencing results with HLA-B\*58:01:01 frequencies of 32.1% in the ADRs/SCARs samples and 28.57% in the healthy volunteer samples. Matching the Sanger sequencing results to the NGS results further revealed common HLA-B alleles with high frequencies (17.9% to 36.5%), namely, HLA-B\*58:01:01, HLA-B\*58 variants, HLA-B\*15 variants and HLA-B\*44 variants. Hence, this shows the possible presence of multiple pharmacogenetic markers again, by both methods.

Three novel SNPs were identified linked to HLA-B\*58:01 via NGS (rs11423052, rs151341211 and rs9279154) and they showed high possibilities of being Malaysian specific due to their absence from literature. These three SNPs can be used in the future to detect the HLA-B\*58:01 allele in Malaysians.

In conclusion, all the hypotheses set at the start of this research were investigated to the best of our capabilities, with all the available resources and limitations experienced. All the specific objectives and aims set initially were also met to a certain extent throughout the project. A new theory was investigated throughout this research. Many years ago, Freud defined neurosis as 'the inability to tolerate ambiguity', and it seems that this is the best principle to adhere to for the future of the HLA-omics field.

## 6.2 Future work

Future studies should consider hospitals with a fully computerized system, the number of doctors willing to participate in the study and the availability of an open designated blood taking place. A nationwide study should be carried out in Malaysia, with all hospitals reporting their ADRs/SCARs along with the offending drugs. Proper SCARs classification must be implemented in all Malaysian hospitals (SCORETEN scale), along with the propagation of higher awareness of the importance of such cases. The creation of a pharmacogenomic passport, as previously done in the Netherlands (Van der Wouden et al., 2019), will then be possible, where patients have their full HLA genes listed in a panel for pre-emptive pharmacogenetic testing. Numerous controls must be found and included in the HRM screening, namely a sample homozygous for the HLA-B\*58:01 allele and one heterozygous for the allele, thus allowing easier and more accurate HLA-B\*58:01 detection (Montgomery et al., 2007). More than one HRM platform must be used to test for the presence of the HLA-B\*58:01 allele or a HRM specific machine must be used, as machine resolution is a big influencing factor on results (Herrmann et al., 2006). This will help find the true potential of the HRM method and show a high sensitivity and specificity as shown in literature.

In the future, the whole gout cohort sample must be sent for Sanger sequencing in order to find the true strength of association between ethnicities in a Malaysian gout population and the HLA-B\*58:01 allele. More studies are needed to check which HLA-B alleles are involved in double allergy to allopurinol and febuxostat. Future screening systems need to focus more on amplification of SNPs instead of exon 2 and 3 for more specific results (Saskit et al., 2017). The three novel HLA-B\*58:01 linked SNPs (rs11423052, rs151341211 and rs9279154) discovered in this study can be investigated in the Malaysian population. This SNP-based focus will also help to upgrade the HRM method for HLA-B alleles detection. HRM method needs to be tested on multiple platforms and on more HRM-specific platforms for better resolution (Herrmann et al., 2006; He et al., 2016).

The role of HLA-B\*35:01:01 and HLA-B\*58:02 must be confirmed in Malaysian gout patients as possible pharmacogenetic markers in ADRs/SCARs. Moreover, protective HLA alleles must also be identified in Malaysians. The peptide-binding altered repertoire mechanism needs to be investigated, as it contributes to the presence of protective HLA alleles (Mori et al.,

2014). HLA molecules with a more restricted peptide-binding repertoire, such as HLA-B27 and HLA-B57, and in this case, HLA-B35:05, are more likely to be protective. The crystallography method and molecular-docking studies need to be used to see which residues are making the whole difference. The newly found altered TCR-repertoire mechanism of action must also be investigated in all HLA-B alleles. Linkage disequilibrium between all those identified HLA-B genetic markers must be investigated globally (Chang et al., 2020). The NGS method can be used for haplotype calculation, docking modelling of the alleles, 3D protein docking, statistical analysis, linkage disequilibrium present and homozygotes and heterozygotes identification (Chang et al., 2020). HLA-B alleles and their link to diseases are shown to depend on numerous changing factors, such as, mutations, evolution, selective advantage, protective/susceptible alleles and geographic location, which all need to be investigated (Chang et al., 2020).

The complexity of HLA alleles has been simplified to a search for associations between one HLA allele and one drug linked to SCARs in a specific disease. However, this hunt for stand-alone effective genetic markers shifted the focus from the big picture for years, that is, the complexity and interconnection of the whole HLA system in the human immune response. Now that we have a list of multiple, population-specific pharmacogenetic markers, it is time to combine them all together and investigate their roles as an interconnected, fool-proof barrier in the immune system. This new theory of the presence of multiple pharmacogenetic markers in populations need to be tested out via reviews, NGS studies and meta-analyses in the future. The strength of association of the HLA-B\*58:01, HLA-B\*15:02, HLA-B\*57:01 and numerous other validated markers must be reassessed globally (Wu et al., 2016; Then & Raymond 2019; Chang et al., 2020). The strength of association of numerous variants need to be assessed and not only the first variant, as the latter are able to bind to the drug molecules directly, or grooves not participating in antigen-presentations (Jaruthamsophon et al., 2017; Pompeu et al., 2012).

**“Our imagination is the only limit to what we can hope to have in the future.”**

**-Charles Kettering**

## REFERENCES

- A Ministry of Frontier Ventures, *The Joshua Project, 2001*, Available online: <https://joshuaproject.net/about/website>. Accessed, 07/12/2021.
- Complete sequence and gene map of a human major histocompatibility complex. (1999) The MHC sequencing consortium. *Nature*, 401(6756), 921-3.
- Abbott, R. D., Brand, F. N., Kannel, W. B. & Castelli, W. P. (1988) Gout and coronary heart disease: the Framingham Study. *J Clin Epidemiol*, 41(3), 237-42.
- Alfirevic, A., Gonzalez-Galarza, F., Bell, C., Martinsson, K., Platt, V., Bretland, G., Evely, J., Lichtenfels, M., Cederbrant, K. & French, N. (2012) In silicoanalysis of HLA associations with drug-induced liver injury: use of a HLA-genotyped DNA archive from healthy volunteers. *Genome medicine*, 4(6), 51.
- Alqasem, A., Alanazi, H., Aloufi, F., Alanazi, A. & Hajeer, A. H. (2019) A novel HLA-B allele, HLA-B\*44:03:01:19, identified by next-generation sequencing in a Saudi individual. *Hla*, 94(4), 381-382.
- Althaf, M. M., El Kossi, M., Jin, J. K., Sharma, A. & Halawa, A. M. (2017) Human leukocyte antigen typing and crossmatch: A comprehensive review. *World journal of transplantation*, 7(6), 339.
- Anton, F. M., Garcia Puig, J., Ramos, T., Gonzalez, P. & Ordas, J. (1986) Sex differences in uric acid metabolism in adults: evidence for a lack of influence of estradiol-17 beta (E2) on the renal handling of urate. *Metabolism*, 35(4), 343-8.
- Atdjian, M. & Fernandez-Madrid, F. (1981) Coexistence of chronic tophaceous gout and rheumatoid arthritis. *The Journal of rheumatology*, 8(6), 989-992.
- Bai, M., Wang, C., Yin, H., Tian, Y. & Li, J. (2012) Evaluation of different reaction systems for HRM analysis in apple. *Bioscience Methods*, 3.
- Bain, B. J. (1996) Ethnic and sex differences in the total and differential white cell count and platelet count. *Journal of clinical pathology*, 49(8), 664-666.
- Bardin, T., Chalès, G., Pascart, T., Flipo, R. M., Korng Ea, H., Roujeau, J. C., Delayen, A. & Clerson, P. (2016) Risk of cutaneous adverse events with febuxostat treatment in patients with skin reaction to allopurinol. A retrospective, hospital-based study of 101 patients with consecutive allopurinol and febuxostat treatment. *Joint Bone Spine*, 83(3), 314-7.
- Bardin, T., Flipo, R.-M., Richette, P. & Clerson, P. (2014) Gouty Patients with History of Adverse Reaction to Allopurinol Are Not at Higher Risk of Reaction to Febuxostat.: 164. *Arthritis & Rheumatology*, 66.

Bastuji-Garin, S., Fouchard, N., Bertocchi, M., Roujeau, J. C., Revuz, J. & Wolkenstein, P. (2000) SCORTEN: a severity-of-illness score for toxic epidermal necrolysis. *J Invest Dermatol*, 115(2), 149-53.

Bedir, A., Topbas, M., Tanyeri, F., Alvur, M. & Arik, N. (2003) Leptin might be a regulator of serum uric acid concentrations in humans. *Japanese heart journal*, 44(4), 527-536.

Bentley, G., Higuchi, R., Hoglund, B., Goodridge, D., Sayer, D., Trachtenberg, E. A. & Erlich, H. A. (2009) High-resolution, high-throughput HLA genotyping by next-generation sequencing. *Tissue Antigens*, 74(5), 393-403.

Bharadwaj, M., Illing, P., Theodosis, A., Purcell, A. W., Rossjohn, J. & McCluskey, J. (2012) Drug hypersensitivity and human leukocyte antigens of the major histocompatibility complex. *Annual review of pharmacology and toxicology*, 52, 401-431.

Bidwell, J. (1994) Advances in DNA-based HLA-typing methods. *Immunology today*, 15(7), 303-307.

Bon, M. A. M., van Oeveren-Dybicz, A. & van den Bergh, F. A. J. T. M. (2000) Genotyping of HLA-B27 by Real-Time PCR without Hybridization Probes. *Clinical Chemistry*, 46(7), 1000-1002.

Buchanan, W. W., Klinenberg, J. R. & Seegmiller, J. (1965) The inflammatory response to injected microcrystalline monosodium urate in normal, hyperuricemic, gouty and uremic subjects. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 8(3), 361-367.

Bustin, S. A., Benes, V., Garson, J. A., Hellems, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J. & Wittwer, C. T. (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem*, 55(4), 611-22.

Bădulescu, M., Macovei, L. & Rezuş, E. (2014) Acute gout attack with normal serum uric acid levels. *Rev Med Chir Soc Med Nat Iasi*, 118, 942-945.

Caboux, E., Lallemand, C., Ferro, G., Hémon, B., Mendy, M., Biessy, C., Sims, M., Wareham, N., Britten, A., Boland, A., Hutchinson, A., Siddiq, A., Vineis, P., Riboli, E., Romieu, I., Rinaldi, S., Gunter, M. J., Peeters, P. H. M., van der Schouw, Y. T., Travis, R., Bueno-de-Mesquita, H. B., Canzian, F., Sánchez, M.-J., Skeie, G., Olsen, K. S., Lund, E., Bilbao, R., Sala, N., Barricarte, A., Palli, D., Navarro, C., Panico, S., Redondo, M. L., Polidoro, S., Dossus, L., Boutron-Ruault, M. C., Clavel-Chapelon, F., Trichopoulou, A., Trichopoulos, D., Lagiou, P., Boeing, H., Fisher, E., Tumino, R., Agnoli, C. & Hainaut, P. (2012) Sources of pre-analytical variations in yield of DNA extracted from blood samples: analysis of 50,000 DNA samples in EPIC. *PloS one*, 7(7), e39821-e39821.

Cakir, N., Pamuk, O. N., Dervis, E., Imeryuz, N., Uslu, H., Benian, O., Elelci, E., Erdem, G., Sarvan, F. O. & Senocak, M. (2012) The prevalences of some rheumatic diseases in western Turkey: Havs study. *Rheumatol Int*, 32(4), 895-908.

Campion, E. W., Glynn, R. J. & DeLabry, L. O. (1987a) Asymptomatic hyperuricemia. Risks and consequences in the Normative Aging Study. *Am J Med*, 82(3), 421-6.

Cao, Z.-h., Wei, Z.-y., Zhu, Q.-y., Zhang, J.-y., Yang, L., Qin, S.-y., Shao, L.-y., Zhang, Y.-t., Xuan, J.-k. & Li, Q.-l. (2012a) HLA-B\* 58: 01 allele is associated with augmented risk for both mild and severe cutaneous adverse reactions induced by allopurinol in Han Chinese. *Pharmacogenomics*, 13(10), 1193-1201.

Capule, F., Tragulpiankit, P., Mahasirimongkol, S., Jittikoon, J., Wichukchinda, N., Theresa Alentajan-Aleta, L., James Barit, J. V., Casanova-Gutierrez, J., Cabral-Lim, L., Baltazar Reyes, J. P., Roa, F., Salonga, R., San Gabriel, K. F. & Lynn Silao, C. (2020) Association of carbamazepine-induced Stevens-Johnson syndrome/toxic epidermal necrolysis with the HLA-B75 serotype or HLA-B\*15:21 allele in Filipino patients. *Pharmacogenomics J*, 20(3), 533-541.

Carrington, M., Nelson, G. W., Martin, M. P., Kissner, T., Vlahov, D., Goedert, J. J., Kaslow, R., Buchbinder, S., Hoots, K. & O'Brien, S. J. (1999) HLA and HIV-1: heterozygote advantage and B\*35-Cw\*04 disadvantage. *Science*, 283(5408), 1748-52.

Carstens, J., Wendelboe, P., Sjøgaard, H. & Thestrup-Pedersen, K. (1994) Toxic epidermal necrolysis and erythema multiforme following therapy with terbinafine. *Acta Derm Venereol*, 74(5), 391-2.

Chambers, G. & Edinur, H. (2013) Genetic relationships between Malays and Maori. *The Changing Values of Malays, Maori and Pacific Islander*, 6-37.

Chang, C.-J., Chen, C.-B., Hung, S.-I., Ji, C. & Chung, W.-H. (2020) Pharmacogenetic Testing for Prevention of Severe Cutaneous Adverse Drug Reactions. *Frontiers in Pharmacology*, 11, 969.

Chang, S. J., Ko, Y. C., Wang, T. N., Chang, F. T., Cinkotai, F. F. & Chen, C. J. (1997) High prevalence of gout and related risk factors in Taiwan's Aborigines. *J Rheumatol*, 24(7), 1364-9.

Chen, S. Y., Chen, C. L., Shen, M. L. & Kamatani, N. (2003) Trends in the manifestations of gout in Taiwan. *Rheumatology (Oxford)*, 42(12), 1529-33.

Chen, S. Y. & Shen, M. L. (2007) Juvenile gout in Taiwan associated with family history and overweight. *J Rheumatol*, 34(11), 2308-11.

Chen, W. T., Wang, C. W., Lu, C. W., Chen, C. B., Lee, H. E., Hung, S. I., Choon, S. E., Yang, C. H., Liu, M. T., Chen, T. J., Fan, W. L., Su, S. C., Lin, Y. Y., Chang, Y. C. & Chung, W. H. (2018) The Function of HLA-B\*13:01 Involved in the Pathomechanism of Dapsone-Induced Severe Cutaneous Adverse Reactions. *J Invest Dermatol*, 138(7), 1546-1554.

Chen, Z., Zhang, S., Zhang, J., Zhang, Y., Xue, L. & Miao, L. (2015) rs9263726 is a specific genetic marker for allopurinol-induced severe cutaneous adverse reactions in Chinese patients. *Personalized Medicine*, 12(6), 585-592.

- Cheng, L., Xiong, Y., Qin, C. Z., Zhang, W., Chen, X. P., Li, J. & Zhou, H. H. (2015) HLA-B\* 58: 01 is strongly associated with allopurinol-induced severe cutaneous adverse reactions in Han Chinese patients: a multicentre retrospective case–control clinical study. *British Journal of Dermatology*, 173(2), 555-558.
- Chhana, A. & Dalbeth, N. (2015) The gouty tophus: a review. *Current rheumatology reports*, 17(3), 19.
- Chiu, M. L. S., Hu, M., Ng, M. H. L., Yeung, C. K., Chan, J. Y., Chang, M. M., Cheng, S. H., Li, L. & Tomlinson, B. (2012) Association between HLA-B\* 58: 01 allele and severe cutaneous adverse reactions with allopurinol in Han Chinese in Hong Kong. *British Journal of Dermatology*, 167(1), 44-49.
- Cho, S. K., Kim, S., Chung, J.-Y. & Jee, S. H. (2015) Discovery of URAT1 SNPs and association between serum uric acid levels and URAT1. *BMJ open*, 5(11), e009360.
- Chohan, S. (2011) Safety and efficacy of febuxostat treatment in subjects with gout and severe allopurinol adverse reactions. *J Rheumatol*, 38(9), 1957-9.
- Choi, H. K. & Curhan, G. (2005) Gout: epidemiology and lifestyle choices. *Current opinion in rheumatology*, 17(3), 341-345.
- Choi, H. K., De Vera, M. A. & Krishnan, E. (2008) Gout and the risk of type 2 diabetes among men with a high cardiovascular risk profile. *Rheumatology*, 47(10), 1567-1570.
- Choi, H. K. & Ford, E. S. (2007) Prevalence of the metabolic syndrome in individuals with hyperuricemia. *The American journal of medicine*, 120(5), 442-447.
- Choi, H. K., Zhu, Y. & Mount, D. B. (2010) Genetics of gout. *Current opinion in rheumatology*, 22(2), 144-151.
- Chong, H. Y., Lim, Y. H., Prawjaeng, J., Tassaneeyakul, W., Mohamed, Z. & Chaiyakunapruk, N. (2018a) Cost-effectiveness analysis of HLA-B\* 58: 01 genetic testing before initiation of allopurinol therapy to prevent allopurinol-induced Stevens–Johnson syndrome/toxic epidermal necrolysis in a Malaysian population. *Pharmacogenetics and genomics*, 28(2), 56-67.
- Chong, K. S., Devi, N. A., Gan, K. B. & Then, S.-M. (2018b) Design and development of polymerase chain reaction thermal cycler using proportional-integral temperature controller. *Malaysian Journal of Fundamental and Applied Sciences*, 14(2), 213-218.
- Choo, K. J. L., Pang, S. M. & Lee, H. Y. (2014) Multiple drug hypersensitivity associated with severe cutaneous adverse reaction. *Clinical and Translational Allergy*, 4(S3), P135.
- Choo, S. Y. (2007) The HLA system: genetics, immunology, clinical testing, and clinical implications. *Yonsei medical journal*, 48(1), 11-23.

- Choon, S. E. & Lai, N. M. (2012) An epidemiological and clinical analysis of cutaneous adverse drug reactions seen in a tertiary hospital in Johor, Malaysia. *Indian J Dermatol Venereol Leprol*, 78(6), 734-9.
- Chou, C. T. & Lai, J. S. (1998) The epidemiology of hyperuricaemia and gout in Taiwan aborigines. *Br J Rheumatol*, 37(3), 258-62.
- Chou, C. T., Pei, L., Chang, D. M., Lee, C. F., Schumacher, H. R. & Liang, M. H. (1994) Prevalence of rheumatic diseases in Taiwan: a population study of urban, suburban, rural differences. *J Rheumatol*, 21(2), 302-6.
- Choudhary, S., McLeod, M., Torchia, D. & Romanelli, P. (2013) Drug reaction with eosinophilia and systemic symptoms (DRESS) syndrome. *The Journal of clinical and aesthetic dermatology*, 6(6), 31.
- Chowalloor, P. V., Siew, T. K. & Keen, H. I. (2014) Imaging in gout: A review of the recent developments. *Therapeutic advances in musculoskeletal disease*, 6(4), 131-143.
- Chuang, L. Y., Cheng, Y. H. & Yang, C. H. (2013) Specific primer design for the polymerase chain reaction. *Biotechnol Lett*, 35(10), 1541-9.
- Chung, W.-H., Chang, W.-C., Lee, Y.-S., Wu, Y.-Y., Yang, C.-H., Ho, H.-C., Chen, M.-J., Lin, J.-Y., Hui, R. C.-Y. & Ho, J.-C. (2014) Genetic variants associated with phenytoin-related severe cutaneous adverse reactions. *Jama*, 312(5), 525-534.
- Chung, W.-H., Hung, S.-I., Hong, H.-S., Hsieh, M.-S., Yang, L.-C., Ho, H.-C., Wu, J.-Y. & Chen, Y.-T. (2004) A marker for Stevens–Johnson syndrome. *Nature*, 428(6982), 486-486.
- Chung, W.-H., Hung, S.-I., Yang, J.-Y., Su, S.-C., Huang, S.-P., Wei, C.-Y., Chin, S.-W., Chiou, C.-C., Chu, S.-C. & Ho, H.-C. (2008) Granulysin is a key mediator for disseminated keratinocyte death in Stevens-Johnson syndrome and toxic epidermal necrolysis. *Nature medicine*, 14(12), 1343-1350.
- Cocco, E., Meloni, A., Murru, M. R., Corongiu, D., Tranquilli, S., Fadda, E., Murru, R., Schirru, L., Secci, M. A., Costa, G., Asunis, I., Cuccu, S., Fenu, G., Loreface, L., Carboni, N., Mura, G., Rosatelli, M. C. & Marrosu, M. G. (2012) Vitamin D responsive elements within the HLA-DRB1 promoter region in Sardinian multiple sclerosis associated alleles. *PLoS One*, 7(7), e41678.
- Collins, F. S. & McKusick, V. A. (2001) Implications of the Human Genome Project for medical science. *Jama*, 285(5), 540-544.
- Cozier, Y. C., Palmer, J. R. & Rosenberg, L. (2004) Comparison of methods for collection of DNA samples by mail in the Black Women's Health Study. *Ann Epidemiol*, 14(2), 117-22.
- Cristallo, A. F., Schroeder, J., Citterio, A., Santori, G., Ferrioli, G. M., Rossi, U., Bertani, G., Cassano, S., Gottardi, P. & Ceschini, N. (2011) A study of HLA class I and class II 4-digit

allele level in Stevens–Johnson syndrome and toxic epidermal necrolysis. *International journal of immunogenetics*, 38(4), 303-309.

Cronstein, B. N. & Sunkureddi, P. (2013) Mechanistic aspects of inflammation and clinical management of inflammation in acute gouty arthritis. *J Clin Rheumatol*, 19(1), 19-29.

Cui, G., Zhang, L., Xu, Y., Cianflone, K., Ding, H. & Wang, D. W. (2013) Development of a high resolution melting method for genotyping of risk HLA-DQA1 and PLA2R1 alleles and ethnic distribution of these risk alleles. *Gene*, 514(2), 125-130.

Curran, M. D., Williams, F., Earle, J. A. P., Rima, B. K., Van Dam, M. G., Bunce, M. & Middleton, D. (1996) Long-range PCR amplification as an alternative strategy for characterizing novel HLA-B alleles. *International Journal of Immunogenetics*, 23(4), 297-309.

D'Ans, G., Gottlieb, D. & Kokotovic, P. (1972) Optimal control of bacterial growth. *Automatica*, 8(6), 729-736.

Dalbeth, N., Choi, H. K., Joosten, L. A. B., Khanna, P. P., Matsuo, H., Perez-Ruiz, F. & Stamp, L. K. (2019) Gout. *Nat Rev Dis Primers*, 5(1), 69.

Dalbeth, N., Clark, B., Gregory, K., Gamble, G., Sheehan, T., Doyle, A. & McQueen, F. M. (2009) Mechanisms of bone erosion in gout: a quantitative analysis using plain radiography and computed tomography. *Annals of the rheumatic diseases*, 68(8), 1290-1295.

Dalbeth, N., Clark, B., Gregory, K., Gamble, G. D., Doyle, A. & McQueen, F. M. (2007) Computed tomography measurement of tophus volume: comparison with physical measurement. *Arthritis Care & Research*, 57(3), 461-465.

Dalbeth, N., Merriman, T. R. & Stamp, L. K. (2016) Gout. *Lancet*, 388(10055), 2039-2052.

Dalbeth, N., Pool, B., Gamble, G. D., Smith, T., Callon, K. E., McQueen, F. M. & Cornish, J. (2010) Cellular characterization of the gouty tophus: a quantitative analysis. *Arthritis Rheum*, 62(5), 1549-56.

Das Ghosh, D., Mukhopadhyay, I., Bhattacharya, A., Roy Chowdhury, R., Mandal, N. R., Roy, S. & Sengupta, S. (2017) Impact of genetic variations and transcriptional alterations of HLA class I genes on cervical cancer pathogenesis. *International Journal of Cancer*, 140(11), 2498-2508.

Das Gupta, E., Sakthiswary, R., Lee, S. L., Wong, S. F., Hussein, H. & Gun, S. C. (2018) Clinical significance of SLC2A9/GLUT9 rs11722228 polymorphisms in gout. *Int J Rheum Dis*, 21(3), 705-709.

Davatchi, F., Jamshidi, A. R., Banihashemi, A. T., Gholami, J., Forouzanfar, M. H., Akhlaghi, M., Barghamdi, M., Noorolahzadeh, E., Khabazi, A. R., Salesi, M., Salari, A. H., Karimifar, M., Essalat-Manesh, K., Hajjaliloo, M., Soroosh, M., Farzad, F., Moussavi, H. R., Samadi, F., Ghaznavi, K., Asgharifard, H., Zangiabadi, A. H., Shahram, F., Nadji, A., Akbarian, M. &

- Gharibdoost, F. (2008) WHO-ILAR COPCORD Study (Stage 1, Urban Study) in Iran. *J Rheumatol*, 35(7), 1384.
- Day, R. O., Graham, G. G., Hicks, M., McLachlan, A. J., Stocker, S. L. & Williams, K. M. (2007) Clinical pharmacokinetics and pharmacodynamics of allopurinol and oxypurinol. *Clinical pharmacokinetics*, 46(8), 623-644.
- De Leonardis, F., Govoni, M., Colina, M., Bruschi, M. & Trotta, F. (2007) Elderly-onset gout: a review. *Rheumatol Int*, 28(1), 1-6.
- De Santis, D., Dinauer, D., Duke, J., Erlich, H. A., Holcomb, C. L., Lind, C., Mackiewicz, K., Monos, D., Moudgil, A. & Norman, P. (2013) 16th IHIW: review of HLA typing by NGS. *International journal of immunogenetics*, 40(1), 72-76.
- Dessein, P., Shipton, E., Stanwix, A., Joffe, B. & Ramokgadi, J. (2000) Beneficial effects of weight loss associated with moderate calorie/carbohydrate restriction, and increased proportional intake of protein and unsaturated fat on serum urate and lipoprotein levels in gout: a pilot study. *Annals of the rheumatic diseases*, 59(7), 539-543.
- Dincer, H. E., Dincer, A. P. & Levinson, D. J. (2002) Asymptomatic hyperuricemia: to treat or not to treat. *Cleveland Clinic journal of medicine*, 69(8), 594-608.
- Ding, W. Y., Lee, C. K. & Choon, S. E. (2010) Cutaneous adverse drug reactions seen in a tertiary hospital in Johor, Malaysia. *Int J Dermatol*, 49(7), 834-41.
- Do, M. D., Le, L. G. H., Nguyen, V. T., Dang, T. N., Nguyen, N. H., Vu, H. A. & Mai, T. P. (2020) High-Resolution HLA Typing of HLA-A, -B, -C, -DRB1, and -DQB1 in Kinh Vietnamese by Using Next-Generation Sequencing. *Front Genet*, 11, 383.
- Dong, D., Tan-Koi, W.-C., Teng, G. G., Finkelstein, E. & Sung, C. (2015) Cost-effectiveness analysis of genotyping for HLA-B\* 5801 and an enhanced safety program in gout patients starting allopurinol in Singapore. *Pharmacogenomics*, 16(16), 1781-1793.
- Dubois, V., Tiercy, J. M., Labonne, M. P., Dormoy, A. & Gebuhrer, L. (2004) A new HLA-B44 allele (B\*44020102S) with a splicing mutation leading to a complete deletion of exon 5. *Tissue Antigens*, 63(2), 173-80.
- Duskin-Bitan, H., Cohen, E., Goldberg, E., Shochat, T., Levi, A., Garty, M. & Krause, I. (2014) The degree of asymptomatic hyperuricemia and the risk of gout. A retrospective analysis of a large cohort. *Clin Rheumatol*, 33(4), 549-53.
- Edinur, H. A., Dunn, P. P., Hammond, L., Selwyn, C., Brescia, P., Askar, M., Reville, P., Velickovic, Z. M., Lea, R. A. & Chambers, G. K. (2013) HLA and MICA polymorphism in Polynesians and New Zealand Maori: implications for ancestry and health. *Hum Immunol*, 74(9), 1119-29.
- Elnifro, E. M., Ashshi, A. M., Cooper, R. J. & Klapper, P. E. (2000) Multiplex PCR: optimization and application in diagnostic virology. *Clin Microbiol Rev*, 13(4), 559-70.

Elsner, H. A., Bernard, G., Eiz-Vesper, B., de Matteis, M., Bernard, A. & Blasczyk, R. (2002) Non-expression of HLA-A\*2901102 N is caused by a nucleotide exchange in the mRNA splicing site at the beginning of intron 4. *Tissue Antigens*, 59(2), 139-41.

Emmerson, B. (1998) Hyperlipidaemia in hyperuricaemia and gout. *Annals of the rheumatic diseases*, 57(9), 509-510.

Enomoto, A. & Endou, H. (2005) Roles of organic anion transporters (OATs) and a urate transporter (URAT1) in the pathophysiology of human disease. *Clinical and experimental nephrology*, 9(3), 195-205.

Erkeller-Yuksel, F. M., Deneys, V., Yuksel, B., Hannel, I., Hulstaert, F., Hamilton, C., Mackinnon, H., Stokes, L. T., Munhyeshuli, V., Vanlangendonck, F. & et al. (1992) Age-related changes in human blood lymphocyte subpopulations. *J Pediatr*, 120(2 Pt 1), 216-22.

Erlich, H. (2012) HLA DNA typing: past, present, and future. *Tissue antigens*, 80(1), 1-11.

Erlich, R. L., Jia, X., Anderson, S., Banks, E., Gao, X., Carrington, M., Gupta, N., DePristo, M. A., Henn, M. R., Lennon, N. J. & de Bakker, P. I. (2011) Next-generation sequencing for HLA typing of class I loci. *BMC Genomics*, 12, 42.

Faner, R., Casamitjana, N., Colobran, R., Ribera, A., Pujol-Borrell, R., Palou, E. & Juan, M. (2004) HLA-B27 genotyping by fluorescent resonance emission transfer (FRET) probes in real-time PCR. *Hum Immunol*, 65(8), 826-38.

Filippucci, E., Delle Sedie, A., Riente, L., Di Geso, L., Carli, L., Fulvia, C., Sakellariou, G., Iagnocco, A. & Grassi, W. (2013) Ultrasound imaging for the rheumatologist XLVII. Ultrasound of the shoulder in patients with gout and calcium pyrophosphate deposition disease.

Firestein, G. S. (2009) Kelley's textbook of rheumatology. Philadelphia, PA: Saunders. Elsevier.

Fortini, D., Ciammaruconi, A., De Santis, R., Fasanella, A., Battisti, A., D'Amelio, R., Lista, F., Cassone, A. & Carattoli, A. (2007) Optimization of high-resolution melting analysis for low-cost and rapid screening of allelic variants of *Bacillus anthracis* by multiple-locus variable-number tandem repeat analysis. *Clinical chemistry*, 53(7), 1377-1380.

Fraser, K. J. (1992) William Stukeley and the gout. *Medical history*, 36(2), 160-186.

Fricke-Galindo, I., Martínez-Juárez, I. E., Monroy-Jaramillo, N., Jung-Cook, H., Falfán-Valencia, R., Ortega-Vázquez, A., Alonso-Vilatela, M. E. & López-López, M. (2014) HLA-A\*02:01:01/-B\*35:01:01/-C\*04:01:01 haplotype associated with lamotrigine-induced maculopapular exanthema in Mexican Mestizo patients. *Pharmacogenomics*, 15(15), 1881-1891.

Garcia-Closas, M., Egan, K. M., Abruzzo, J., Newcomb, P. A., Titus-Ernstoff, L., Franklin, T., Bender, P. K., Beck, J. C., Le Marchand, L., Lum, A., Alavanja, M., Hayes, R. B., Rutter, J., Buetow, K., Brinton, L. A. & Rothman, N. (2001) Collection of genomic DNA from adults in epidemiological studies by buccal cytobrush and mouthwash. *Cancer Epidemiol Biomarkers Prev*, 10(6), 687-96.

Gauri, S., Gan, K. B. & Then, S.-M. (2019) Simulation of DNA extraction and separation from salivary fluid by superparamagnetic beads and electromagnetic field in microfluidic platform. *Microsystem Technologies*, 25(4), 1379-1385.

Gerster, J., Landry, M., Dufresne, L. & Meuwly, J. (2002) Imaging of tophaceous gout: computed tomography provides specific images compared with magnetic resonance imaging and ultrasonography. *Annals of the Rheumatic Diseases*, 61(1), 52-54.

Ghattaoraya, G. S., Dundar, Y., González-Galarza, F. F., Maia, M. H. T., Santos, E. J. M., da Silva, A. L. S., McCabe, A., Middleton, D., Alfirevic, A. & Dickson, R. (2016) A web resource for mining HLA associations with adverse drug reactions: HLA-ADR. *Database*, 2016.

Gonzalez-Galarza, F. F., McCabe, A., Santos, E., Jones, J., Takeshita, L., Ortega-Rivera, N. D., Cid-Pavon, G. M. D., Ramsbottom, K., Ghattaoraya, G., Alfirevic, A., Middleton, D. & Jones, A. R. (2020) Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. *Nucleic Acids Res*, 48(D1), D783-d788.

González-Galarza, F. F., Takeshita, L. Y., Santos, E. J., Kempson, F., Maia, M. H., da Silva, A. L., Teles e Silva, A. L., Ghattaoraya, G. S., Alfirevic, A., Jones, A. R. & Middleton, D. (2015) Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic Acids Res*, 43(Database issue), D784-8.

Goulder, P. J. & Walker, B. D. (2012) HIV and HLA class I: an evolving relationship. *Immunity*, 37(3), 426-40.

Gundry, C. N., Vandersteen, J. G., Reed, G. H., Pryor, R. J., Chen, J. & Wittwer, C. T. (2003) Amplicon melting analysis with labeled primers: a closed-tube method for differentiating homozygotes and heterozygotes. *Clinical chemistry*, 49(3), 396-406.

Hak, A. E., Curhan, G. C., Grodstein, F. & Choi, H. K. (2010a) Menopause, postmenopausal hormone use and risk of incident gout. *Ann Rheum Dis*, 69(7), 1305-9.

Hak, A. E., Curhan, G. C., Grodstein, F. & Choi, H. K. (2010b) Menopause, postmenopausal hormone use and risk of incident gout. *Annals of the Rheumatic Diseases*, 69(7), 1305.

Halevy, S., Ghislain, P. D., Mockenhaupt, M., Fagot, J. P., Bouwes Bavinck, J. N., Sidoroff, A., Naldi, L., Dunant, A., Viboud, C. & Roujeau, J. C. (2008) Allopurinol is the most common cause of Stevens-Johnson syndrome and toxic epidermal necrolysis in Europe and Israel. *J Am Acad Dermatol*, 58(1), 25-32.

Han, J., Liu, Y., Rao, F., Nievergelt, C. M., O'Connor, D. T., Wang, X., Liu, L., Bu, D., Liang, Y. & Wang, F. (2013) Common genetic variants of the human uromodulin gene regulate transcription and predict plasma uric acid levels. *Kidney international*, 83(4), 733-740.

Handoko, K. B., van Puijenbroek, E. P., Bijl, A. H., Hermens, W. A. J. J., Zwart-van Rijkom, J. E. F., Hekster, Y. A. & Egberts, T. C. G. (2008) Influence of chemical structure on hypersensitivity reactions induced by antiepileptic drugs: the role of the aromatic ring. *Drug safety*, 31(8), 695.

Hartung, E. F. (1957) Symposium on gout; historical considerations. *Metabolism*, 6(3), 196-208.

Harty, L. C., Garcia-Closas, M., Rothman, N., Reid, Y. A., Tucker, M. A. & Hartge, P. (2000) Collection of Buccal Cell DNA Using Treated Cards. *Cancer Epidemiology Biomarkers & Prevention*, 9(5), 501.

Hashim, R., Zainal, Z. A. & Taib, H. M. Preliminary report: Review of adverse drug reactions (ADRs) reporting in the Malaysian elderly.

He, Y., Wang, W., Han, Z., He, J., Chen, N., Dong, L., Tao, S., Zhang, W. & Zhu, F. (2016) HLA-B allele dropout in PCR sequence-specific oligonucleotide probe typing due to intronic polymorphism in the novel B\* 58: 01: 01: 02 allele. *International Journal of Immunogenetics*, 43(3), 180-183.

Health, M. o. (2008) *Clinical Practice Guidelines, Management of Gout*. Malaysia.

Hendel, H., Caillat-Zucman, S., Lebuane, H., Carrington, M., O'Brien, S., Andrieu, J. M., Schächter, F., Zagury, D., Rappaport, J., Winkler, C., Nelson, G. W. & Zagury, J. F. (1999) New class I and II HLA alleles strongly associated with opposite patterns of progression to AIDS. *J Immunol*, 162(11), 6942-6.

Herrmann, M. G., Durtschi, J. D., Bromley, L. K., Wittwer, C. T. & Voelkerding, K. V. (2006) Amplicon DNA melting analysis for mutation scanning and genotyping: cross-platform comparison of instruments and dyes. *Clin Chem*, 52(3), 494-503.

Hilmi, B. A., Asmahan, M. I. & Rosman, A. (2012) Use of newly available febuxostat in a case of chronic tophaceous gout contraindicated to allopurinol and probenecid. *Med J Malaysia*, 67(1), 125.

Hirayama, K., Zaidi, A. S. M., Hakim, S. L., Kimura, A., Ong, K. J., Kikuchi, M., Nasuruddin, H. A., Kojima, S. & Mak, J. W. (1996) Molecular analysis of HLA-B in the Malaysian aborigines. *Tissue Antigens*, 48(6), 692-697.

Hosomichi, K., Jinam, T. A., Mitsunaga, S., Nakaoka, H. & Inoue, I. (2013) Phase-defined complete sequencing of the HLA genes by next-generation sequencing. *BMC genomics*, 14(1), 355.

Huang, C., Chen, S.-P., Huang, Y.-H., Chen, H.-Y., Wang, Y.-F., Lee, M.-H. & Wang, S.-J. (2020) HLA class I alleles are associated with clinic-based migraine and increased risks of chronic migraine and medication overuse. *Cephalalgia*, 40(5), 493-502.

Huang, X., Du, H., Gu, J., Zhao, D., Jiang, L., Li, X., Zuo, X., Liu, Y., Li, Z., Zhu, P., Li, J., Zhang, Z., Huang, A., Zhang, Y. & Bao, C. (2014) An allopurinol-controlled, multicenter, randomized, double-blind, parallel between-group, comparative study of febuxostat in Chinese patients with gout and hyperuricemia. *Int J Rheum Dis*, 17(6), 679-86.

Huh, J. Y., Han, O. J. & Park, G. (2018) Rapid, Reliable, and Inexpensive HLA-B\*58:01 Detection Method Using DNA Binding Dye-based Duplex Allele-specific Melting Curve Analysis. *Ann Clin Lab Sci*, 48(3), 296-300.

Hung, S.-I., Chung, W.-H., Liou, L.-B., Chu, C.-C., Lin, M., Huang, H.-P., Lin, Y.-L., Lan, J.-L., Yang, L.-C. & Hong, H.-S. (2005a) HLA-B\* 5801 allele as a genetic marker for severe cutaneous adverse reactions caused by allopurinol. *Proceedings of the National Academy of Sciences*, 102(11), 4134-4139.

Hung, S.-I., Chung, W.-H., Liu, Z.-S., Chen, C.-H., Hsih, M.-S., Hui, R. C.-y., Chu, C.-Y. & Chen, Y.-T. (2010) Common risk allele in aromatic antiepileptic-drug induced Stevens–Johnson syndrome and toxic epidermal necrolysis in Han Chinese. *Pharmacogenomics*, 11(3), 349-356.

Hung, S. I., Chung, W. H., Liou, L. B., Chu, C. C., Lin, M., Huang, H. P., Lin, Y. L., Lan, J. L., Yang, L. C., Hong, H. S., Chen, M. J., Lai, P. C., Wu, M. S., Chu, C. Y., Wang, K. H., Chen, C. H., Fann, C. S., Wu, J. Y. & Chen, Y. T. (2005b) HLA-B\*5801 allele as a genetic marker for severe cutaneous adverse reactions caused by allopurinol. *Proc Natl Acad Sci U S A*, 102(11), 4134-9.

Huppertz, A., Hermann, K.-G. A., Diekhoff, T., Wagner, M., Hamm, B. & Schmidt, W. A. (2014) Systemic staging for urate crystal deposits with dual-energy CT and ultrasound in patients with suspected gout. *Rheumatology international*, 34(6), 763-771.

Ichida, K., Matsuo, H., Takada, T., Nakayama, A., Murakami, K., Shimizu, T., Yamanashi, Y., Kasuga, H., Nakashima, H. & Nakamura, T. (2012) Decreased extra-renal urate excretion is a common cause of hyperuricemia. *Nature communications*, 3, 764.

Imperiali, C., Alía-Ramos, P. & Padró-Miquel, A. (2015) Rapid detection of HLA-B\*51 by real-time polymerase chain reaction and high-resolution melting analysis. *Tissue Antigens*, 86(2), 139-42.

Isomaki, H. A. & Takkunen, H. (1969) Gout and hyperuricemia in a Finnish rural population. *Acta Rheumatol Scand*, 15(2), 112-20.

Jain, K. K. (2002) Personalized medicine. *Current opinion in molecular therapeutics*, 4(6), 548.

Jantararoungtong, T., Rerkpattanapipat, T., Prommas, S., Koomdee, N., Santon, S., Chammanphol, M., Puangpetch, A. & Sukasem, C. (2014) HLA-B\* 58: 01 allele is strongly associated with allopurinol-induced severe cutaneous adverse reactions in a Thai population. *Clinical and Translational Allergy*, 4(S3), P120.

Jaruthamsophon, K., Tipmanee, V., Sangiemchoey, A., Sukasem, C. & Limprasert, P. (2017) HLA-B\*15:21 and carbamazepine-induced Stevens-Johnson syndrome: pooled-data and in silico analysis. *Sci Rep*, 7, 45553.

Joosten, L. A., Netea, M. G., Mylona, E., Koenders, M. I., Malireddi, R. K., Oosting, M., Stienstra, R., van de Veerdonk, F. L., Stalenhoef, A. F., Giamarellos-Bourboulis, E. J., Kanneganti, T. D. & van der Meer, J. W. (2010) Engagement of fatty acids with Toll-like receptor 2 drives interleukin-1beta production via the ASC/caspase 1 pathway in monosodium urate monohydrate crystal-induced gouty arthritis. *Arthritis Rheum*, 62(11), 3237-48.

Jorgez, C. J., Dang, D. D., Simpson, J. L., Lewis, D. E. & Bischoff, F. Z. (2006) Quantity versus quality: optimal methods for cell-free DNA isolation from plasma of pregnant women. *Genet Med*, 8(10), 615-9.

Jung, J.-W., Song, W.-J., Kim, Y.-S., Joo, K. W., Lee, K. W., Kim, S.-H., Park, H.-W., Chang, Y.-S., Cho, S.-H. & Min, K.-U. (2011) HLA-B58 can help the clinical decision on starting allopurinol in patients with chronic renal insufficiency. *Nephrology Dialysis Transplantation*, 26(11), 3567-3572.

Jutkowitz, E., Dubreuil, M., Lu, N., Kuntz, K. M. & Choi, H. K. The cost-effectiveness of HLA-B\* 5801 screening to guide initial urate-lowering therapy for gout in the United States 2017. Elsevier.

Kamei, K., Konta, T., Hirayama, A., Suzuki, K., Ichikawa, K., Fujimoto, S., Iseki, K., Moriyama, T., Yamagata, K., Tsuruya, K., Kimura, K., Narita, I., Kondo, M., Asahi, K. & Watanabe, T. (2014) A slight increase within the normal range of serum uric acid and the decline in renal function: associations in a community-based population. *Nephrology Dialysis Transplantation*, 29(12), 2286-2292.

Kanbara, A. & Seyama, I. (2011) Effect of urine pH on uric acid excretion by manipulating food materials. *Nucleosides, Nucleotides and Nucleic Acids*, 30(12), 1066-1071.

Kang, H.-R., Jee, Y. K., Kim, Y.-S., Lee, C. H., Jung, J.-W., Kim, S. H., Park, H.-W., Chang, Y.-S., Jang, I.-J. & Cho, S.-H. (2011) Positive and negative associations of HLA class I alleles with allopurinol-induced SCARs in Koreans. *Pharmacogenetics and genomics*, 21(5), 303-307.

Kang, X., Chen, R., Han, M., Liu, Z., Liu, J., Dai, P., Chen, C. & Wang, H. (2016a) Rapid and reliable genotyping of HLA-B\*58:01 in four Chinese populations using a single-tube duplex real-time PCR assay. *Pharmacogenomics*, 17(1), 47-57.

Kang, X., Chen, R., Han, M., Liu, Z., Liu, J., Dai, P., Chen, C. & Wang, H. (2016b) Rapid and reliable genotyping of HLA-B\* 58: 01 in four Chinese populations using a single-tube duplex real-time PCR assay. *Pharmacogenomics*, 17(1), 47-57.

Kaniwa, N., Saito, Y., Aihara, M., Matsunaga, K., Tohkin, M., Kurose, K., Sawada, J.-i., Furuya, H., Takahashi, Y. & Muramatsu, M. (2008) HLA-B locus in Japanese patients with anti-epileptics and allopurinol-related Stevens–Johnson syndrome and toxic epidermal necrolysis.

Kano, Y., Hirahara, K., Mitsuyama, Y., Takahashi, R. & Shiohara, T. (2007) Utility of the lymphocyte transformation test in the diagnosis of drug sensitivity: dependence on its timing and the type of drug eruption. *Allergy*, 62(12), 1439-44.

Karnes, J. H., Rettie, A. E., Somogyi, A. A., Huddart, R., Fohner, A. E., Formea, C. M., Lee, M. T. M., Llerena, A., Whirl-Carrillo, M. & Klein, T. E. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2C9 and HLA-B Genotypes and Phenytoin Dosing: 2020 Update. *Clinical Pharmacology & Therapeutics*.

Kayser, M., Lao, O., Saar, K., Brauer, S., Wang, X., Nürnberg, P., Trent, R. J. & Stoneking, M. (2008) Genome-wide analysis indicates more Asian than Melanesian ancestry of Polynesians. *Am J Hum Genet*, 82(1), 194-8.

Ke, C.-H., Chung, W.-H., Wen, Y.-H., Huang, Y.-B., Chuang, H.-Y., Tain, Y.-L., Wang, Y.-C. L., Wu, C.-C. & Hsu, C.-N. (2017) Cost-effectiveness analysis for genotyping before allopurinol treatment to prevent severe cutaneous adverse drug reactions. *The Journal of rheumatology*, 44(6), 835-843.

Keag, O. E., Murphy, L., Bradley, A., Deakin, N., Whyte, S., Norman, J. E. & Stock, S. J. (2020) Postal recruitment for genetic studies of preterm birth: A feasibility study. *Wellcome open research*, 5, 26-26.

Khoo, A. K. M. & Foo, C. L. (1996) Toxic epidermal necrolysis in a burns centre: a 6-year review. *Burns*, 22(4), 275-278.

Klemp, P., Stansfield, S. A., Castle, B. & Robertson, M. C. (1997) Gout is on the increase in New Zealand. *Ann Rheum Dis*, 56(1), 22-6.

Klosinski, D. D. (1997) Collecting Specimens From the Elderly Patient. *Laboratory Medicine*, 28(8), 518-522.

Ko, T.-M., Tsai, C.-Y., Chen, S.-Y., Chen, K.-S., Yu, K.-H., Chu, C.-S., Huang, C.-M., Wang, C.-R., Weng, C.-T. & Yu, C.-L. (2015) Use of HLA-B\* 58: 01 genotyping to prevent allopurinol induced severe cutaneous adverse reactions in Taiwan: national prospective cohort study. *Bmj*, 351.

Kolz, M., Johnson, T., Sanna, S., Teumer, A., Vitart, V., Perola, M., Mangino, M., Albrecht, E., Wallace, C. & Farrall, M. (2009) Meta-analysis of 28,141 individuals identifies common

variants within five new loci that influence uric acid concentrations. *PLoS genetics*, 5(6), e1000504.

Kottgen, A., Albrecht, E., Teumer, A., Vitart, V., Krumsiek, J., Hundertmark, C., Pistis, G., Ruggiero, D., O'Seaghdha, C. M., Haller, T., Yang, Q., Tanaka, T., Johnson, A. D., Kutalik, Z., Smith, A. V., Shi, J., Struchalin, M., Middelberg, R. P., Brown, M. J., Gaffo, A. L., Pirastu, N., Li, G., Hayward, C., Zemunik, T., Huffman, J., Yengo, L., Zhao, J. H., Demirkan, A., Feitosa, M. F., Liu, X., Malerba, G., Lopez, L. M., van der Harst, P., Li, X., Kleber, M. E., Hicks, A. A., Nolte, I. M., Johansson, A., Murgia, F., Wild, S. H., Bakker, S. J., Peden, J. F., Dehghan, A., Steri, M., Tenesa, A., Lagou, V., Salo, P., Mangino, M., Rose, L. M., Lehtimäki, T., Woodward, O. M., Okada, Y., Tin, A., Müller, C., Oldmeadow, C., Putku, M., Czamara, D., Kraft, P., Frogger, L., Thun, G. A., Grotevendt, A., Gislason, G. K., Harris, T. B., Launer, L. J., McArdle, P., Shuldiner, A. R., Boerwinkle, E., Coresh, J., Schmidt, H., Schallert, M., Martin, N. G., Montgomery, G. W., Kubo, M., Nakamura, Y., Munroe, P. B., Samani, N. J., Jacobs, D. R., Jr., Liu, K., D'Adamo, P., Ulivi, S., Rotter, J. I., Psaty, B. M., Vollenweider, P., Waeber, G., Campbell, S., Devuyst, O., Navarro, P., Kolcic, I., Hastie, N., Balkau, B., Froguel, P., Esko, T., Salumets, A., Khaw, K. T., Langenberg, C., Wareham, N. J., Isaacs, A., Kraja, A., Zhang, Q., Wild, P. S., et al (2013) Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nat Genet*, 45(2), 145-54.

Kramer, H. M. & Curhan, G. (2002) The association between gout and nephrolithiasis: the National Health and Nutrition Examination Survey III, 1988-1994. *American Journal of Kidney Diseases*, 40(1), 37-42.

Kuo, C.-F., Grainge, M. J., Zhang, W. & Doherty, M. (2015a) Global epidemiology of gout: prevalence, incidence and risk factors. *Nature Reviews Rheumatology*, 11(11), 649-662.

Kuo, C. F., Grainge, M. J., Mallen, C., Zhang, W. & Doherty, M. (2015b) Rising burden of gout in the UK but continuing suboptimal management: a nationwide population study. *Ann Rheum Dis*, 74(4), 661-7.

Kuo, C. F., Grainge, M. J., See, L. C., Yu, K. H., Luo, S. F., Valdes, A. M., Zhang, W. & Doherty, M. (2015c) Familial aggregation of gout and relative genetic and environmental contributions: a nationwide population study in Taiwan. *Ann Rheum Dis*, 74(2), 369-74.

Kuo, C. F., Grainge, M. J., See, L. C., Yu, K. H., Luo, S. F., Zhang, W. & Doherty, M. (2015d) Epidemiology and management of gout in Taiwan: a nationwide population study. *Arthritis Res Ther*, 17, 13.

Kwok, J. & Kwong, K. M. (2013) Detection of HLA-B\*58:01, the susceptible allele for allopurinol-induced hypersensitivity, by loop-mediated isothermal amplification. *Br J Dermatol*, 168(3), 526-32.

Le Marchand, L., Lum-Jones, A., Saltzman, B., Visaya, V., Nomura, A. M. & Kolonel, L. N. (2001) Feasibility of collecting buccal cell DNA by mail in a cohort study. *Cancer Epidemiol Biomarkers Prev*, 10(6), 701-3.

Lee, H. Y., Ariyasinghe, J. T. N. & Thirumoorthy, T. (2008) Allopurinol hypersensitivity syndrome: a preventable severe cutaneous adverse reaction? *Singapore medical journal*, 49(5), 384.

Lee, M. H., Stocker, S. L., Anderson, J., Phillips, E. J., Nolan, D., Williams, K. M., Graham, G. G., Sullivan, J. R. & Day, R. O. (2012) Initiating allopurinol therapy: do we need to know the patient's human leucocyte antigen status? *Internal medicine journal*, 42(4), 411-416.

Leitman, E. M., Hurst, J., Mori, M., Kublin, J., Ndung'u, T., Walker, B. D., Carlson, J., Gray, G. E., Matthews, P. C., Frahm, N. & Goulder, P. J. (2016) Lower Viral Loads and Slower CD4+ T-Cell Count Decline in MRKAd5 HIV-1 Vaccinees Expressing Disease-Susceptible HLA-B\*58:02. *J Infect Dis*, 214(3), 379-89.

Li, G., Li, W. X. & Shi, W. (2020a) The novel HLA-B\*58:118 allele was identified during high-resolution HLA typing. *Hla*, 95(3), 214-215.

Li, Y. J., Phillips, E., Dellinger, A., Nicoletti, P., Schutte, R., Li, D., Ostrov, D. A., Fontana, R. J., Watkins, P. B. & Stolz, A. (2020b) HLA-B\* 14: 01 and HLA-B\* 35: 01 are associated with trimethoprim-sulfamethoxazole induced liver injury. *Hepatology*.

Lien, Y. H. & Logan, J. L. (2017) Cross-Reactions Between Allopurinol and Febuxostat. *Am J Med*, 130(2), e67-e68.

Lin, K. C., Lin, H. Y. & Chou, P. (2000) The interaction between uric acid level and other risk factors on the development of gout among asymptomatic hyperuricemic men in a prospective study. *J Rheumatol*, 27(6), 1501-5.

Lin, M. S., Dai, Y. S., Pwu, R. F., Chen, Y. H. & Chang, N. C. (2005) Risk estimates for drugs suspected of being associated with Stevens-Johnson syndrome and toxic epidermal necrolysis: a case-control study. *Intern Med J*, 35(3), 188-90.

Lind, C., Ferriola, D., Mackiewicz, K., Heron, S., Rogers, M., Slavich, L., Walker, R., Hsiao, T., McLaughlin, L., D'Arcy, M., Gai, X., Goodridge, D., Sayer, D. & Monos, D. (2010) Next-generation sequencing: the solution for high-resolution, unambiguous human leukocyte antigen typing. *Hum Immunol*, 71(10), 1033-42.

Liu, C., Chen, G., Bentley, A. R., Doumatey, A., Zhou, J., Adeyemo, A., Yang, J. & Rotimi, C. (2019) Genome-wide association study for proliferative diabetic retinopathy in Africans. *NPJ genomic medicine*, 4(1), 1-7.

Low DE, Nurul-Aain AF, Tan WC, Tang JJ, Bakhtiar MF, Murad S, Chang CC, Too CL, Tang MM. (2020) HLA-B\*58: 01 association in allopurinol-induced severe cutaneous adverse reactions: the implication of ethnicity and clinical phenotypes in multiethnic Malaysia. *Pharmacogenet Genomics*, 30(7), 153-160.

Longman-Jacobsen, N., Williamson, J. F., Dawkins, R. L. & Gaudieri, S. (2003) In polymorphic genomic regions indels cluster with nucleotide polymorphism: Quantum Genomics. *Gene*, 312, 257-61.

Lonjou, C., Borot, N., Sekula, P., Ledger, N., Thomas, L., Halevy, S., Naldi, L., Bouwes-Bavinck, J.-N., Sidoroff, A. & de Toma, C. (2008) A European study of HLA-B in Stevens–Johnson syndrome and toxic epidermal necrolysis related to five high-risk drugs. *Pharmacogenetics and genomics*, 18(2), 99-107.

Loo, C. H., Tan, W. C., Khor, Y. H. & Chan, L. C. (2018) A 10-years retrospective study on Severe Cutaneous Adverse Reactions (SCARs) in a tertiary hospital in Penang, Malaysia. *Med J Malaysia*, 73(2), 73-77.

Lundgren, A., Kim, S., Stadnisky, M. D. & Brown, M. G. (2012) Rapid discrimination of MHC class I and killer cell lectin-like receptor allele variants by high-resolution melt analysis. *Immunogenetics*, 64(8), 633-640.

Mack, S. J. & Erlich, H. A. (1998) HLA class II polymorphism in the Ticuna of Brazil: evolutionary implications of the DRB1\*0807 allele. *Tissue Antigens*, 51(1), 41-50.

Madden, K. & Chabot-Richards, D. (2019) HLA testing in the molecular diagnostic laboratory. *Virchows Arch*, 474(2), 139-147.

Madrac (2002) Malaysian Guidelines for the Reporting and Monitoring of ADR. National Pharmaceutical Control Bureau, Ministry of Health Malaysia Kuala Lumpur.

Mahmoud, H., Leverger, G., Patte, C., Harvey, E. & Lascombes, F. (1998) Advances in the management of malignancy-associated hyperuricaemia. *British journal of cancer*, 77(S4), 18.

Malik, A., Schumacher, H. R., Dinnella, J. E. & Clayburne, G. M. (2009) Clinical diagnostic criteria for gout: comparison with the gold standard of synovial fluid crystal analysis. *JCR: Journal of Clinical Rheumatology*, 15(1), 22-24.

Mallal, S., Phillips, E., Carosi, G., Molina, J.-M., Workman, C., Tomažič, J., Jägel-Guedes, E., Rugina, S., Kozyrev, O. & Cid, J. F. (2008) HLA-B\* 5701 screening for hypersensitivity to abacavir. *New England Journal of Medicine*, 358(6), 568-579.

Mandal, A. K. & Mount, D. B. (2015) The Molecular Physiology of Uric Acid Homeostasis. *Annual Review of Physiology*, 77(1), 323-345.

Martinon, F., Petrilli, V., Mayor, A., Tardivel, A. & Tschopp, J. (2006) Gout-associated uric acid crystals activate the NALP3 inflammasome. *Nature*, 440(7081), 237-41.

Mauck, M., DeGueme, A., Taintor, A. & Jha, P. CROSS-SENSITIVITY OF ALLOPURINOL AND FEBUXOSTAT INDUCED DRUG RASH WITH EOSINOPHILIA AND SYSTEMIC SYMPTOMS (DRESS) SYNDROME 2010. SPRINGER 233 SPRING ST, NEW YORK, NY 10013 USA.

Maynard, J. W., McAdams-DeMarco, M. A., Law, A., Kao, L., Gelber, A. C., Coresh, J. & Baer, A. N. (2014) Racial differences in gout incidence in a population-based cohort: Atherosclerosis Risk in Communities Study. *Am J Epidemiol*, 179(5), 576-83.

MCCARTY, D. J. & HOLLANDER, J. L. (1961) Identification of urate crystals in gouty synovial fluid. *Annals of internal medicine*, 54(3), 452-460.

McCrum-Gardner, E. (2008) Which is the correct statistical test to use? *British Journal of Oral and Maxillofacial Surgery*, 46(1), 38-41.

McDonagh, E. M., Thorn, C. F., Callaghan, J. T., Altman, R. B. & Klein, T. E. (2014) PharmGKB summary: uric acid-lowering drugs pathway, pharmacodynamics. *Pharmacogenetics and genomics*, 24(9), 464.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S. & Daly, M. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, 20(9), 1297-1303.

McQueen, F. M., Doyle, A. & Dalbeth, N. (2011) Imaging in gout-What can we learn from MRI, CT, DECT and US? *Arthritis research & therapy*, 13(6), 246.

Mellen, P. B., Bleyer, A. J., Erlinger, T. P., Evans, G. W., Nieto, F. J., Wagenknecht, L. E., Wofford, M. R. & Herrington, D. M. (2006) Serum uric acid predicts incident hypertension in a biethnic cohort: the atherosclerosis risk in communities study. *Hypertension*, 48(6), 1037-1042.

Merriman, T. R., Choi, H. K. & Dalbeth, N. (2014) The genetic basis of gout. *Rheum Dis Clin North Am*, 40(2), 279-90.

Min, F.-L., Mao, B.-J., Zheng, Z.-Z., He, N., Fan, C.-X., Cai, R.-Y., Wang, J., Ou, Y.-M., Qin, B., Liao, W.-P., Yi, Y.-H., Li, Z. & Shi, Y.-W. (2019) HLA-B\*13:01 as a Risk Allele for Antiepileptic Drugs-Induced Cutaneous Adverse Reactions: Higher Risk for Cross-Reactivity? *Frontiers in Neurology*, 10, 614.

Misawa, T., Takahama, M., Kozaki, T., Lee, H., Zou, J., Saitoh, T. & Akira, S. (2013) Microtubule-driven spatial arrangement of mitochondria promotes activation of the NLRP3 inflammasome. *Nat Immunol*, 14(5), 454-60.

Mohd, A., Gun, S. C. & De'Souza, B. (2006) The study of clinical characteristics of gout patients on follow up in Rheumatology Clinic, Hospital Seremban. *Aplar Journal of Rheumatology*, 9, A89.

Mohd, A., Gupta, E. d., Loh, Y., Gandhi, C., D'Souza, B. & Gun, S. (2011) Clinical characteristics of gout: a hospital case series. *Malaysian family physician : the official journal of the Academy of Family Physicians of Malaysia*, 6(2-3), 72-73.

- Monshi, M. M., Faulkner, L., Gibson, A., Jenkins, R. E., Farrell, J., Earnshaw, C. J., Alfirevic, A., Cederbrant, K., Daly, A. K. & French, N. (2013) Human leukocyte antigen (HLA)-B\* 57: 01-restricted activation of drug-specific T cells provides the immunological basis for flucloxacillin-induced liver injury. *Hepatology*, 57(2), 727-739.
- Montgomery, J., Wittwer, C. T., Palais, R. & Zhou, L. (2007) Simultaneous mutation scanning and genotyping by high-resolution DNA melting analysis. *Nature protocols*, 2(1), 59-66.
- Moon, J., Kim, T. J., Lim, J. A., Sunwoo, J. S., Byun, J. I., Lee, S. T., Jung, K. H., Park, K. I., Jung, K. Y. & Jeon, D. (2016) HLA-B\* 40: 02 and DRB 1\* 04: 03 are risk factors for oxcarbazepine-induced maculopapular eruption. *Epilepsia*, 57(11), 1879-1886.
- Mori, M., Wichukchinda, N., Miyahara, R., Rojanawiwat, A., Pathipvanich, P., Maekawa, T., Miura, T., Goulder, P., Yasunami, M. & Ariyoshi, K. (2014) HLA-B\* 35: 05 is a protective allele with a unique structure among HIV-1 CRF01\_AE-infected Thais, in whom the B\* 57 frequency is low. *Aids*, 28(7), 959-967.
- Mosaad, Y. M. (2015) Clinical Role of Human Leukocyte Antigen in Health and Disease. *Scand J Immunol*, 82(4), 283-306.
- Nath, S. D., Voruganti, V. S., Arar, N. H., Thameem, F., Lopez-Alvarenga, J. C., Bauer, R., Blangero, J., MacCluer, J. W., Comuzzie, A. G. & Abboud, H. E. (2007) Genome scan for determinants of serum uric acid variability. *J Am Soc Nephrol*, 18(12), 3156-63.
- Nauwelaerts, S. J. D., Van Geel, D., Delvoye, M., De Cremer, K., Bernard, A., Roosens, N. H. C. & De Keersmaecker, S. C. J. (2020) Selection of a Noninvasive Source of Human DNA Envisaging Genotyping Assays in Epidemiological Studies: Urine or Saliva? *Journal of biomolecular techniques : JBT*, jbt.20-3101-004.
- Nebert, D. W. & Menon, A. G. (2001) Pharmacogenomics, ethnicity, and susceptibility genes. *The pharmacogenomics journal*, 1(1), 19-22.
- Negeri, Y. d.-P. (2015) Sultanate of Malacca. *Population*, 2.
- Nestorova, R. & Fodor, D. (2015) Crystal-induced arthritis, *Musculoskeletal Ultrasonography in Rheumatic Diseases* Springer, 137-167.
- Nguyen, D. V., Vida, C., Chu, H. C., Fulton, R., Li, J. & Fernando, S. L. (2017) Validation of a Rapid, Robust, Inexpensive Screening Method for Detecting the HLA-B\*58:01 Allele in the Prevention of Allopurinol-Induced Severe Cutaneous Adverse Reactions. *Allergy Asthma Immunol Res*, 9(1), 79-84.
- Nguyen, D. V., Vidal, C., Chu, H. C. & van Nunen, S. (2019) Human leukocyte antigen-associated severe cutaneous adverse drug reactions: from bedside to bench and beyond. *Asia Pacific Allergy*, 9(3).

Niihara, H., Kaneko, S., Ito, T., Sugamori, T., Takahashi, N., Kohno, K. & Morita, E. (2013) HLA-B\* 58: 01 strongly associates with allopurinol-induced adverse drug reactions in a Japanese sample population. *Journal of dermatological science*, 71(2), 150-152.

Nunes, A. P., Oliveira, I. O., Santos, B. R., Millech, C., Silva, L. P., González, D. A., Hallal, P. C., Menezes, A. M. B., Araújo, C. L. & Barros, F. C. (2012) Quality of DNA extracted from saliva samples collected with the Oragene™ DNA self-collection kit. *BMC Medical Research Methodology*, 12(1), 65.

O'Sullivan, J. B. (1972) Gout in a New England town. A prevalence study in Sudbury, Massachusetts. *Ann Rheum Dis*, 31(3), 166-9.

Obregón-Ponce, A., Iraheta, I., García-Ferrer, H., Mejia, B. & García-Kutzbach, A. (2012) Prevalence of musculoskeletal diseases in Guatemala, Central America: the COPCORD study of 2 populations. *J Clin Rheumatol*, 18(4), 170-4.

Omoumi, P., Becce, F., Ott, J. G., Racine, D. & Verdun, F. R. (2015a) Optimization of radiation dose and image quality in musculoskeletal CT: emphasis on iterative reconstruction techniques (Part 1), *Seminars in musculoskeletal radiology*. Thieme Medical Publishers.

Omoumi, P., Becce, F., Racine, D., Ott, J. G., Andreisek, G. & Verdun, F. R. (2015b) Dual-energy CT: basic principles, technical approaches, and applications in musculoskeletal imaging (part 1), *Seminars in musculoskeletal radiology*. Thieme Medical Publishers.

Omoumi, P., Verdun, F. R. & Becce, F. (2015c) Optimization of radiation dose and image quality in musculoskeletal CT: emphasis on iterative reconstruction techniques (part 2), *Seminars in musculoskeletal radiology*. Thieme Medical Publishers.

Omoumi, P., Zufferey, P., Malghem, J. & So, A. (2016) Imaging in gout and other crystal-related arthropathies. *Rheumatic Disease Clinics*, 42(4), 621-644.

Palais, R. A., Liew, M. A. & Wittwer, C. T. (2005) Quantitative heteroduplex analysis for single nucleotide polymorphism genotyping. *Analytical biochemistry*, 346(1), 167-175.

Papúchová, H., Meissner, T. B., Li, Q., Strominger, J. L. & Tilburgs, T. (2019) The Dual Role of HLA-C in Tolerance and Immunity at the Maternal-Fetal Interface. *Frontiers in Immunology*, 10, 2730.

Parikh, R., Mathai, A., Parikh, S., Chandra Sekhar, G. & Thomas, R. (2008) Understanding and using sensitivity, specificity and predictive values. *Indian journal of ophthalmology*, 56(1), 45-50.

Park, D. J., Kang, J. H., Lee, J. W., Lee, K. E., Wen, L., Kim, T. J., Park, Y. W., Park, S. H. & Lee, S. S. (2015) Cost-Effectiveness Analysis of HLA-B5801 Genotyping in the Treatment of Gout Patients With Chronic Renal Insufficiency in Korea. *Arthritis care & research*, 67(2), 280-287.

- Pascart, T., Oehler, E. & Flipo, R. M. (2014) Gout in French Polynesia: a survey of common practices. *Joint Bone Spine*, 81(4), 374-5.
- Pascual, E., Batlle-Gualda, E., Martínez, A., Rosas, J. & Vela, P. (1999) Synovial fluid analysis for diagnosis of intercritical gout. *Annals of Internal Medicine*, 131(10), 756-759.
- Pascual, E., Tovar, J. & Ruiz, M. (1989) The ordinary light microscope: an appropriate tool for provisional detection and identification of crystals in synovial fluid. *Annals of the rheumatic diseases*, 48(12), 983-985.
- Pavlos, R., Mallal, S., Ostrov, D., Buus, S., Metushi, I., Peters, B. & Phillips, E. (2015) T cell-mediated hypersensitivity reactions to drugs. *Annu Rev Med*, 66, 439-54.
- Perez-Ruiz, F., Castillo, E., Chinchilla, S. P. & Herrero-Beites, A. M. (2014) Clinical manifestations and diagnosis of gout. *Rheumatic Disease Clinics*, 40(2), 193-206.
- Phipps-Green, A., Merriman, M., Topless, R., Altaf, S., Montgomery, G., Franklin, C., Jones, G., Van Rij, A., White, D. & Stamp, L. (2016) Twenty-eight loci that influence serum urate levels: analysis of association with gout. *Annals of the rheumatic diseases*, 75(1), 124-130.
- Pichler, W. J., Adam, J., Daubner, B., Gentinetta, T., Keller, M. & Yerly, D. (2010) Drug hypersensitivity reactions: pathomechanism and clinical symptoms. *Med Clin North Am*, 94(4), 645-64, xv.
- Plumpton, C. O., Alfirevic, A., Pirmohamed, M. & Hughes, D. A. (2017) Cost effectiveness analysis of HLA-B\* 58: 01 genotyping prior to initiation of allopurinol for gout. *Rheumatology*, 56(10), 1729-1739.
- Poláková, K. M., Lopotová, T., Klamová, H. & Moravcová, J. (2008) High-resolution melt curve analysis: initial screening for mutations in BCR-ABL kinase domain. *Leukemia research*, 32(8), 1236-1243.
- Pompeu, Y. A., Stewart, J. D., Mallal, S., Phillips, E., Peters, B. & Ostrov, D. A. (2012) The structural basis of HLA-associated drug hypersensitivity syndromes. *Immunological reviews*, 250(1), 158-166.
- Popert, A. J. & Hewitt, J. V. (1962) Gout and hyperuricaemia in rural and urban populations. *Ann Rheum Dis*, 21, 154-63.
- Pozzi, S., Longo, A. & Ferrara, G. B. (1999) HLA-B locus sequence-based typing. *Tissue antigens*, 53(3), 275-281.
- Poór, G. & Mituszova, M. (2003) History, classification and epidemiology of crystal-related arthropathies. *Rheumatology*, 2, 1893-1902.
- Practitioners, G. B. O. o. P. C. a. S. G. B. D. o. H. a. S. S. R. C. o. G. (1982) *Morbidity statistics from general practice 1970-71 : socio-economic analyses / Royal College of General*

*Practitioners, Office of Population Censuses and Surveys, Department of Health and Social Security.* London : H.M.S.O., 1982.: H.M.S.O.

Prior, I. A., Welby, T. J., Ostbye, T., Salmond, C. E. & Stokes, Y. M. (1987) Migration and gout: the Tokelau Island migrant study. *Br Med J (Clin Res Ed)*, 295(6596), 457-61.

Puangpetch, A., Koomdee, N., Chamnanphol, M., Jantararoungtong, T., Santon, S., Prommas, S., Hongkaew, Y. & Sukasem, C. (2015) HLA-B allele and haplotype diversity among Thai patients identified by PCR-SSOP: evidence for high risk of drug-induced hypersensitivity. *Frontiers in genetics*, 5, 478.

Ragab, G., Elshahaly, M. & Bardin, T. (2017) Gout: An old disease in new perspective - A review. *J Adv Res*, 8(5), 495-511.

Ramalingam, R. (2018) Severe cutaneous adverse reactions: a 5-year experience in a tertiary referral hospital in Malaysia. *Malaysian J Dermatol*, 41, 28-34.

Rani, Z. Z. M., Murad, N. A. A., Then, S.-M., Bernam, S. P., Abdullah, A., Saimun, S., Othman, S. N., Ali, R. A. & Jamal, R. (2018) HLA-B\* 15: 02 screening in epileptic patients using a high resolution melting-real time PCR (HRM-QPCR) method. *Neurology Asia*, 23(2).

Ravi, R. K., Walton, K. & Khosroheidari, M. (2018) MiSeq: a next generation sequencing platform for genomic analysis, *Disease Gene Identification* Springer, 223-232.

Reed, G. H., Kent, J. O. & Wittwer, C. T. (2007) High-resolution DNA melting analysis for simple and efficient molecular diagnostics.

Reed, G. H. & Wittwer, C. T. (2004) Sensitivity and specificity of single-nucleotide polymorphism scanning by high-resolution melting analysis. *Clinical chemistry*, 50(10), 1748-1754.

Reginato, A. M. & Olsen, B. R. (2007) Genetics and experimental models of crystal-induced arthritis. Lessons learned from mice and men: is it crystal clear? *Current opinion in rheumatology*, 19(2), 134-145.

Reuss-Borst, M. A., Pape, C. A. & Tausche, A. K. (2014) Hidden gout-Ultrasound findings in patients with musculo-skeletal problems and hyperuricemia. *Springerplus*, 3(1), 592.

Rice, T., Vogler, G. P., Perry, T. S., Laskarzewski, P. M., Province, M. A. & Rao, D. C. (1990) Heterogeneity in the familial aggregation of fasting serum uric acid level in five North American populations: the Lipid Research Clinics Family Study. *Am J Med Genet*, 36(2), 219-25.

Richardson, A. J., Narendran, N., Guymer, R. H., Vu, H. & Baird, P. N. (2006) Blood storage at 4 degrees C-factors involved in DNA yield and quality. *J Lab Clin Med*, 147(6), 290-4.

Richette, P. & Bardin, T. (2010) Gout. *Lancet*, 375(9711), 318-28.

Richette, P., Doherty, M., Pascual, E., Barskova, V., Becce, F., Castaneda, J., Coyfish, M., Guillo, S., Jansen, T., Janssens, H., Liote, F., Mallen, C. D., Nuki, G., Perez-Ruiz, F., Pimentao, J., Punzi, L., Pywell, A., So, A. K., Tausche, A. K., Uhlig, T., Zavada, J., Zhang, W., Tubach, F. & Bardin, T. (2020) 2018 updated European League Against Rheumatism evidence-based recommendations for the diagnosis of gout. *Ann Rheum Dis*, 79(1), 31-38.

Robinson, P. C., Merriman, T. R., Herbison, P. & Highton, J. (2013) Hospital admissions associated with gout and their comorbidities in New Zealand and England 1999-2009. *Rheumatology (Oxford)*, 52(1), 118-26.

Roddy, E. (2011) Revisiting the pathogenesis of podagra: why does gout target the foot? *Journal of foot and ankle research*, 4(1), 13.

Roddy, E. & Doherty, M. (2010) Epidemiology of gout. *Arthritis Res Ther*, 12(6), 223.

Roddy, E., Zhang, W. & Doherty, M. (2007) Are joints affected by gout also affected by osteoarthritis? *Annals of the rheumatic diseases*, 66(10), 1374-1377.

Ronaghi, M., Nygren, M., Lundeberg, J. & Nyrén, P. (1999) Analyses of secondary structures in DNA by pyrosequencing. *Analytical biochemistry*, 267(1), 65-71.

ROSE, B. S. & PRIOR, I. A. (1963) A SURGERY OF RHEUMATISM IN A RURAL NEW ZEALAND MAORI COMMUNITY. *Ann Rheum Dis*, 22, 410-5.

Roubenoff, R., Klag, M. J., Mead, L. A., Liang, K.-Y., Seidler, A. J. & Hochberg, M. C. (1991) Incidence and risk factors for gout in white men. *Jama*, 266(21), 3004-3007.

Roujeau, J.-C., Kelly, J. P., Naldi, L., Rzany, B., Stern, R. S., Anderson, T., Auquier, A., Bastuji-Garin, S., Correia, O. & Locati, F. (1995) Medication use and the risk of Stevens–Johnson syndrome or toxic epidermal necrolysis. *New England Journal of Medicine*, 333(24), 1600-1608.

Rylander-Rudqvist, T., Håkansson, N., Tybring, G. & Wolk, A. (2006) Quality and Quantity of Saliva DNA Obtained from the Self-administrated Oragene Method—A Pilot Study on the Cohort of Swedish Men. *Cancer Epidemiology Biomarkers & Prevention*, 15(9), 1742.

Saksit, N., Nakkam, N., Konyoung, P., Khunarkornsiri, U., Tassaneeyakul, W., Chumworathayi, P., Kanjanawart, S., Sukasem, C., Sangviroon, A. & Pattanacheewapull, O. (2017) Comparison between the HLA-B58: 01 Allele and Single-Nucleotide Polymorphisms in Chromosome 6 for Prediction of Allopurinol-Induced Severe Cutaneous Adverse Reactions. *Journal of Immunology Research*, 2017.

Sambrook, J. (2001) A laboratory manual. *Molecular cloning*, 1.

Samla, G., Gan, K. B. & Then, S.-M. (2017) Modeling microfluidic DNA extraction using superparamagnetic bead particles in COMSOL multiphysics simulation. *Microsystem Technologies*, 23(10), 4435-4440.

Santangelo, M. F., Sciuto, E. L., Lombardo, S. A., Busacca, A. C., Petralia, S., Conoci, S. & Libertino, S. (2015) Si photomultipliers for bio-sensing applications. *IEEE Journal of Selected Topics in Quantum Electronics*, 22(3), 335-341.

Saokaew, S., Tassaneeyakul, W., Maenthaisong, R. & Chaiyakunapruk, N. (2014) Cost-effectiveness analysis of HLA-B\* 5801 testing in preventing allopurinol-induced SJS/TEN in Thai population. *PloS one*, 9(4), e94294.

Satapornpong, P., Jinda, P., Jantararoungtong, T., Koomdee, N., Chaichan, C., Pratoomwun, J., Na Nakorn, C., Aekplakorn, W., Wilantho, A. & Ngamphiw, C. (2020) Genetic Diversity of HLA Class I and Class II Alleles in Thai Populations: Contribution to Genotype-Guided Therapeutics. *Frontiers in Pharmacology*, 11, 78.

Schauer, C., Janko, C., Munoz, L. E., Zhao, Y., Kienhofer, D., Frey, B., Lell, M., Manger, B., Rech, J., Naschberger, E., Holmdahl, R., Krenn, V., Harrer, T., Jeremic, I., Bilyy, R., Schett, G., Hoffmann, M. & Herrmann, M. (2014) Aggregated neutrophil extracellular traps limit inflammation by degrading cytokines and chemokines. *Nat Med*, 20(5), 511-7.

Scheele, C. W. & Beddoes, T. (1786) *The Chemical Essays of Charles-William Scheele*, 1786.

Sciuto, E. L., Coniglio, M. A., Corso, D., van der Meer, J. R., Acerbi, F., Gola, A. & Libertino, S. (2019) Biosensors in Monitoring Water Quality and Safety: An Example of a Miniaturizable Whole-Cell Based Sensor for Hg<sup>2+</sup> Optical Detection in Water. *Water*, 11(10), 1986.

Shaikh Abdul Rahman, S. & Aziz, Z. (2020) Complementary and alternative medicine: Pharmacovigilance in Malaysia and predictors of serious adverse reactions. *J Clin Pharm Ther*.

Shiina, T., Hosomichi, K., Inoko, H. & Kulski, J. K. (2009) The HLA genomic loci map: expression, interaction, diversity and disease. *J Hum Genet*, 54(1), 15-39.

Shiina, T., Suzuki, S., Ozaki, Y., Taira, H., Kikkawa, E., Shigenari, A., Oka, A., Umemura, T., Joshita, S., Takahashi, O., Hayashi, Y., Paumen, M., Katsuyama, Y., Mitsunaga, S., Ota, M., Kulski, J. K. & Inoko, H. (2012) Super high resolution for single molecule-sequence-based typing of classical HLA loci at the 8-digit level using next generation sequencers. *Tissue Antigens*, 80(4), 305-16.

Stamp, L. K. & Chapman, P. T. (2013) Gout and its comorbidities: implications for therapy. *Rheumatology*, 52(1), 34-44.

Stamp, L. K., Day, R. O. & Yun, J. (2016) Allopurinol hypersensitivity: investigating the cause and minimizing the risk. *Nature Reviews Rheumatology*, 12(4), 235.

Stewart, C. A., Horton, R., Allcock, R. J., Ashurst, J. L., Atrazhev, A. M., Coggill, P., Dunham, I., Forbes, S., Halls, K., Howson, J. M., Humphray, S. J., Hunt, S., Mungall, A. J., Osoegawa, K., Palmer, S., Roberts, A. N., Rogers, J., Sims, S., Wang, Y., Wilming, L. G., Elliott, J. F., de

Jong, P. J., Sawcer, S., Todd, J. A., Trowsdale, J. & Beck, S. (2004) Complete MHC haplotype sequencing for common disease gene mapping. *Genome Res*, 14(6), 1176-87.

Strasinger, S. K. & Di Lorenzo, M. S. (2014) *Urinalysis and body fluids* FA Davis.

Stuart, P. E., Nair, R. P., Hiremagalore, R., Kullavanijaya, P., Kullavanijaya, P., Tejasvi, T., Lim, H. W., Voorhees, J. J. & Elder, J. T. (2010) Comparison of MHC class I risk haplotypes in Thai and Caucasian psoriatics shows locus heterogeneity at PSORS1. *Tissue Antigens*, 76(5), 387-397.

Sukasem, C., Jantararoungtong, T., Kuntawong, P., Puangpetch, A., Koomdee, N., Satapornpong, P., Supapsophon, P., Klaewsongkram, J. & Rerkpattanapipat, T. (2016) HLA-B\* 58: 01 for allopurinol-induced cutaneous adverse drug reactions: implication for clinical interpretation in Thailand. *Frontiers in pharmacology*, 7, 186.

Sulaiman, N., Othman, A. Z., Shahril, N. S., Abdul Rashid, A. M. & Md Noh, M. S. F. (2017a) Successful febuxostat desensitization in a patient with febuxostat hypersensitivity: A Malaysian experience, *SAGE Open Med Case Rep*, 2050313x17749080.

Sulaiman, N., Othman, A. Z., Shahril, N. S., Abdul Rashid, A. M. & Md Noh, M. S. F. (2017b) Successful febuxostat desensitization in a patient with febuxostat hypersensitivity: a Malaysian experience. *SAGE open medical case reports*, 5, 2050313X17749080.

Sulaiman, W., Wahida, N., Zuki, M., Zamri, N., Sugathan, S., Chandran, A., Ong, P. & Seung (2019) EPIDEMIOLOGY AND MANAGEMENT OF GOUT PATIENTS ATTENDING RHEUMATOLOGY TERTIARY CENTRE IN PERAK, MALAYSIA.

Sungkanuparph, S., Kiertiburanakul, S., Sura, T., Chantratita, W., Chantarangsu, S., Charoenyingwattana, A., Mahasirimongkol, S., Kubo, M., Mushiroda, T. & Nakamura, Y. (2017) Risk assessment for cutaneous adverse drug reactions from antiretroviral agent. Google Patents.

Talmaci, R., Traeger-Synodinos, J., Kanavakis, E., Coriu, D., Colita, D. & Gavrila, L. (2004) Scanning of beta-globin gene for identification of beta-thalassemia mutation in Romanian population. *J Cell Mol Med*, 8(2), 232-40.

Tan, L. L., Ooi, S. T., Wong, S. M., Ch'ng, C. C., Kwan, Z. & Yong, A. S. W. (2017) An Epidemiological Study of Stevens-Johnson Syndrome (SJS), Toxic Epidermal Necrolysis (TEN) and Stevens-Johnson Syndrome/Toxic Epidermal Necrolysis (SJS/TEN) overlap in University Malaya Medical Centre. *Notice to Authors*, 52.

Tan, P. K., Ostertag, T. M. & Miner, J. N. (2016) Mechanism of high affinity inhibition of the human urate transporter URAT1. *Scientific reports*, 6, 34995.

Tan-Koi, W. C., Sung, C., Chong, Y. Y., Lateef, A., Pang, S. M., Vasudevan, A., Aw, D., Lui, N. L., Lee, S. X. & Ren, E. C. (2017) Tailoring of recommendations to reduce serious cutaneous adverse drug reactions: a pharmacogenomics approach. *Pharmacogenomics*, 18(9), 881-890.

Tana, T. & Vadas, P. (2019) Febuxostat Desensitization in a Patient with Previous Stevens-Johnson Syndrome and HLA-B\*58:01 Genotype, *J Rheumatol. Canada*, 1250-1251.

Tassaneeyakul, W., Jantararoungtong, T., Chen, P., Lin, P.-Y., Tiamkao, S., Khunarkornsiri, U., Chucherd, P., Konyoung, P., Vannaprasaht, S. & Choonhakarn, C. (2009) Strong association between HLA-B\* 5801 and allopurinol-induced Stevens–Johnson syndrome and toxic epidermal necrolysis in a Thai population. *Pharmacogenetics and genomics*, 19(9), 704-709.

Tausche, A.-K., Jansen, T. L., Schröder, H.-E., Bornstein, S. R., Aringer, M. & Müller-Ladner, U. (2009) Gout—current diagnosis and treatment. *Deutsches Ärzteblatt International*, 106(34-35), 549.

Taylor, W. J., Fransen, J., Jansen, T. L., Dalbeth, N., Schumacher, H. R., Brown, M., Louthrenoo, W., Vazquez-Mellado, J., Eliseev, M. & McCarthy, G. (2015) Study for updated gout classification criteria: identification of features to classify gout. *Arthritis care & research*, 67(9), 1304-1315.

Tee, S. H. & Ng, T. G. (2014) A 5 Year Retrospective Study on the Clinical Patterns and Treatment Outcomes of Severe Cutaneous Adverse Drug Reactions (SCARS) in Hospital Tengku Ampuan Rahimah, Malaysia. *Malaysian Journal of Dermatology*, 33, 9-17.

Teh, C. L., Chew, K. F. & Ling, G. R. (2014) Acute Gout in Hospitalized patients in Sarawak General Hospital. *Med J Malaysia*, 69(3), 126-8.

Teng, G. G., Ang, L. W., Saag, K. G., Yu, M. C., Yuan, J. M. & Koh, W. P. (2012) Mortality due to coronary heart disease and kidney disease among middle-aged and elderly men and women with gout in the Singapore Chinese Health Study. *Ann Rheum Dis*, 71(6), 924-8.

Teng, G. G., Tan-Koi, W.-C., Dong, D. & Sung, C. (2020) Is HLA-B\* 58: 01 genotyping cost effective in guiding allopurinol use in gout patients with chronic kidney disease? *Pharmacogenomics*, 21(4), 279-291.

The Sequencing Center. 2022. [online] Available at: <<https://thesequencingcenter.com/knowledge-base/hla-nomenclature/>> [Accessed 12 June 2022].

Then, S.-M., Rani, Z. Z. M., Raymond, A. A. & Jamal, R. (2013) Pharmacogenomics screening of HLA-B\* 1502 in epilepsy patients: How we do it in the UKM Medical Centre, Malaysia. *Neurology Asia*, 18(Supplement 1), 27-29.

Then, S.-M., Rani, Z. Z. M., Raymond, A. A., Ratnaningrum, S. & Rahman, J. (2011) Frequency of the HLA-B\* 1502 allele contributing to carbamazepine-induced hypersensitivity reactions in a cohort of Malaysian epilepsy patients. *Asian Pacific journal of allergy and immunology*, 29(3), 290.

Then, S.-M. & Raymond, A. A. (2019) An update on HLA alleles as pharmacogenetic markers for antiepileptic drug-induced cutaneous adverse reaction. *Neuroscience Research Notes*, 2(2), 1-17.

Thomas, M., Hopkins, C., Duffy, E., Lee, D., Loulergue, P., Ripamonti, D., Ostrov, D. A. & Phillips, E. (2017) Association of the HLA-B\*53:01 Allele With Drug Reaction With Eosinophilia and Systemic Symptoms (DRESS) Syndrome During Treatment of HIV Infection With Raltegravir. *Clin Infect Dis*, 64(9), 1198-1203.

Tohkin, M., Kaniwa, N., Saito, Y., Sugiyama, E., Kurose, K., Nishikawa, J., Hasegawa, R., Aihara, M., Matsunaga, K. & Abe, M. (2013) A whole-genome association study of major determinants for allopurinol-related Stevens–Johnson syndrome and toxic epidermal necrolysis in Japanese patients. *The pharmacogenomics journal*, 13(1), 60-69.

Ton, K. N. T., Cree, S. L., Gronert-Sum, S. J., Merriman, T. R., Stamp, L. K. & Kennedy, M. A. (2018) Multiplexed Nanopore Sequencing of HLA-B Locus in Māori and Pacific Island Samples. *Front Genet*, 9, 152.

Torres, R. J. & Puig, J. G. (2007) Hypoxanthine-guanine phosphoribosyltransferase (HPRT) deficiency: Lesch-Nyhan syndrome. *Orphanet journal of rare diseases*, 2(1), 48.

Towiwat, P. & Li, Z. G. (2015) The association of vitamin C, alcohol, coffee, tea, milk and yogurt with uric acid and gout. *International journal of rheumatic diseases*, 18(5), 495-501.

Traherne, J. A., Horton, R., Roberts, A. N., Miretti, M. M., Hurles, M. E., Stewart, C. A., Ashurst, J. L., Atrazhev, A. M., Coggill, P., Palmer, S., Almeida, J., Sims, S., Wilming, L. G., Rogers, J., de Jong, P. J., Carrington, M., Elliott, J. F., Sawcer, S., Todd, J. A., Trowsdale, J. & Beck, S. (2006) Genetic analysis of completely sequenced disease-associated MHC haplotypes identifies shuffling of segments in recent human history. *PLoS Genet*, 2(1), e9.

Tsai, T.-F. & Yeh, T.-Y. (2010) Allopurinol in dermatology. *American journal of clinical dermatology*, 11(4), 225-232.

Underwood, M. (2006) Diagnosis and management of gout. *Bmj*, 332(7553), 1315-1319.

van der Wouden, C. H., van Rhenen, M. H., Jama, W. O. M., Ingelman-Sundberg, M., Lauschke, V. M., Konta, L., Schwab, M., Swen, J. J. & Guchelaar, H. J. (2019) Development of the PGx-Passport: A Panel of Actionable Germline Genetic Variants for Pre-Emptive Pharmacogenetic Testing. *Clin Pharmacol Ther*, 106(4), 866-873.

Vidal, C., Li, J., Fulton, R. & Fernando, S. L. (2016) A polymorphism within the psoriasis susceptibility 1 candidate 1 (PSORS1C1) gene is not linked to HLA-B\* 58: 01 in an Australian cohort. *Drug metabolism and pharmacokinetics*, 31(3), 252-255.

Wan Rohani, W. T., Mahfudzah, A., Nazihah, M. Y., Tan, H. L., Wan Syamimee, W. G., Amanda Jane, P. G. & Tony Richard, M. (2018) Association of solute carrier family 2, member 9 (SLC2A9) genetic variant rs3733591 with gout in a Malay sample set. *Med J Malaysia*, 73(5), 307-310.

Wang, C., Krishnakumar, S., Wilhelmy, J., Babrzadeh, F., Stepanyan, L., Su, L. F., Levinson, D., Fernandez-Viña, M. A., Davis, R. W., Davis, M. M. & Mindrinos, M. (2012) High-throughput, high-fidelity HLA genotyping with deep sequencing. *Proc Natl Acad Sci U S A*, 109(22), 8676-81.

Wang, W. Y., Tian, W., Zhu, F. M., Liu, X. X., Li, L. X. & Wang, F. (2016) MICA, MICB Polymorphisms and Linkage Disequilibrium with HLA-B in a Chinese Mongolian Population. *Scand J Immunol*, 83(6), 456-62.

Wang, Y., Deng, X., Xu, Y., Ji, L. & Zhang, Z. (2018) Detection of uric acid crystal deposition by ultrasonography and dual-energy computed tomography: A cross-sectional study in patients with clinically diagnosed gout. *Medicine*, 97(42).

Wijnands, J. M., Viechtbauer, W., Thevissen, K., Arts, I. C., Dagnelie, P. C., Stehouwer, C. D., van der Linden, S. & Boonen, A. (2015) Determinants of the prevalence of gout in the general population: a systematic review and meta-regression. *Eur J Epidemiol*, 30(1), 19-33.

Wilk, J. B., Djousse, L., Borecki, I., Atwood, L. D., Hunt, S. C., Rich, S. S., Eckfeldt, J. H., Arnett, D. K., Rao, D. C. & Myers, R. H. (2000) Segregation analysis of serum uric acid in the NHLBI Family Heart Study. *Hum Genet*, 106(3), 355-9.

Williams, T. M. (2001) Human leukocyte antigen gene polymorphism and the histocompatibility laboratory. *J Mol Diagn*, 3(3), 98-104.

Williams, T. M., Wu, J., Foutz, T., McAuley, J. D. & Troup, G. M. (1994) A new DRB1 allele (DRB1\* 0811) identified in Native Americans. *Immunogenetics*, 40(4), 314-314.

Wittig, M., Anmarkrud, J. A., Kässens, J. C., Koch, S., Forster, M., Ellinghaus, E., Hov, J. R., Sauer, S., Schimmler, M. & Ziemann, M. (2015) Development of a high-resolution NGS-based HLA-typing and analysis pipeline. *Nucleic acids research*, 43(11), e70-e70.

Wong, H. B. & Lim, G. H. (2011) Measures of diagnostic accuracy: sensitivity, specificity, PPV and NPV. *Proceedings of Singapore healthcare*, 20(4), 316-318.

Wortmann, R. L., Schumacher, H. R., Becker, M. A. & Ryan, L. M. (2006) *Crystal-induced arthropathies: gout, pseudogout, and apatite-associated syndromes* New York: Taylor & Francis Group.

Wu, R., Cheng, Y.-j., Zhu, L.-l., Yu, L., Zhao, X.-k., Jia, M., Wen, C.-h., Long, X.-z., Tang, T. & He, A.-j. (2016) Impact of HLA-B\* 58: 01 allele and allopurinol-induced cutaneous adverse drug reactions: evidence from 21 pharmacogenetic studies. *Oncotarget*, 7(49), 81870.

Wu, S.-B., Wirthensohn, M. G., Hunt, P., Gibson, J. P. & Sedgley, M. (2008) High resolution melting analysis of almond SNPs derived from ESTs. *Theoretical and Applied Genetics*, 118(1), 1-14.

- Yang, F., Gu, B., Zhang, L., Xuan, J., Luo, H., Zhou, P., Zhu, Q., Yan, S., Chen, S.-a. & Cao, Z. (2014) HLA-B\* 13: 01 is associated with salazosulfapyridine-induced drug rash with eosinophilia and systemic symptoms in Chinese Han population. *Pharmacogenomics*, 15(11), 1461-1469.
- Yap, F. B. B., Wahiduzzaman, M. & Pubalan, M. (2008) Stevens-Johnson syndrome (SJS) and toxic epidermal necrolysis (TEN) in Sarawak: A four years review. *Egyptian Dermatology Online Journal*, 4(1), 1-13.
- Yu, F., Ma, N., Zhang, X., Tian, S., Geng, L., Xu, W., Wang, M., Jia, Y., Liu, X. & Ma, J. (2018) Comprehensive investigating of cytokine and receptor related genes variants in patients with chronic hepatitis B virus infection. *Cytokine*, 103, 10-14.
- Yu, F., Zhang, X., Tian, S., Geng, L., Xu, W., Ma, N., Wang, M., Jia, Y., Liu, X. & Ma, J. (2017) Comprehensive investigation of cytokine-and immune-related gene variants in HBV-associated hepatocellular carcinoma patients. *Bioscience reports*, 37(6).
- Yu, K. H. & Luo, S. F. (2003) Younger age of onset of gout in Taiwan. *Rheumatology (Oxford)*, 42(1), 166-70.
- Zalokar, J., Lellouch, J., Claude, J. R. & Kuntz, D. (1972) Serum uric acid in 23,923 men and gout in a subsample of 4257 men in France. *J Chronic Dis*, 25(5), 305-12.
- Zhang, B., Fang, W., Zeng, X., Zhang, Y., Ma, Y., Sheng, F. & Zhang, X. (2016) Clinical characteristics of early- and late-onset gout: A cross-sectional observational study from a Chinese gout clinic. *Medicine*, 95(47), e5425-e5425.
- Zhang, W., Doherty, M., Bardin, T., Pascual, E., Barskova, V., Conaghan, P., Gerster, J., Jacobs, J., Leeb, B. & Liote, F. (2006) EULAR evidence based recommendations for gout. Part II: Management. Report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Annals of the rheumatic diseases*, 65(10), 1312-1324.
- Zhang, X., Ma, H., Hu, C., Yu, B., Ma, W., Wu, Z., Luo, X., Zou, H. & Guan, M. (2015) Detection of HLA-B\*58:01 with TaqMan assay and its association with allopurinol-induced sCADR. *Clin Chem Lab Med*, 53(3), 383-90.
- Zhou, L., Vandersteen, J., Wang, L., Fuller, T., Taylor, M., Palais, B. & Wittwer, C. T. (2004) High-resolution DNA melting curve analysis to establish HLA genotypic identity. *Tissue antigens*, 64(2), 156-164.

## APPENDICES

### Appendix A- Consent form for volunteers

#### PATIENT INFORMATION SHEET

**Title:**

Developing a High Molecular Melting (HRM) method for detection of the HLA-B\*5801 allele to prevent allopurinol-induced Steven-Johnson Syndrome.

**Purpose:**

To develop the method to detect the allele HLA-B\*5801 in patients before giving allopurinol to prevent Steven Johnson Syndrome.

**Introduction:**

Allopurinol is a drug that decreases the production of uric acid. It is most commonly used in the treatment of gout and high levels of uric acid. It is also used to prevent or treat uric acid kidney stones and chronic kidney diseases. However, a small number of patients carry the risk of getting a skin hypersensitivity reaction that is both painful and dangerous called Steven Johnson Syndrome (SJS) and toxic epidermal necrolysis (TEN) if they take allopurinol. Research has shown that patients carrying the HLA-B\*5801 gene are more at risk of getting SJS/TEN caused by allopurinol. This gene/allele is more common among Asian people, especially people of Korean, Han Chinese and Thai descent. Hence, individuals from these populations might have a higher risk of getting SJS/TEN caused by allopurinol.

In order to reduce the risk of patients getting the hypersensitivity reaction, screening should be carried out before prescribing allopurinol. By developing a rapid method to detect the HLA-B\*5801 gene, patients testing positive can be identified and alternatives to allopurinol given. Therefore, the project aims to develop and optimize a method using High Resolution Melting (HRM) which is based on PCR to detect the HLA-B\*5801 allele. This involves collection of blood from consented participants to obtain the DNA for the method optimization. The participants will be informed of the nature of the study, and the confidentiality of the participants will be maintained.

**What will be done?**

5 to 10 ml of blood will be taken like that which is carried out during routine treatment. No additional procedure will be carried out. The sample will be processed and analyzed in a lab. The samples will be coded and no personal information will be attached to the samples. Any excess specimens would

also be stored in the BIOBANK PPUKM-UMBI for future use. Future use of these samples for other studies would be anonymized.

**Benefits of the study:**

The screening method that will be developed will help doctor identify whether the patient carries the HLA-B\*5801 allele and avoid giving allopurinol to prevent skin hypersensitivity reaction caused by allopurinol.

.

**Risk:**

No additional risks apart from that which is involved in routine blood taking procedures as no additional procedures are carried out.

**Data handling/ Confidentiality:**

The data and results of the study will be reported in a collected manner with no reference to a specific individual. Hence the results of the study will be kept as private and confidential and only you and the researcher will be allowed to know the results of the test. Personal information will not be attached to the sample.

**Do I need to take part?**

Your participation in this study is entirely voluntary. If you agree, the study will not affect your relationship with your doctor and you do not have to give any reasons for participating. You are also free to withdraw from the study without giving a reason at any point of time.

**What if I am tested positive?**

Patients who are tested positive in the study will be notified so that precautions can be taken to prevent Steven Johnson Syndrome caused by allopurinol use.

**Payment/Reimbursement:**

There will no payment charged or paid to you for participating in this study. In the event that this study results in the development of any marketable products, you will have no ownership interest in the product and no right to share in any profits from its commercialization whatever.

**If you have any questions, please contact:** Prof. Dr. Mohd Shahrir Bin Mohamed Said:  
[drobiwan@gmail.com](mailto:drobiwan@gmail.com), Dr Then Sue-Mian: [then.sue-mian@nottingham.edu.my](mailto:then.sue-mian@nottingham.edu.my)

**CONFIDENTIAL**

**CONSENT FORM FOR ADULT PATIENT**

**Title of project:**

Developing a method for screening of HLA-B\*5801 allele to prevent hypersensitivity reaction caused by allopurinol.

**Consent:**

I have read the information on the research project stated above and have also been given the explanation by a doctor about the purpose of this document. I understand the objectives of keeping my blood for storage and research purpose. I also understand that I retained the absolute right over the specimens and may withdraw at any time. I also have the right to know about the research conducted on my blood including information on the results of the research.

I \_\_\_\_\_ IC number: \_\_\_\_\_

\* agree / disagree for my blood be kept and used for medical research including genetic research.

I \*would like to know / don't want to know the result of the genetic analysis or other research which will be performed on the samples.

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

Witness

Medical officer

Name :

Name :

IC :

IC :

Signature :

Signature :

Date :

Date :

**CONFIDENTIAL**

**CONSENT FORM FOR CHILDREN PATIENT**

**Title of project:**

Developing a method for screening of HLA-B\*5801 allele to prevent hypersensitivity reaction caused by allopurinol.

**Consent:**

I have read the information on the research project stated above and have also been given the explanation by a doctor about the purpose of this document. I understand the objectives of keeping my blood for storage and research purpose. I also understand that I retained the absolute right over the specimens and may withdraw at any time. I also have the right to know about the research conducted on our child blood including information on the results of the research.

I \_\_\_\_\_ IC number: \_\_\_\_\_

\* agree / disagree for my child blood be kept and used for medical research including genetic research.

I \*would like to know / don't want to know the result of the genetic analysis or other research which will be performed on the samples.

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

Witness

Medical officer

Name :

Name :

IC :

IC :

Signature :

Signature :

Date :

Date :

**CONFIDENTIAL**

**QUESTIONNAIRE**

**Blood sample collection for developing the method for detecting the HLA-B\*5801 gene in patients to prevent allopurinol associated hypersensitivity reaction.**

**Investigators:** Dr Then Sue-Mian & Prof. Dr. Mohd Shahrir Bin Mohamed Said:

The participant should complete the whole of this sheet himself/herself. Please cross out as necessary:

• Have you read and understood the participant information sheet?	YES/NO
---	--------

• Have you had the opportunity to ask questions and discuss the study?	YES/NO
• Have all the questions been answered satisfactorily?	YES/NO
• Have you received enough information about the study?	YES/NO
• Do you understand that you are free to withdraw from the study:	
- at any time?	YES/NO
- without having to give a reason?	YES/NO
• Do you agree to take part in the study?	YES/NO

**“This study has been explained to me to my satisfaction, and I agree to take part. I understand that I am free to at any time.”**

Signature of the Participant: \_\_\_\_\_ Date: \_\_\_\_\_

Name (in block capitals): \_\_\_\_\_

I have explained the study to the above participant and he/she has agreed to take part.

Signature of the Researcher: \_\_\_\_\_ Date: \_\_\_\_\_

Name (in block capitals): \_\_\_\_\_

## Appendix B- Full 290 samples details

Table 1- Full details for the 80 Pusat Perubatan Universiti Kebangsaan Malaysia (PPUKM) patients, labelled as HLAG1 to HLAG80. Abbreviations were used for Allopurinol (ALLO), Febuxostat (FEB) and Colchicine (CLC) in the drugs column.

HLA Code	Concentration (ng/ $\mu$ L)	Purity (260/280 nm)	Age	Race	Gender	Drugs	Hypersensitivity observed
<b>HLAG 1</b>	359.1	1.94	60	Indian	Male	ALLO	No ADRs noted
<b>HLAG 2</b>	1564.1	1.91	75	Malay	Male	ALLO	No ADRs noted
<b>HLAG 3</b>	223.5	1.93	74	Malay	Male	ALLO	No ADRs noted
<b>HLAG 4</b>	640	1.86	26	Malay	Male	ALLO	No ADRs noted
<b>HLAG 5</b>	608.4	1.86	61	Malay	Male	ALLO	No ADRs noted
<b>HLAG 6</b>	625.9	1.91	71	Malay	Male	ALLO	No ADRs noted
<b>HLAG 7</b>	307.2	1.9	65	Malay	Male	ALLO	No ADRs noted
<b>HLAG 8</b>	1452.1	1.86	46	Malay	Male	ALLO	No ADRs noted
<b>HLAG 9</b>	1868.8	1.86	69	Chinese	Male	ALLO	No ADRs noted
<b>HLAG 10</b>	1269.5	1.91	28	Malay	Male	ALLO	No ADRs noted
<b>HLAG 11</b>	759.8	1.94	73	Chinese	Male	ALLO	No ADRs noted
<b>HLAG 12</b>	1188.4	1.92	81	Malay	Male	ALLO, FEB	Vasculitis rash
<b>HLAG 13</b>	2858.1	1.89	63	Malay	Male	ALLO, FEB	mild allergy- itchiness
<b>HLAG 14</b>	1782.6	1.86	65	Chinese	Male	ALLO	No ADRs noted
<b>HLAG 15</b>	1334.4	1.9	69	Malay	Male	ALLO	No ADRs noted
<b>HLAG 16</b>	2709.8	1.87	69	Chinese	Female	ALLO	No ADRs noted
<b>HLAG 17</b>	3851.5	1.89	55	Malay	Female	ALLO	No ADRs noted
<b>HLAG 18</b>	1263.1	1.92	45	Malay	Male	ALLO, FEB	mild allergy- rash
<b>HLAG 19</b>	848.8	1.9	32	Malay	Male	ALLO, FEB	mild allergy- itchiness
<b>HLAG 20</b>	2016.7	1.89	46	Malay	Male	ALLO	No ADRs noted

<b>HLAG 21</b>	1571.6	1.92	40	Malay	Male	ALLO	No ADRs noted
<b>HLAG 22</b>	1380.7	1.04	54	Malay	Male	ALLO	No ADRs noted
<b>HLAG 23</b>	1253.6	1.94	55	Chinese	Male	ALLO	No ADRs noted
<b>HLAG 24</b>	2257.9	1.91	43	Malay	Male	ALLO	No ADRs noted
<b>HLAG 25</b>	2408.5	1.91	36	Malay	Male	ALLO	No ADRs noted
<b>HLAG 26</b>	1333	1.91	32	Malay	Male	ALLO, FEB	Exfoliative dermatitis
<b>HLAG 27</b>	3725.2	1.82	68	Malay	Male	ALLO	No ADRs noted
<b>HLAG 28</b>	911.4	1.94	64	Malay	Female	ALLO, FEB	itchiness
<b>HLAG 29</b>	2079.4	1.93	68	Chinese	Male	ALLO, FEB	DRESS Syndrome
<b>HLAG 30</b>	354	1.95	33	Malay	Male	ALLO	No ADRs noted
<b>HLAG 31</b>	404.6	1.97	82	Malay	Male	ALLO	No ADRs noted
<b>HLAG 32</b>	227.1	2.00	64	Malay	Male	ALLO	No ADRs noted
<b>HLAG 33</b>	3376.7	1.81	46	Chinese	Male	ALLO	No ADRs noted
<b>HLAG 34</b>	708.1	1.9	69	Chinese	Male	ALLO	No ADRs noted
<b>HLAG 35</b>	2199.8	1.81	61	Malay	Male	ALLO	No ADRs noted
<b>HLAG 36</b>	2921.2	1.88	60	Malay	Male	ALLO	No ADRs noted
<b>HLAG 37</b>	1314.1	1.94	76	Chinese	Male	ALLO	No ADRs noted
<b>HLAG 38</b>	1101.3	1.97	72	Malay	Male	ALLO, FEB, CLC	Stevens Johnson Syndrome
<b>HLAG 39</b>	3801.6	1.85	67	Others	Female	ALLO, FEB	Stevens Johnson Syndrome
<b>HLAG 40</b>	3153.8	1.89	75	Chinese	Male	ALLO	No ADRs noted
<b>HLAG 41</b>	4365.6	1.75	38	Chinese	Male	ALLO	No ADRs noted
<b>HLAG 42</b>	4240.1	1.83	34	Malay	Male	ALLO	No ADRs noted
<b>HLAG 43</b>	251.5	1.89	69	Chinese	Male	ALLO	No ADRs noted
<b>HLAG 44</b>	3334.5	1.88	62	Malay	Male	ALLO	No ADRs noted
<b>HLAG 45</b>	1306.7	1.86	71	Chinese	Female	ALLO	No ADRs noted
<b>HLAG 46</b>	500	1.8	53	Malay	Male	ALLO, FEB, CLC	ALLO hypersensitivity syndrome-severe itchiness and red eye.
<b>HLAG 47</b>	1395.4	1.88	73	Malay	Male	ALLO, CLC	No ADRs noted

<b>HLAG 48</b>	1271.7	1.87	52	Malay	Male	ALLO, FEB, CLC	Allergic to ALLO
<b>HLAG 49</b>	1266.9	1.93	31	Malay	Male	ALLO	No ADRs noted
<b>HLAG 50</b>	1293.4	1.9	57	Malay	Female	ALLO, CLC	No ADRs noted
<b>HLAG 51</b>	2813.6	1.93	64	Malay	MALE	ALLO	No ADRs noted
<b>HLAG 52</b>	1085.6	1.85	63	Malay	Male	ALLO	No ADRs noted
<b>HLAG 53</b>	970.8	1.94	73	Malay	Male	ALLO	Rashes
<b>HLAG 54</b>	3512.54	1.87	56	Chinese	Male	ALLO	No ADRs noted
<b>HLAG 55</b>	1839	1.89	33	Malay	Male	ALLO	No ADRs noted
<b>HLAG56</b>	1452.2	1.9	64	Malay	Male	ALLO	DRESS syndrome and rash
<b>HLAG57</b>	1242.03	1.92	82	Malay	Male	ALLO, CLC	No ADRs noted
<b>HLAG58</b>	1083.67	1.93	57	Malay	Male	ALLO, CLC	No ADRs noted
<b>HLAG59</b>	1625.58	1.91	39	Malay	Male	ALLO, CLC	No ADRs noted
<b>HLAG60</b>	1528.88	1.93	53	Malay	Male	ALLO, FEB	ALLO doesn't work, changed to FEB, bilateral renal calculi (renal stone)
<b>HLAG61</b>	3659	1.84	56	Chinese	Male	ALLO, CLC	No ADRs noted
<b>HLAG62</b>	2807.78	1.9	75	Chinese	Male	ALLO, CLC	No ADRs noted
<b>HLAG63</b>	4055.94	1.8	59	Chinese	Male	ALLO, CLC	No ADRs noted
<b>HLAG64</b>	3598.44	1.84	66	Malay	Male	ALLO, FEB, CLC	Allergic to allopurinol, rash
<b>HLAG65</b>	2449.51	1.87	54	Chinese	Male	ALLO, CLC	No ADRs noted
<b>HLAG66</b>	1392.48	1.88	80	Malay	Male	ALLO, FEB	Allergic to allopurinol, itchiness
<b>HLAG67</b>	4473.8	1.8	61	Malay	Male	ALLO, CLC	No ADRs noted
<b>HLAG68</b>	2224.24	1.91	68	Malay	Male	ALLO, FEB	Allergic to allopurinol, anaphylactic shock
<b>HLAG69</b>	961.77	1.87	49	Malay	Male	ALLO, FEB, CLC	Allergic to allopurinol
<b>HLAG70</b>	1293.12	1.97	58	Malay	Male	ALLO	No ADRs noted
<b>HLAG71</b>	1555.86	1.88	55	Malay	Male	ALLO, CLC	No ADRs noted

<b>HLAG72</b>	2095.8	1.91	35	Malay	Male	ALLO, CLC	No ADRs noted
<b>HLAG 73</b>	3567.6	1.85	76	Chinese	Female	ALLO, CLC	No ADRs noted
<b>HLAG 74</b>	2467.7	1.92	44	Malay	Male	ALLO, CLC	No ADRs noted
<b>HLAG 75</b>	2500.7	1.85	67	Malay	Male	ALLO, CLC	No ADRs noted
<b>HLAG 76</b>	2206.8	1.95	47	Malay	Male	ALLO, CLC	No ADRs noted
<b>HLAG 77</b>	2858.4	1.88	52	Malay	Male	ALLO, CLC	No ADRs noted
<b>HLAG 78</b>	2136.6	1.92	70	Malay	Male	ALLO, FEB	No ADRs noted
<b>HLAG 79</b>	3486.2	1.89	79	Malay	Male	ALLO	No ADRs noted
<b>HLAG 80</b>	3289.8	1.8	36	Malay	Male	ALLO, CLC	No ADRs noted

Table 2- Full details for the 65 Hospital Putrajaya patients, labelled as HLAP1 to HLAP65. Abbreviations were used for Allopurinol (ALLO), Febuxostat (FEB) and Colchicine (CLC) in the drugs column.

<b>HLA code</b>	<b>Concentration (ng/ <math>\mu</math>L)</b>	<b>Purity (260/280 nm)</b>	<b>Age</b>	<b>Race</b>	<b>Gender</b>	<b>Disease</b>	<b>Drugs</b>	<b>Hypersensitivity observed</b>
<b>HLAP1</b>	2657.0	1.82	61	Chinese	Male	Gout	ALLO	No ADRs noted.
<b>HLAP2</b>	1892.5	1.81	60	Malay	Male	Gout	ALLO	Generalized rash. Undergoing desensitization. Allergy to allopurinol.
<b>HLAP3</b>	1883.6	1.81	63	Malay	Male	Gout	ALLO, FEB	Allopurinol and febuxostat allergy, generalized rash and facial swelling.
<b>HLAP4</b>	692.50	1.78	74	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP5</b>	2405.5	1.90	33	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP6</b>	2498.2	1.85	39	Malay	Male	Gout	ALLO	SJS in mouth and genitalia, skin and chest, ulceration in lips.
<b>HLAP7</b>	3239.5	1.95	29	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP8</b>	2757.8	1.80	60	Malay	Male	Gout	ALLO, CLC	No ADRs noted.
<b>HLAP9</b>	2715.2	1.94	32	Malay	Male	Gout	ALLO	No ADRs noted.

<b>HLAP10</b>	2734.8	1.93	68	Malay	Female	Gout	ALLO	No ADRs noted.
<b>HLAP11</b>	2197.7	1.87	59	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP12</b>	2668.8	1.82	53	Indian	Male	Gout	ALLO	No ADRs noted.
<b>HLAP13</b>	2532.0	1.95	50	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP14</b>	2588.4	1.90	46	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP15</b>	2539.8	1.90	32	Malay	Male	Gout	ALLO, CLC	No ADRs noted.
<b>HLAP16</b>	3266.9	1.87	63	Chinese	Male	Gout	ALLO, CLC	No ADRs noted.
<b>HLAP17</b>	1758.3	1.89	58	Malay	Male	Gout	ALLO, CLC	No ADRs noted.
<b>HLAP18</b>	2163.6	1.91	27	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP19</b>	1543.6	1.83	63	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP20</b>	2224.8	1.80	63	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP21</b>	1438.6	1.90	80	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP22</b>	2291.5	1.92	54	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP23</b>	2163.8	1.88	56	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP24</b>	1888.4	1.83	29	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP25</b>	2688.3	1.82	42	Malay	Male	Gout	ALLO, FEB	Allergy to Allopurinol. Generalized skin rash.
<b>HLAP26</b>	1314.1	1.94	66	Chinese	Male	Gout	ALLO	No ADRs noted.
<b>HLAP27</b>	1117.6	1.84	58	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP28</b>	507.50	1.80	44	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP29</b>	1399.0	1.83	65	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP30</b>	1703.3	1.80	77	Chinese	Male	Gout	ALLO	No ADRs noted.
<b>HLAP31</b>	1766.3	1.85	65	Malay	Female	Gout	ALLO	No ADRs noted.
<b>HLAP32</b>	1041.5	1.84	55	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP33</b>	1824.5	1.82	53	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP34</b>	1282.0	1.88	33	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP35</b>	1041.7	1.94	49	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP36</b>	1653.6	1.92	70	Malay	Male	Gout	ALLO	No ADRs noted.
<b>HLAP37</b>	2964.5	1.86	58	Malay	Male	Gout	ALLO	Allergic to Allopurinol, generalised rash.
<b>HLAP38</b>	3308.5	1.89	58	Malay	Male	Gout	ALLO	No ADRs noted.

<b>HLAP39</b>	1768.9	1.80	54	Chinese	Male	Gout	ALLO	No ADRs noted
<b>HLAP40</b>	2461.8	1.84	38	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP41</b>	2859.8	1.85	53	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP42</b>	2625.8	1.92	33	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP43</b>	2129.9	1.91	64	Malay	Female	Gout	ALLO	No ADRs noted
<b>HLAP44</b>	2789.3	1.89	22	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP45</b>	3113.0	1.91	35	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP46</b>	2875.7	1.87	37	Malay	Female	Gout	ALLO	rash on face after 2 weeks of using allopurinol.
<b>HLAP47</b>	2876.9	1.88	30	Chinese	Male	Gout	ALLO	No ADRs noted
<b>HLAP48</b>	2879.4	1.84	68	Chinese	Male	Gout	ALLO	No ADRs noted
<b>HLAP49</b>	3050.9	1.95	42	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP50</b>	2762.4	1.92	67	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP51</b>	2335.0	1.83	66	Malay	Male	Gout	ALLO	Exfoliative dermatitis and allergy to febuxostat
<b>HLAP52</b>	2059.6	1.85	40	Malay	Male	Gout	ALLO	Allergy to Allopurinol-Transaminitis, no rash seen.
<b>HLAP53</b>	1848.0	1.80	75	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP54</b>	2637.7	1.84	55	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP55</b>	2323.6	1.83	30	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP56</b>	3172.6	1.86	41	Malay	Male	Gout	ALLO	Allergic to allopurinol, generalised rashes and macular papular rash.
<b>HLAP57</b>	3238.0	1.90	62	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP58</b>	3217.8	1.82	45	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP59</b>	2847.2	1.80	35	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP60</b>	2737.0	1.83	62	Chinese	Male	Gout	ALLO	Allergic to allo, generalised rashes
<b>HLAP61</b>	2973.1	1.93	67	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP62</b>	2188.9	1.88	27	Malay	Male	Gout	ALLO	No ADRs noted
<b>HLAP63</b>	2136.7	1.83	68	Malay	Male	Gout	ALLO	No ADRs noted

<b>HLAP64</b>	3196.3	1.86	55	Malay	Male	Gout	ALLO	Severe allergy to allopurinol, angioedema and swelling of the eye.
<b>HLAP65</b>	3239.6	1.95	30	Malay	Male	Gout	ALLO	No ADRs noted

Table 3- Full details for the 145 healthy volunteers samples from UMBI's biobank, labelled either as N or NS, depending on their UMBI codes.

<b>Normal sample code</b>	<b>Concentration (ng/ <math>\mu</math>L)</b>	<b>Purity (260/280 nm)</b>	<b>Age</b>	<b>Race</b>	<b>Gender</b>
<b>N1</b>	523.5	1.80	54	Malay	Female
<b>N3</b>	918.3	1.79	63	Malay	Male
<b>N4</b>	689.4	1.89	60	Malay	Male
<b>N5</b>	711.4	1.88	57	Malay	Female
<b>N6</b>	568.0	1.97	63	Malay	Female
<b>N8</b>	746.3	1.67	40	Malay	Female
<b>N9</b>	753.5	1.89	55	Malay	Female
<b>N12</b>	1217.4	1.89	70	Malay	Female
<b>N11</b>	781.4	1.86	70	Malay	Male
<b>N16</b>	576.8	1.79	57	Malay	Female
<b>N17</b>	1482.3	1.87	54	Malay	Male
<b>N18</b>	864.4	1.84	77	Malay	Male
<b>N19</b>	828.4	1.86	59	Malay	Female
<b>N20</b>	696.0	1.97	77	Malay	Male
<b>N22</b>	837.5	1.84	64	Malay	Male
<b>N24</b>	982.3	1.87	63	Malay	Female
<b>N26</b>	1041.4	1.85	53	Malay	Male
<b>N27</b>	548.3	1.90	55	Malay	Female
<b>N28</b>	395.3	1.90	63	Malay	Female

<b>Ns3</b>	1040.5	1.93	72	Malay	Female
<b>Ns4</b>	980.5	1.80	55	Malay	Male
<b>Ns5</b>	317.4	1.91	40	Malay	Male
<b>Ns6</b>	345.7	1.83	44	Malay	Female
<b>Ns11</b>	570.2	1.88	43	Malay	Male
<b>N33</b>	400.3	1.93	34	Malay	Male
<b>N34</b>	1211.7	1.93	31	Malay	Female
<b>Ns12</b>	687.3	1.89	57	Indian	Female
<b>Ns14</b>	860.2	1.86	43	Chinese	Female
<b>Ns15</b>	530.2	1.76	57	Malay	Male
<b>Ns17</b>	2100.8	1.82	45	Malay	Male
<b>Ns18</b>	641.4	1.80	60	Malay	Male
<b>Ns19</b>	758.4	1.79	39	Malay	Male
<b>Ns20</b>	2000.0	1.89	36	Malay	Male
<b>Ns21</b>	2065.8	1.88	40	Malay	Female
<b>Ns22</b>	754.3	1.97	34	Malay	Female
<b>Ns23</b>	691.3	1.67	57	Malay	Male
<b>Ns24</b>	963.5	1.89	56	Malay	Female
<b>Ns25</b>	1114.0	1.89	62	Malay	Female
<b>Ns26</b>	2000	1.86	58	Malay	Female
<b>Ns27</b>	1000	1.79	63	Malay	Female
<b>Ns28</b>	2171.8	1.87	56	Malay	Female
<b>Ns29</b>	844.5	1.84	43	Malay	Male
<b>Ns30</b>	327.8	1.86	55	Malay	Female
<b>Ns31</b>	640.4	1.97	35	Malay	Male
<b>Ns34</b>	1264.5	1.84	34	Malay	Female
<b>Ns35</b>	1191.9	1.87	58	Malay	Female
<b>Ns37</b>	430.3	1.85	58	Malay	Female
<b>Ns39</b>	246.7	1.90	35	Malay	Male

<b>Ns40</b>	398.6	1.90	34	Malay	Male
<b>Ns41</b>	495.4	1.93	45	Malay	Male
<b>Ns42</b>	986.9	1.80	45	Malay	Female
<b>Ns46</b>	259.0	1.91	55	Malay	Male
<b>N35</b>	1182.8	1.83	45	Malay	Female
<b>N39</b>	877.2	1.88	40	Malay	Male
<b>N40</b>	585.2	1.93	35	Malay	Female
<b>Ns67</b>	379.3	1.93	45	Chinese	Female
<b>Ns68</b>	817.2	1.89	61	Malay	Female
<b>N43</b>	1135.9	1.86	52	Malay	Female
<b>Ns70</b>	460.1	1.76	63	Malay	Female
<b>Ns72</b>	532.9	1.82	53	Malay	Female
<b>Ns75</b>	1558.4	1.80	57	Chinese	Female
<b>N46</b>	1264.2	1.79	40	Malay	Male
<b>N47</b>	2551.5	1.89	41	Malay	Female
<b>N48</b>	902.66	1.88	57	Chinese	Male
<b>Ns78</b>	456.01	1.97	73	Indian	Female
<b>N49</b>	495.20	1.67	34	Malay	Female
<b>Ns79</b>	563.6	1.89	70	Chinese	Male
<b>Ns80</b>	620.6	1.89	76	Chinese	Female
<b>N50</b>	977.3	1.86	36	Malay	Male
<b>N51</b>	2567.9	1.79	38	Malay	Female
<b>N52</b>	1222.8	1.87	42	Malay	Male
<b>Ns81</b>	1498.6	1.84	72	Chinese	Female
<b>Ns92</b>	898.6	1.86	53	Chinese	Female
<b>Ns95</b>	661.6	1.97	62	Chinese	Male
<b>N53</b>	930.3	1.84	27	Chinese	Female
<b>N54</b>	1500.0	1.87	27	Chinese	Male
<b>N55</b>	310.6	1.85	27	Chinese	Male

<b>N56</b>	2000.0	1.90	27	Chinese	Male
<b>N35</b>	539.7	1.90	45	Malay	Female
<b>N39</b>	545.8	1.93	40	Malay	Male
<b>N58</b>	1330.3	1.80	40	Malay	Female
<b>N59</b>	304.1	1.91	41	Malay	Female
<b>Ns101</b>	707.4	1.83	67	Malay	Male
<b>Ns298</b>	1281.1	1.88	54	Chinese	Female
<b>Ns104</b>	2273.6	1.93	56	Malay	Female
<b>Ns105</b>	1134.0	1.93	67	Chinese	Female
<b>Ns106</b>	1423.0	1.89	70	Chinese	Male
<b>Ns107</b>	2384.5	1.86	76	Chinese	Female
<b>Ns108</b>	759.8	1.76	81	Chinese	Female
<b>Ns109</b>	862.5	1.82	79	Chinese	Male
<b>Ns110</b>	1296.5	1.80	69	Chinese	Female
<b>Ns296</b>	2669.9	1.79	40	Malay	Male
<b>Ns300</b>	1681.5	1.89	38	Malay	Male
<b>Ns301</b>	764.0	1.88	57	Malay	Male
<b>Ns304</b>	1922.1	1.97	76	Malay	Male
<b>N63</b>	1033.4	1.67	22	Malay	Male
<b>Ns33</b>	984.9	1.89	49	Malay	Female
<b>Ns45</b>	232.8	1.89	47	Malay	Male
<b>Ns13</b>	367.5	1.86	48	Chinese	Male
<b>Ns299</b>	976.5	1.79	46	Malay	Male
<b>N2</b>	844.9	1.87	50	Malay	Female
<b>N13</b>	802.7	1.84	49	Malay	Female
<b>Ns81</b>	847.6	1.86	49	Malay	Female
<b>N30</b>	590.3	1.97	50	Malay	Male
<b>N38</b>	258.7	1.84	48	Malay	Male
<b>N37</b>	479.0	1.80	48	Malay	Female

<b>N15</b>	605.7	1.78	51	Malay	Female
<b>N36</b>	678.8	1.89	50	Malay	Male
<b>Ns99</b>	633.0	1.88	49	Chinese	Male
<b>Ns32</b>	852.9	1.98	52	Malay	Female
<b>Ns44</b>	566.8	1.68	52	Malay	Female
<b>N10</b>	1000.5	1.90	53	Malay	Female
<b>Ns76</b>	578.5	1.88	63	Chinese	Male
<b>Ns103</b>	1280.4	1.85	62	Malay	Male
<b>Ns305</b>	1370.3	1.78	61	Malay	Male
<b>N29</b>	959.1	1.87	65	Malay	Female
<b>N25</b>	426.3	1.83	65	Malay	Male
<b>N62</b>	270.3	1.87	65	Chinese	Male
<b>Ns93</b>	302.2	1.98	64	Chinese	Male
<b>Ns303</b>	255.3	1.81	63	Malay	Female
<b>Ns71</b>	571.1	1.87	65	Malay	Male
<b>Ns74</b>	759.8	1.87	65	Chinese	Female
<b>N14</b>	1034.5	1.91	67	Malay	Female
<b>Ns49</b>	1084.7	1.90	67	Chinese	Male
<b>Ns82</b>	834.0	1.93	66	Others	Male
<b>Ns73</b>	1019.6	1.80	67	Malay	Female
<b>Ns302</b>	961.9	1.91	66	Chinese	Female
<b>Ns97</b>	1478.0	1.88	68	Chinese	Male
<b>Ns306</b>	848.0	1.87	57	Chinese	Female
<b>N71</b>	398.7	1.94	26	Malay	Female
<b>N70</b>	549.9	1.93	26	Malay	Female
<b>N69</b>	543.9	1.89	43	Malay	Male
<b>N68</b>	668.9	1.84	27	Malay	Male
<b>N67</b>	979.4	1.77	27	Malay	Female
<b>N66</b>	768.6	1.81	30	Malay	Female

<b>N76</b>	835.0	1.80	1	Others	Female
<b>N64</b>	686.8	1.78	27	Malay	Male
<b>Ns322</b>	821.7	1.89	72	Malay	Female
<b>Ns1</b>	721.4	1.88	64	Malay	Male
<b>N45</b>	408.07	1.98	47	Others	Male
<b>N44</b>	603.0	1.68	40	Others	Female
<b>N72</b>	402.9	1.90	26	Chinese	Male
<b>Ns307</b>	591.9	1.88	57	Indian	Male
<b>NS301</b>	935.7	1.85	50	Malay	Female
<b>N78</b>	986.5	1.78	60	Malay	Male

## Appendix C- PCR and Primers

### 1) Area targeted by P2 and P3 on HLA-B\*58:01

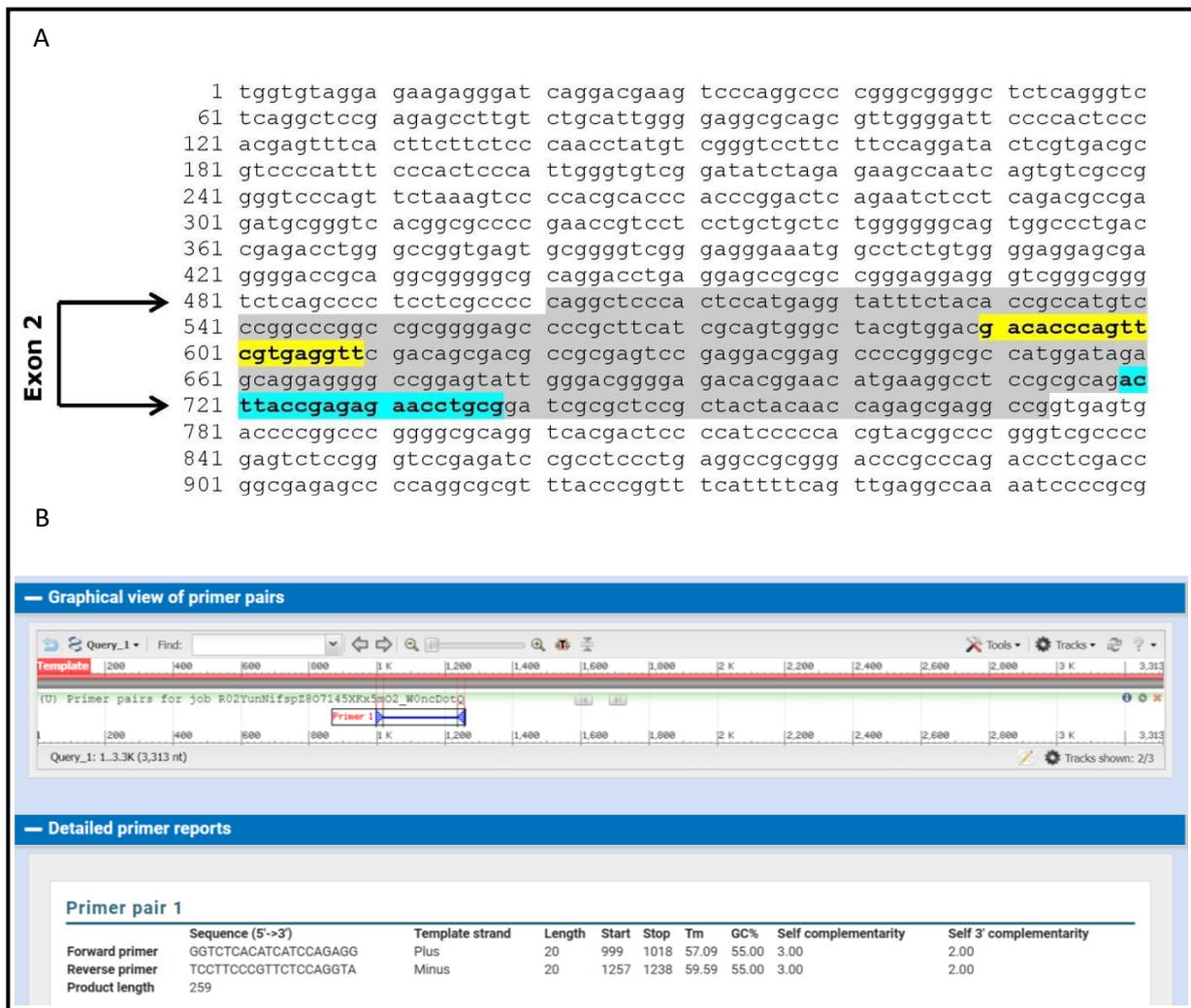
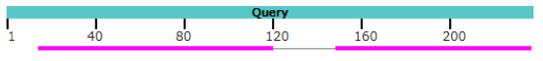
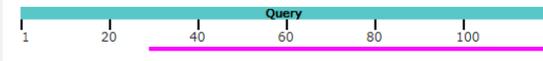
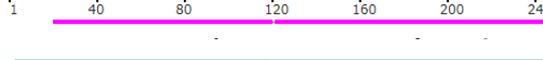
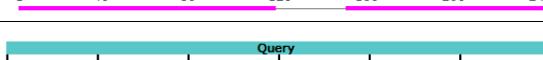
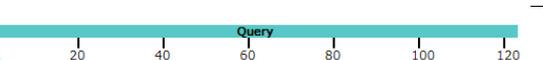
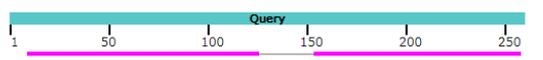
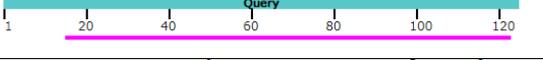
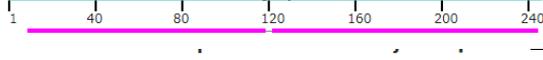
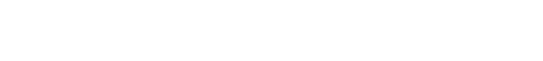


Figure 1- Figure 1-A The partial DNA sequence of the HLA-B\*58:01 allele is shown, along with its exon 2 highlighted in grey. Forward primers are highlighted in yellow, while reverse primers are highlighted in blue. Figure 1-B The location of primers for exon 3 on the HLA-B\*58:01 allele is shown with the blue arrow, along with primer details in the table below.

## 2) Sanger sequencing results for primers P2 and P3

Two tables were compiled with all the Sanger sequencing from NCBI for the two different primers used to target exon 2 and 3, namely P2 and P3. The graphical representation, along with the percentage covered on the original exon and the percentage identical to the exons targeted were all gathered in both tables. The average percentage was calculated for all the samples and this showed that P3 had a higher amplification similarity compared to P2. Hence, the results for P3 was favoured over P2 in the results section, specifically for the HRM section. Results shown in both tables were analysed at different times from 2015 to 2018, but another analysis round was performed in 2019 in order to keep up with the growing number of HLA alleles included in NCBI and to present the most up to date data.

Table 4 - Exon 2 amplification with primer P2 done on 15 samples. This results were analyzed again in August 2019 in order to keep up with new HLA alleles discovered and entered in NCBI.

Sample	Graphical representation of the samples (query) aligned against 149bp targeted part of exon 2 (bp)	Percentage covered on original exon	Percentage identical to exon 2
HLAG1		81%	93.41%
HLAG2		75%	97.78%
HLAG4		83%	96.67%
HLAG5		91%	88.28%
HLAG6		85%	96.67%
HLAG7		81%	94.44%
HLAG26		86%	97.17%
HLAG29		85%	94.23%
HLAG46		86%	95.41%
HLAG56		94%	95.20%
HLAG66		88%	92.79%

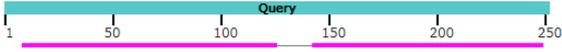
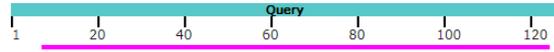
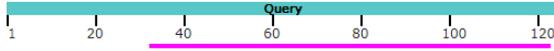
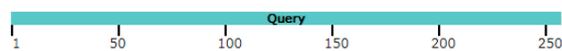
<b>HLAP6</b>		89%	93.52%
<b>HLAP25</b>		92%	96.23%
<b>HLAP46</b>		72%	100%
<b>HLAP60</b>		88%	95.41%
<b>Average percentage</b>		85.1%	95.15%

Table 5 - Exon 3 amplification with primer P3 done on 28 samples. These results were analyzed again in August 2019 in order to keep up with new HLA alleles discovered and entered in NCBI.

Sample	Graphical representation of the samples (query) aligned against 249bp targeted part of exon 3 (bp)	Percentage covered on original exon	Percentage identical to exon 3
HLAG12		98%	99%
HLAG13		96%	95%
HLAG18		98%	95%
HLAG19		97%	95%
HLAG26		97%	100%
HLAG28		98%	98%
HLAG29		97%	99%
HLAG38		91%	99%
HLAG39		94%	95%
HLAG46		91%	97%
HLAG48		92%	99%
HLAG53		92%	99%
HLAG56		93%	100%
HLAG64		90%	99%
HLAG66		91%	100%

<b>HLAG68</b>	<p>Color key for alignment scores</p> <p>■ &lt;40 ■ 40-50 ■ 50-80 ■ 80-200 ■ &gt;=200</p>	92%	99%
<b>HLAG69</b>	<p>Color key for alignment scores</p> <p>■ &lt;40 ■ 40-50 ■ 50-80 ■ 80-200 ■ &gt;=200</p>	91%	95%
<b>HLAP2</b>	<p>Color key for alignment scores</p> <p>■ &lt;40 ■ 40-50 ■ 50-80 ■ 80-200 ■ &gt;=200</p>	91%	95%
<b>HLAP3</b>	<p>Color key for alignment scores</p> <p>■ &lt;40 ■ 40-50 ■ 50-80 ■ 80-200 ■ &gt;=200</p>	91%	98%
<b>HLAP6</b>	<p>Color key for alignment scores</p> <p>■ &lt;40 ■ 40-50 ■ 50-80 ■ 80-200 ■ &gt;=200</p>	92%	99%
<b>HLAP25</b>	<p>Color key for alignment scores</p> <p>■ &lt;40 ■ 40-50 ■ 50-80 ■ 80-200 ■ &gt;=200</p>	93%	99%
<b>HLAP37</b>	<p>Color key for alignment scores</p> <p>■ &lt;40 ■ 40-50 ■ 50-80 ■ 80-200 ■ &gt;=200</p>	92%	98%
<b>HLAP46</b>	<p>Color key for alignment scores</p> <p>■ &lt;40 ■ 40-50 ■ 50-80 ■ 80-200 ■ &gt;=200</p>	92%	98%
<b>HLAP51</b>	<p>Color key for alignment scores</p> <p>■ &lt;40 ■ 40-50 ■ 50-80 ■ 80-200 ■ &gt;=200</p>	92%	99%
<b>HLAP52</b>	<p>Color key for alignment scores</p> <p>■ &lt;40 ■ 40-50 ■ 50-80 ■ 80-200 ■ &gt;=200</p>	91%	96%
<b>HLAP56</b>	<p>Color key for alignment scores</p> <p>■ &lt;40 ■ 40-50 ■ 50-80 ■ 80-200 ■ &gt;=200</p>	91%	98%
<b>HLAP60</b>	<p>Color key for alignment scores</p> <p>■ &lt;40 ■ 40-50 ■ 50-80 ■ 80-200 ■ &gt;=200</p>	91%	100%
<b>HLAP64</b>	<p>Color key for alignment scores</p> <p>■ &lt;40 ■ 40-50 ■ 50-80 ■ 80-200 ■ &gt;=200</p>	90%	95%
<b>Average percentage</b>		93%	98%

### 3) Sanger Alignment of closely related sequences to HLA-B\*58:01

- Alignment of Exon 2 of the HLA-B\*58:01 to other closely related alleles.

AA Codon					5					10					15					20					25
B*58:01:01	GC	TCC	CAC	TCC	ATG	AGG	TAT	TTC	TAC	ACC	GCC	ATG	TCC	CGG	CCC	GGC	CGC	GGG	GAG	CCC	CGC	TTC	ATC	GCA	GTG
B*58:02	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:04	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:05	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:06	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:07	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:08:01	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	T	--
B*58:08:02	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:09	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:10N	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:11	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:12	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	T	--
B*58:13	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:14	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:15	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:16:01	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:16:02	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:17N	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:18	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:19	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:20	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:21	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:22	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:23	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:24	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:25	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:26	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:27	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
B*58:28	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
AA Codon					30					35					40					45					50
B*58:01:01	GGC	TAC	GTG	GAC	GAC	ACC	CAG	TTC	GTG	AGG	TTC	GAC	AGC	GAC	GCC	GCG	AGT	CCG	AGG	ACG	GAG	CCC	CGG	GCG	CCA
B*58:02	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--





```

B*58:26      --- --- --- --- --- --- --- --- --- --- --- --- --- ---
B*58:27      --- --- --- --- --- --- --- --- --- --- --- --- --- ---
B*58:28      --- --- --- --- --- --- --- --- --- --- --- --- --- ---

```

- Alignment of Exon 3 of the HLA-B\*58:01 to other closely related alleles.

AA Codon	95				100				105				110				115								
	GG	TCT	CAC	ATC	ATC	CAG	AGG	ATG	TAT	GGC	TGC	GAC	CTG	GGG	CCC	GAC	GGG	CGC	CTC	CTC	CGC	GGG	CAT	GAC	CAG
B*58:01:01	GG	TCT	CAC	ATC	ATC	CAG	AGG	ATG	TAT	GGC	TGC	GAC	CTG	GGG	CCC	GAC	GGG	CGC	CTC	CTC	CGC	GGG	CAT	GAC	CAG
B*58:02	--	---	---	-C-	C--	---	T--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:04	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:05	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:06	--	---	---	-C-	C--	---	T--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:07	--	---	---	-C-	C--	---	T--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:08:01	--	---	---	-CT	TGG	---	-C-	---	---	---	---	G--	---	--G	---	---	---	---	---	---	---	---	---	A--	---
B*58:08:02	--	---	---	-CT	TGG	---	-C-	---	---	---	---	G--	---	--G	---	---	---	---	---	---	---	---	---	A--	---
B*58:09	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:10N	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:11	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:12	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:13	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:14	--	---	---	---	---	GT-	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:15	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:16:01	--	---	---	-C-	C--	---	---	---	--C	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:16:02	--	---	---	-C-	C--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:17N	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:18	--	---	---	-C-	C--	---	--C	---	--C	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:19	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:20	--	---	---	-C-	C--	---	---	---	--C	---	---	G--	---	--G	---	---	---	---	---	---	---	---	---	A--	---
B*58:21	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:22	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:23	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:24	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:25	--	---	---	-C-	C--	---	T--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:26	--	---	---	---	C--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
B*58:27	--	---	---	-C-	C--	---	--C	---	--C	---	---	G--	---	--G	---	---	---	---	---	---	---	---	---	A--	---







## Appendix D- Multiplex PCR optimisation

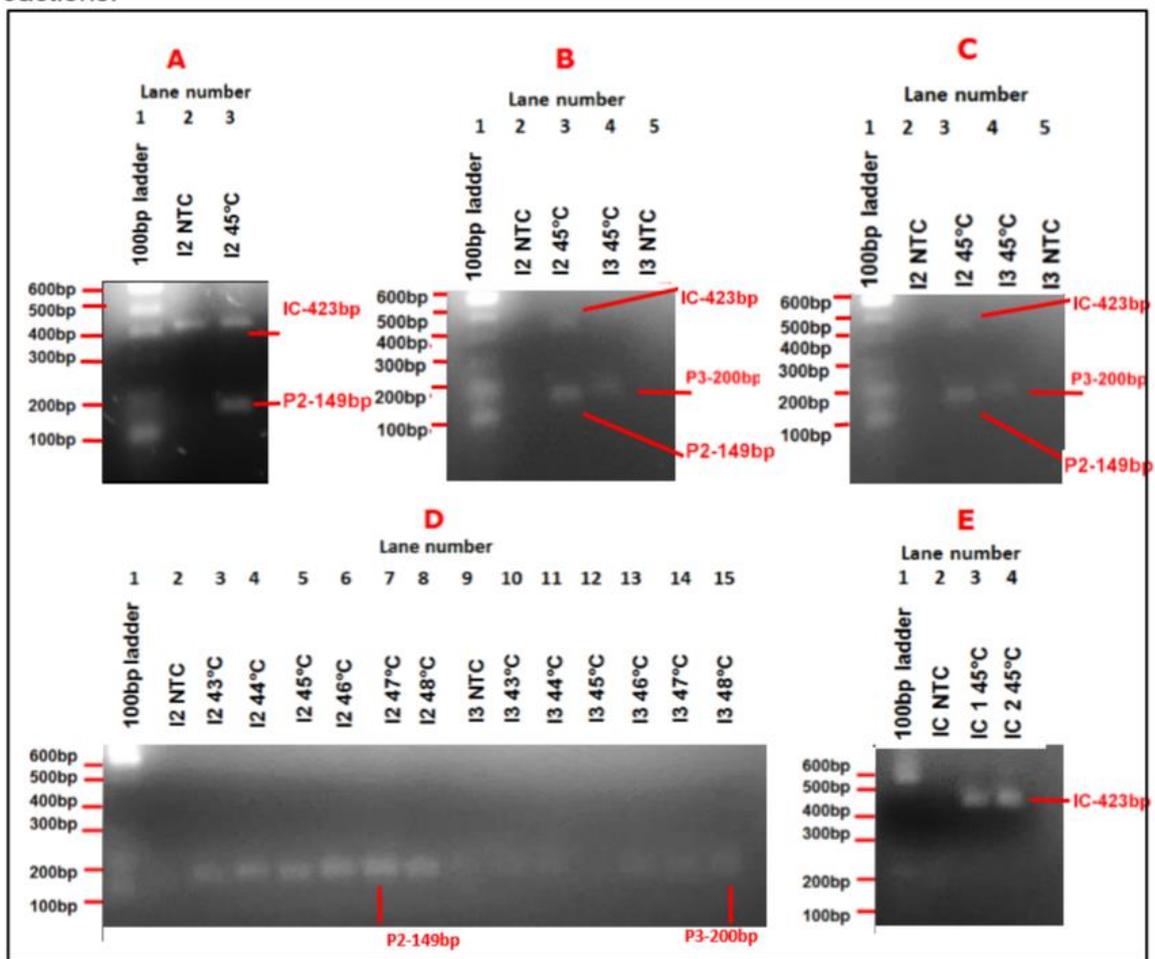


Figure 2- This figure shows the first optimisation for multiplex PCR for I2 and I3. A shows P2 at standard conditions with an annealing temperature of 45°C. A shows contamination of NTC in lane 2. In B, I2 experiments from Figure17-A was repeated (lane 2-3) along with I3 multiplex PCR in lane 4-5. IC bands were invisible for I3 and faint for I2. Thus, C shows an increase in IC concentration(0.3µM) for I2 and I3 which did not ameliorate results. To find a better annealing temperature, a gradient PCR(43°C-48°C) was done for I2 and I3 in D. Due to absence of IC bands, a standard PCR was done in Figure17-E to test for IC presence. Gel E showed IC bands at the expected size (423bp).

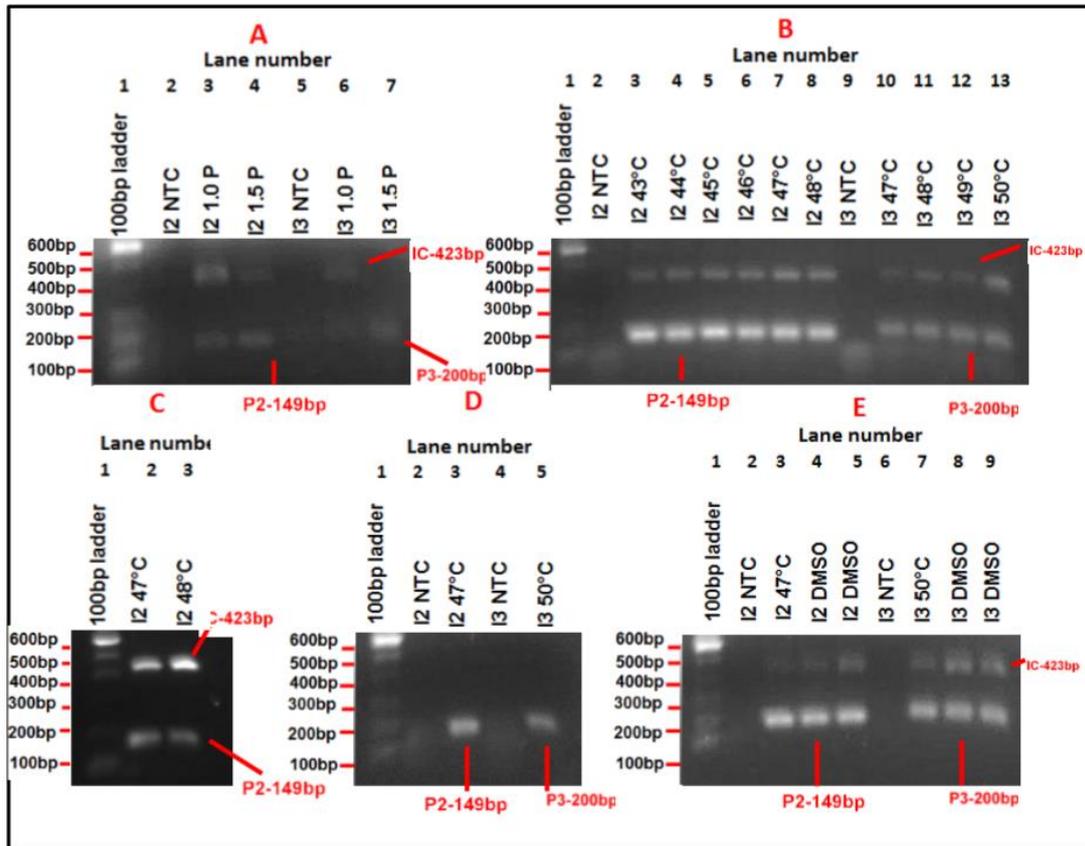


Figure 3-This figure continues optimisation from Figure 2. The 100bp ladder is found in lane 1 of all gels. Figure 18-A shows that with a decreased P2 and P3 concentration (0.1 $\mu$ M) and with increased IC concentration(0.3 $\mu$ M) in lanes 3&6, better results are obtained. B and C improves on this by doing a gradient PCR and finding the best annealing temperatures. I2 has an optimum annealing temperature of 47 $^{\circ}$ C (lane 2 of Figure 18-C) and I3 at 50 $^{\circ}$ C (lane 13, B). A repeat of I2 and I3 with all their optimised conditions yielded improper results, without any IC bands, as shown in Figure D. Using 0.5 $\mu$ L of 100% dimethyl sulfoxide (DMSO) as additive increased stability of reactions as shown in lane 4-5 and 8-9 of E.

## Appendix F- Sanger sequencing results for cloning of exon 2 and exon 3

### 1) Sanger sequencing results for several batches of cloning

This section presents the Sanger sequencing data gathered for randomly selected cloned batches over the years. Both tables show successful cloning, with a higher efficiency for P3 primers in cloning of the exon 3, as shown in Table 7. These set of results again showed that P3 had a higher efficiency in cloning and a larger portion of the exon 3 target was present in the clones.

Table 7- Exon 2 plasmids extracted, with each number representing a new batch extracted and analysed.

Plasmid batch number	P2 targeted region (bp)	Percentage similarity (%)	Cloning Successful
1	149	97.5	Yes
2	149	96.0	Yes
3	149	95.0	Yes
4	149	95.5	Yes
5	149	95.0	Yes
6	149	97.0	Yes
7	149	98.0	Yes
8	149	96.0	Yes
9	149	98.5	Yes
10	149	95.0	Yes
11	149	97.0	Yes
12	149	95.5	Yes

Table 8- Exon 3 plasmids extracted, with each number representing a new batch extracted and analysed.

Plasmid batch number	P3 targeted region (bp)	Percentage similarity (%)	Cloning Successful
1	249	99	Yes
2	249	99	Yes
3	249	100	Yes
4	249	100	Yes
5	249	97	Yes
6	249	97	Yes
7	249	100	Yes
8	249	100	Yes
9	249	99	Yes
10	249	100	Yes
11	249	97	Yes
12	249	100	Yes

**2) EMBL-EBI hits for the inserts cloned**

19 hits show presence of specific HLA-B\*58:01 allele, making up 50% of the 38 hits results table shown from EMBL-EBI.

Align.	DB:ID	Source	Length	Score (Bits)	Identities %	Positives %	E()
<input checked="" type="checkbox"/>	1	IMGTHLAgen:HLA20912	B*58:99 2725 bp	2725	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	2	IMGTHLAgen:HLA20469	B*58:97 2725 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	2725	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	3	IMGTHLAgen:HLA20453	B*58:96 2725 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	2725	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	4	IMGTHLAgen:HLA19642	B*58:95 2725 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	2725	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	5	IMGTHLAgen:HLA15737	B*58:83 2731 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	2731	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	6	IMGTHLAgen:HLA05742	B*58:31N 3004 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	3004	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	7	IMGTHLAgen:HLA26081	B*58:120 2991 bp	2991	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	8	IMGTHLAgen:HLA24402	B*58:117 2725 bp	2725	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	9	IMGTHLAgen:HLA24096	B*58:113 2725 bp	2725	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	10	IMGTHLAgen:HLA23943	B*58:112 2725 bp	2725	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	11	IMGTHLAgen:HLA23434	B*58:111 2921 bp	2921	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	12	IMGTHLAgen:HLA02201	B*58:11 3336 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	3336	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	13	IMGTHLAgen:HLA22642	B*58:108 2725 bp	2725	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	14	IMGTHLAgen:HLA21952	B*58:107 3231 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	3231	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	15	IMGTHLAgen:HLA22114	B*58:106 2725 bp	2725	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	16	IMGTHLAgen:HLA22150	B*58:105 2725 bp	2725	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	17	IMGTHLAgen:HLA22191	B*58:103 2725 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	2725	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	18	IMGTHLAgen:HLA22055	B*58:101 2921 bp	2921	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	19	IMGTHLAgen:HLA21192	B*58:100 3864 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	3864	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	20	IMGTHLAgen:HLA25618	B*58:01:34 2725 bp	2725	513.9	100.0	100.0 1.8E-144
<input checked="" type="checkbox"/>	21	IMGTHLAgen:HLA21519	B*58:01:31 2725 bp	2725	513.9	100.0	100.0 1.8E-144

<input checked="" type="checkbox"/> 22	IMGTHLAgen:HLA20182	B*58:01:28 2725 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	2725	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 23	IMGTHLAgen:HLA19735	B*58:01:27 2725 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	2725	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 24	IMGTHLAgen:HLA19413	B*58:01:25 2973 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	2973	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 25	IMGTHLAgen:HLA18739	B*58:01:24 2725 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	2725	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 26	IMGTHLAgen:HLA13490	B*58:01:19 3336 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	3336	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 27	IMGTHLAgen:HLA12695	B*58:01:17 2689 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	2689	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 28	IMGTHLAgen:HLA05411	B*58:01:07 2898 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	2898	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 29	IMGTHLAgen:HLA05370	B*58:01:06 2689 bp <i>Cross-references and related information in:</i>	2689	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 30	IMGTHLAgen:HLA25699	B*58:01:01:09 2724 bp	2724	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 31	IMGTHLAgen:HLA25922	B*58:01:01:08 2726 bp	2726	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 32	IMGTHLAgen:HLA25396	B*58:01:01:07 2725 bp	2725	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 33	IMGTHLAgen:HLA24591	B*58:01:01:06 2725 bp	2725	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 34	IMGTHLAgen:HLA19009	B*58:01:01:05 3059 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	3059	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 35	IMGTHLAgen:HLA16752	B*58:01:01:04 3336 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	3336	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 36	IMGTHLAgen:HLA16308	B*58:01:01:03 4079 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	4079	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 37	IMGTHLAgen:HLA13394	B*58:01:01:02 3339 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	3339	513.9	100.0	100.0	1.8E-144
<input checked="" type="checkbox"/> 38	IMGTHLAgen:HLA00386	B*58:01:01:01 4094 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	4094	513.9	100.0	100.0	1.8E-144

## Appendix G- High Resolution Melt (HRM) analysis

### 1) Testing primers P2 and P3 on 9 possible positive samples (PPS) with Sanger sequencing

The two tables below show that P3 can catch more SNPs, and therefore more HLA-B alleles are identified correctly. P3 was preferably used for HRM screenings and Sanger sequencing, in order to see the full range of HLA alleles manifesting themselves in each samples. Moreover, P3 primers allowed the identification of the HLA-B\*58 or HLA-B\*58:01 allele in all the 9 samples, where P2 wasn't able to do this. Table 5 shows that with P2 only 3 of the 9 PPS showed presence of the HLA-B\*58/HLA-B\*58:01, with some common alleles seen repeatedly (HLA-B\*35/51/52/15). However, Table 6 with P3 results shows the identification of HLA-B\*58/58:01 for all the 9 PPS. Moreover, all the HLA alleles identified in exon 2 were found in exon 3 results and this was not the same the other way round. Interestingly, HLA-B\*53 was found in all 9 PPS while HLA-B\*35 was found in 5 samples and HLA-B\*15 in 4 samples. Exon 3 amplified with P3 definitely shows more specific and varied results, making it the better choice in order to investigate the HLA alleles present in gout patients.

**Table 9- Exon 2 sequencing with P2 primer for 9 PPS samples.**

Sample code	NCBI BLAST	IMGT/HLA	ADRs manifested
<b>HLAG26</b>	B*35:01:01, 51:01:01	B*53	Exfoliative dermatitis
<b>HLAG29</b>	B*35:01:01, 51:01:01	B*52, B*58:01, 58:44, 58:68, 58:01:25, B*57	DRESS Syndrome
<b>HLAG46</b>	B*35, B*51	B*53, B*51, B*35	Allopurinol hypersensitivity syndrome- red eye + severe itchiness
<b>HLAG56</b>	All B*58- 58:01 to 58:74	B*52	DRESS- rash
<b>HLAG66</b>	B*52 numerous, 52:01, 35	B*52	Itchiness
<b>HLAP6</b>	B*15 numerous, 15:01:01	B*15	SJS ( mouth and genitalia)
<b>HLAP25</b>	B*15 numerous, 15:01:01	B*56 numerous	Generalized skin rash
<b>HLAP46</b>	35:01, 35- numerous, 51, 53	B*58 numerous	rash on face
<b>HLAP60</b>	B*52	B*52	Generalized rashes

**Table 10 – Exon 3 sequencing with P3 primers for 9 PPS.**

<b>Sample</b>	<b>Possible base change + associated allele</b>	<b>NCBI BLAST</b>	<b>IMGT/HLA BLAST</b>	<b>ADRs manifested</b>
<b>HLAG26</b>	CC (58:02)	B*53, B*35, B*58:01, B*58:02	All B*58	Exfoliative dermatitis
<b>HLAG29</b>	CC (58:02)	B*53, B*35, B*58:01, B*58:74, B*58:02	All B*58	DRESS Syndrome
<b>HLAG46</b>	TACC (SS) OR TAA (B*35)	B*53, B*35, B*58:28, B*35:02,	Numerous B*35 , 58:65	Allopurinol hypersensitivity syndrome- red eye + severe itchiness
<b>HLAG56</b>	TAA (B*35)	B*53, B*35, B*58:62, B*58:74	All B*58	DRESS- rash
<b>HLAG66</b>	TAA (B*35) OR CC (58:02)	B*53, B*35, B*58:01, 58:02, B*58:74	All B*58	Itchiness
<b>HLAP6</b>	CCCA (SS)	B*15, B*15:02, 58:06, 58:01, 58:02	B*15, B*44, B*53, 58:01, 58:91, 58:19	SJS ( mouth and genitalia)
<b>HLAP25</b>	TAA (B*35) OR CC (58:02)	B*35, B*53, B*15, , 58:74, 58:01, 58:02	All B*58	Generalized skin rash
<b>HLAP46</b>	CCC (SS) OR B*15(GGGG) OR TAA (B*35)	B*35, B*53, B*15, , 58:01, 58:02, 57	All B*15, 58:97	rash on face
<b>HLAP60</b>	TAA (B*35)	B*35, B*53, B*15, , 58:01, 58:02, 58:74	All B*58	Generalized rashes

## 2) Chromatogram analysis for Sanger sequencing results

All samples had to undergo a round of chromatogram analysis, where the integrity of the sequenced data was checked by hand and base by base. The key below shows how the chromatogram was modified and added to the Sanger sequence for sample HLAG13. The final edited forward and reverse sequence were aligned and mismatched bases were called as polymorphism sites or SNPs. All the different combinations of the final sequenced were inputted into NCBI BLAST and more refined HLA hits came out.

Key for chromatogram analysis:

Editing example: GGGCTG 0 C T/G C/G

0- removed bases, base wrongly called

GTC- Normal bases correctly called by Sanger sequencing

T/G - 2 peaks with the first base having a higher peak. Other peak is significant too (possible SNP)

C/G -Heterozygous peak, same peak height but 2 different bases seen (try to alternate to get different sequences)

A - Added base, which was miscalled by Sanger sequencing

Take out the first 20 bases from sequence as a lot of them were labelled as N.

### Chromatogram analysis for HLAG13

Edited forward sequence-

GCGCCTCCTCCGCGGGT/CATAACCAGTTCGCCTACGACGGCAAGGATTACATCGCCCTGAAC/TGAGGACTCT  
GCGCTCCTGGACCGCCGCGGACACGGCGGCTCAGATCT/ACCCAGCGCAAGTG/TGGAGGCGGCCCGT/ACT  
AGGCGGAGCAGCTGAGAGCCTACCTGGAGGGCA/GCAGTGCGTGGAGTGGCTCCGCAGATACCTGGAGAAC  
GGGAAGGAA

1 <sup>st</sup> sequence- using first base option	GCGCCTCCTCCGCGGGTATAACCAGTTCGCCTACGACGGCAAGGATTACATCGCCCTGAA CGAGGACTCTGCGCTCCTGGACCGCCGCGGACACGGCGGCTCAGATCTCCCAGCGCAAGT GGGAGGCGGCCCGTGCTAGGCGGAGCAGCTGAGAGCCTACCTGGAGGGCACAGTGCGTG GAGTGGCTCCGCAGATACCTGGAGAACGGGAAGGAA
2 <sup>nd</sup> sequence using 2 <sup>nd</sup> base option	GCGCCTCCTCCGCGGGCATAACCAGTTCGCCTACGA CGGCAAGGATTACATCGCCCTGAA TGAGGACTCTGCGCTCCTGGACCGCCGCGGACACGGCGGCTCAGATCACCCAGCGCAAGT TGGAGGCGGCCCGTACTAGGCGGAGCAGCTGAGAGCCTACCTGGAGGGCGCAGTGCGTG GAGTGGCTCCGCAGATACCTGGAGAACGGGAAGGAA

Edited reverse-





**Sample HLAG12-** Aligned to highest Sanger sequencing hit HLA-B\*35:01

Query= sanger sequencing result, Subject= specific HLA allele hit

Here, no base change was seen (no gaps), hence no SNP was found.

**Alignment of both sequences:**

```
Query 1002 ggTCTCACATCATCCAGAGGATGTATGGCTGCGACCTGGGGCCCGACGGGCGCCTCCTCC 1061
          |
Sbjct 1 GGTCTCACATCATCCAGAGGATGTATGGCTGCGACCTGGGGCCCGACGGGCGCCTCCTCC 60

Query 1062 GCGGGCATGACCAGTCCGCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGA 1121
          |
Sbjct 61 GCGGGCATGACCAGTCCGCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGA 120

Query 1122 GTCCTGGACCGCGGCGGACACCGCGGCTCAGATCACCCAGCGCAAGTGGGAGGCGGCC 1181
          |
Sbjct 121 GTCCTGGACCGCGGCGGACACCGCGGCTCAGATCACCCAGCGCAAGTGGGAGGCGGCC 180

Query 1182 GTGTGGCGGAGCAGCTGAGAGCCTACCTGGAGGGCCTGTGCGTGGAGTGGCTCCGCAGAT 1241
          |
Sbjct 181 GTGTGGCGGAGCAGCTGAGAGCCTACCTGGAGGGCCTGTGCGTGGAGTGGCTCCGCAGAT 240

Query 1242 ACCTGGAGAACGGGAAGGA 1260
          |
Sbjct 241 ACCTGGAGAACGGGAAGGA 259
```

**Sample HLAG13-** Aligned to highest Sanger sequencing hit HLA-B\*40's

Query= sanger sequencing result, Subject= specific HLA allele hit

Empty gaps between query and subject= SNP presence. (highlighted in yellow)



```

Query 1062 GCGGGCATGACCAGTCCGCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGA 1121
          |||
Sbjct 61 GCGGGCATGACCAGTCCGCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGA 120

Query 1122 GCTCCTGGACCGCGGGACACGGCGGCTCAGATCACCCAGCGCAAGTGGGAGGCGGCC 1181
          |||
Sbjct 121 GCTCCTGGACCGCGGGACACGGCGGCTCAGATCACCCAGCGCAAGTGGGAGGCGGCC 180

Query 1182 GTGAGGCGGAGCAGTGGAGAGCCTACCTGGAGGGCCTGTGCGTGGAGTGGCTCCGCAGAT 1241
          |||
Sbjct 181 GTGTGGCGGAGCAGCTGAGAGCCTACCTGGAGGGCCTGTGCGTGGAGTGGCTCCGCAGAT 240

Query 1242 ACCTGGAGAACGGGAAGGA 1260
          |||
Sbjct 241 ACCTGGAGAACGGGAAGGA 259

```

**Total base change: TTCC to GGGG**

**Aligned to HLA-B\*15:02**

```

Query 1001 ggTCTCACATCATCCAGAGGATGTATGGCTGCGACGTGGGGCCGGACGGGCGCCTCTCC 1060
          |||
Sbjct 1 GGTCTCACATCATCCAGAGGATGTATGGCTGCGACCTGGGGCCCGACGGGCGCCTCTCC 60

Query 1061 GCGGGTATGACCAGTCCGCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGA 1120
          |||
Sbjct 61 GCGGGCATGACCAGTCCGCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGA 120

Query 1121 GCTCCTGGACCGCGGGACACGGCGGCTCAGATCACCCAGCGCAAGTGGGAGGCGGCC 1180
          |||
Sbjct 121 GCTCCTGGACCGCGGGACACGGCGGCTCAGATCACCCAGCGCAAGTGGGAGGCGGCC 180

Query 1181 GTGAGGCGGAGCAGCTGAGAGCCTACCTGGAGGGCCTGTGCGTGGAGTGGCTCCGCAGAT 1240
          |||
Sbjct 181 GTGTGGCGGAGCAGCTGAGAGCCTACCTGGAGGGCCTGTGCGTGGAGTGGCTCCGCAGAT 240

Query 1241 ACCTGGAGAACGGGAAGGA 1259
          |||
Sbjct 241 ACCTGGAGAACGGGAAGGA 259

```

**Total base change: CCCC to GGGA**

#### 4) Calculating specific base changes for all 28 ADR/SCARs samples

Exon 3 sequence for the HLA-B\*58:01 gene : 276bp (total)

Primer P3 amplified exon 3 (yellow region)- 259bp

```
GGTCTCACATCATCCAGAGGATGTATGGCTGCGACCTGGGGCCCCGACGGGCGCCTCCTCC
GCGGGCATGACCAGTCCGCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGA
GCTCCTGGACCGCGGGACACCGCGGCTCAGATCACCCAGCGCAAGTGGGAGGCGGCC
GTGTGGCGGAGCAGCTGAGAGCCTACCTGGAGGGCCTGTGCGTGGAGTGGCTCCGCAGAT
ACCTGGAGAACGGGAAGGA GACGCTGCAGCGCGCGG
```

#### Details for exon 3:

Total count, all bases:	259
Adenine (A) count:	50
Thymine (T) count:	36
Guanine (G) count:	93
Cytosine (C) count:	80
%G-C content:	66.8

Positive samples- 9 samples out of 30 ADR samples- HLAG26, HLAG29, HLAG46, HLAG56, HLAG66, HLAP6, HLAP25, HLAP46, HLAP60

Sample HLAG26	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	219	223	442
Adenine	43	32	75
Thymine	27	41	68
Guanine	81	73	154
Cytosine	68	77	145
%GC content	68%	67.3%	67.6%

Sample HLAG29	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	229	215	444
Adenine	45	29	74
Thymine	28	41	69
Guanine	85	72	157
Cytosine	71	73	144
%GC content	68.1%	67.4%	67.8%

Sample HLAG46	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	226	218	444
Adenine	48	34	82
Thymine	28	41	69
Guanine	84	68	152
Cytosine	66	75	141
%GC content	66.4	65.6	66

Sample HLAG56	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	224	220	444
Adenine	45	33	78
Thymine	30	40	70
Guanine	81	72	153
Cytosine	68	75	143
%GC content	66.5	66.8	66.7

Sample HLAG66	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	220	216	436
Adenine	44	32	76
Thymine	28	40	68
Guanine	80	71	151
Cytosine	68	73	141
%GC content	67.3	66.7	67

Sample HLAP6	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	221	221	442
Adenine	46	32	78
Thymine	28	43	71
Guanine	80	69	149
Cytosine	67	77	144
%GC content	66.5	66.1	66.3

Sample HLAP25	Forward sequence	Reverse sequence	Forward + Reverse
---------------	------------------	------------------	-------------------

Total bases count	223	220	443
Adenine	44	30	74
Thymine	28	42	70
Guanine	83	72	155
Cytosine	68	76	144
%GC content	67.7	67.3	67.5

Sample HLAP46	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	234	220	454
Adenine	46	32	78
Thymine	30	41	71
Guanine	87	69	156
Cytosine	71	78	149
%GC content	67.5	66.8	67.2

Sample HLAP60	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	223	220	443
Adenine	44	34	78
Thymine	28	40	68
Guanine	83	72	155
Cytosine	68	74	142
%GC content	67.7	66.4	67

Trend noticed:

- Number of G and C is less for all the positive samples, compared with the original exon 3
- % GC content may be inaccurate as the total base count isn't the same for all samples. The range for % GC spreads over a range of 66 to 67.8
- A and T numbers are lower for all positive samples, compared to exon 3

Negative samples- HLAG12, 13, 18, 19, 28,

Sample HLAG12	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	219	225	444
Adenine	42	34	76
Thymine	27	41	68
Guanine	82	73	155
Cytosine	68	77	145
%GC content	68.5%	66.7 %	67.6%

Sample HLAG13	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	220	219	439
Adenine	43	33	76
Thymine	30	41	71
Guanine	82	69	151
Cytosine	65	76	141
%GC content	66.8 %	66.2%	66.5%

Sample HLAG18	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	229	213	442
Adenine	46	31	77
Thymine	28	41	69
Guanine	85	68	153
Cytosine	70	73	143
%GC content	67.7%	66.2	67

Sample HLAG19	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	221	219	440
Adenine	44	31	75
Thymine	27	41	68
Guanine	80	74	154
Cytosine	70	73	143
%GC content	67.9	67.1	67.5

Sample HLAG28	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	221	217	438
Adenine	43	31	74
Thymine	27	41	68
Guanine	82	68	150
Cytosine	69	77	146
%GC content	68.3	66.8	67.6

Negative samples- HLAG12, 13, 18, 19, 28, 38, 39, 48, 53, 64, 68, 69, HLAP2, HLAP3, 37

Sample HLAG38	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	225	221	446
Adenine	46	33	79
Thymine	28	41	69
Guanine	84	69	153
Cytosine	67	78	145
%GC content	67.1	66.5	66.8

Sample HLAG39	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	232	229	461
Adenine	47	32	79
Thymine	27	43	70
Guanine	88	76	164
Cytosine	70	78	148
%GC content	68.1	67.2	67.7

Sample HLAG48	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	222	215	437
Adenine	43	31	74
Thymine	29	40	69
Guanine	83	70	153
Cytosine	67	74	141
%GC content	67.6	67	67.3

Sample HLAG53	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	222	217	439
Adenine	44	31	75
Thymine	29	40	69
Guanine	82	71	153
Cytosine	67	75	142
%GC content	67.1	67.3	67.2

Sample HLAG64	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	225	219	444
Adenine	45	30	75
Thymine	27	45	72
Guanine	86	67	153
Cytosine	67	77	144
%GC content	68	65.8	66.9

Sample HLAG68	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	222	218	440
Adenine	46	31	77
Thymine	29	41	70
Guanine	81	68	149
Cytosine	66	78	144
%GC content	66.2	67	66.6

Sample HLAG69	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	226	225	451
Adenine	47	32	79
Thymine	30	43	73
Guanine	83	71	154

Cytosine	66	79	145
%GC content	65.9	66.7	66.3

Sample HLAP2	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	223	215	438
Adenine	46	33	79
Thymine	28	40	68
Guanine	83	68	151
Cytosine	66	74	140
%GC content	66.8	66	66.4

Sample HLAP3	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	224	221	445
Adenine	45	31	76
Thymine	27	41	68
Guanine	84	70	154
Cytosine	68	79	147
%GC content	67.8	67.4	67.6

Sample HLAP37	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	236	221	457
Adenine	48	31	79
Thymine	33	45	78
Guanine	87	69	156
Cytosine	68	76	144
%GC content	65.7	65.6	65.6

Negative samples- HLAG12, 13, 18, 19, 28, 38, 39, 48, 53, 64, 68, 69, HLAP2, HLAP3, 37, 51, 52, 56, 64,

Sample HLAP51	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	226	233	459
Adenine	45	32	77
Thymine	29	47	76
Guanine	86	76	162
Cytosine	66	78	144
%GC content	67.3	66.1	66.7

Sample HLAP52	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	223	224	447
Adenine	45	32	77
Thymine	27	42	69
Guanine	85	74	159
Cytosine	66	76	142
%GC content	67.7	67	67.3

Sample HLAP56	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	224	222	446
Adenine	46	35	81
Thymine	27	41	68
Guanine	84	69	153
Cytosine	67	77	144
%GC content	67.4	65.8	66.6

Sample HLAP64	Forward sequence	Reverse sequence	Forward + Reverse
Total bases count	220	223	443
Adenine	44	36	80
Thymine	29	41	70
Guanine	81	71	152
Cytosine	66	75	141
%GC content	66.8	65.5	66.1

Trend noticed:

- Number of G and C is less for all the negative samples, compared with the original exon 3
- % GC content may be inaccurate as the total base count isn't the same for all samples. The range for % GC spreads over a range of 65.5% to 68%. A wider range of GC % is seen for the negative samples.
- A and T numbers are lower for all negative samples, compared to exon 3

##### 5) Melting point shift calculation from HRM results

**Table 11- Melting point generated in HRM runs for the 9 possible positive samples and their corresponding positive control used during screening.**

Sample	Tm	Tm1	+VE C 1	+VE C 2	Temperature change (degrees)
HLAG26	88.1	88.2	90.8	90.9	2.7
HLAG29	90.9	91	90.8	90.9	0.1
HLAG46	90.4	90.5	90.8	90.9	0.4
HLAG56	90.8	90.9	90.7	90.8	0.1

HLAG66	91.1	91.2	90.7	90.8	0.4
HLAP6	90.7	90.8	90	89.8	0.85
HLAP25	91.2	91.2	90.5	90.3	0.8
HLAP46	91	91	90.6	90.7	0.35
HLAP60	91.2	91.2	90.6	90.7	0.55

**Table 12- Melting point generated in HRM runs for the 19 negative samples and their corresponding positive control used during screening.**

Sample	Tm	Tm1	+VE C 1	+VE C 2	Temperature change (degrees)
HLAG12	91	91.1	90.8	90.7	0.3
HLAG13	90.6	90.7	90.8	90.7	0.1
HLAG18	90.7	90.6	90.8	90.9	0.2
HLAG19	90.9	90.8	90.8	90.9	0*
HLAG28	90.9	90.8	90.8	90.9	0*
HLAG38	90.7	90.7	90.9	90.8	0.15
HLAG39	90.6	90.6	90.9	90.8	0.25
HLAG48	91	91.1	90.9	90.8	0.2
HLAG53	90.9	90.9	90.7	90.9	-0.05
HLAG64	90.8	90.7	90.8	90.7	0*
HLAG68	90.7	90.6	90.7	90.6	0*
HLAG69	90.7	90.7	90.7	90.6	-0.05
HLAP2	90.5	90.5	90	89.8	-0.6
HLAP3	90.8	90.7	90	89.8	-0.95
HLAP37	90.5	90.4	90.9	90.7	0.35
HLAP51	90.6	90.6	90.8	91	0.3
HLAP52	90.8	90.7	90.8	91	0.15
HLAP56	90.9	91	90.8	91	-0.05
HLAP64	90.7	90.8	90.7	90.6	-0.1

## 6) Analysis of %GC content

The percentage of Guanine and Cytosine were gathered for all the 28 ADR samples from their Sanger sequencing results. The full DNA sequence generated for each sample, being amplified with primer P3, was inputted in an online percentage GC calculator. The numbers of the 4 different bases were calculated, along with the percentage GC automatically. The Table 13 below shows the compiled results for the 28 ADR samples. No specific trend was noticed for %GC for both possible positive and negative samples. HRM should by theory be able to detect unknown variants and SNPs, even in cases with a single base change. Base changes will affect the % GC and melting point of all samples, and there should be a trend seen for all similar positive or negative samples. The results in the table below showed an erratic difference in the melting point for all samples, hence preventing any trend to be identified. This also shows that there are other unknown variable factors affecting the results.

%GC will in turn affect the melting point and change it. A known and consistent %GC and Tm was needed for positive and negative samples to allow the proper differentiation in terms of genotype.

**Table 13- Analysis of the GC percentage in 28 ADR samples, with the possible positive samples and negative samples for the HLA-B\*58:01 allele.**

Sample code	Total base count in sequenced DNA fragment (bp)	Percentage GC content (%)	HLA-B*58:01 positive or negative
<b>Original Exon 3</b>	<b>259</b>	<b>66.8</b>	<b>Positive</b>
HLAG26	221	67.6	Positive
HLAG29	222	67.8	Positive
HLAG46	222	66.0	Positive
HLAG56	222	66.7	Positive
HLAG66	218	67.0	Positive
HLAP6	221	66.3	Positive
HLAP25	222	67.5	Positive
HLAP46	227	67.2	Positive
HLAP60	222	67.0	Positive
HLAG12	222	67.6	Negative
HLAG13	220	66.5	Negative
HLAG18	221	67.0	Negative
HLAG19	220	67.5	Negative
HLAG28	219	67.6	Negative
HLAG38	223	66.8	Negative
HLAG39	230	67.7	Negative
HLAG48	219	67.3	Negative
HLAG53	220	67.2	Negative
HLAG64	222	66.9	Negative
HLAG68	220	66.6	Negative
HLAG69	226	66.3	Negative
HLAP2	219	66.4	Negative
HLAP3	223	67.6	Negative
HLAP37	229	65.6	Negative
HLAP51	230	66.7	Negative
HLAP52	224	67.3	Negative
HLAP56	223	66.6	Negative
HLAP64	222	66.1	Negative

Trend noticed for negative samples:

- Number of G and C is less for all the negative samples, compared with the original exon 3
- % GC content may be inaccurate as the total base count isn't the same for all samples. The range for % GC spreads over a range of 65.5% to 68%. A wider range of GC % is seen for the negative samples.
- A and T numbers are lower for all negative samples, compared to exon 3

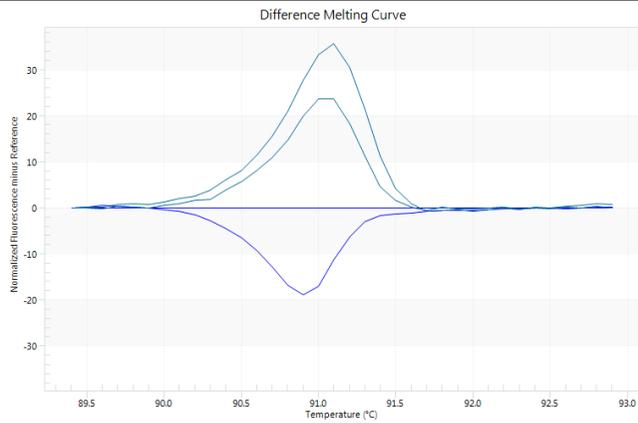
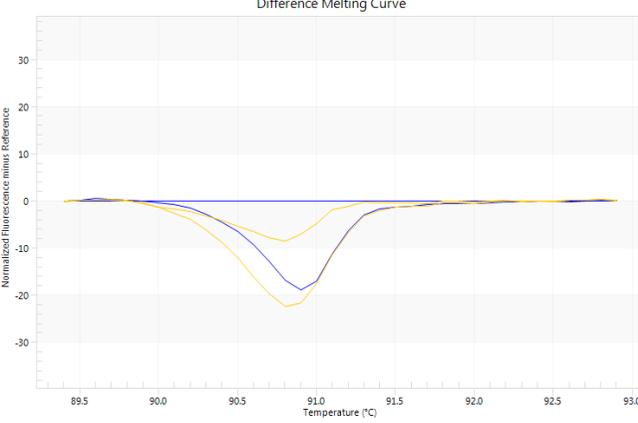
Trend noticed for positive sample

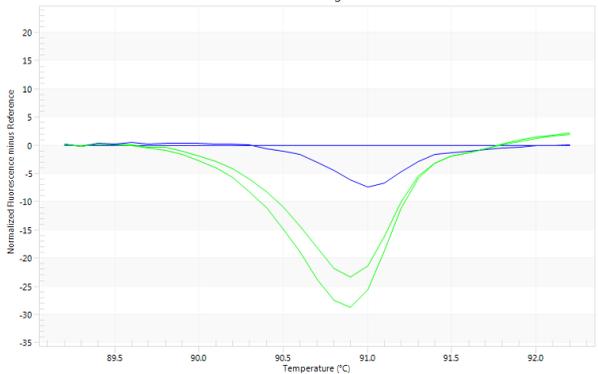
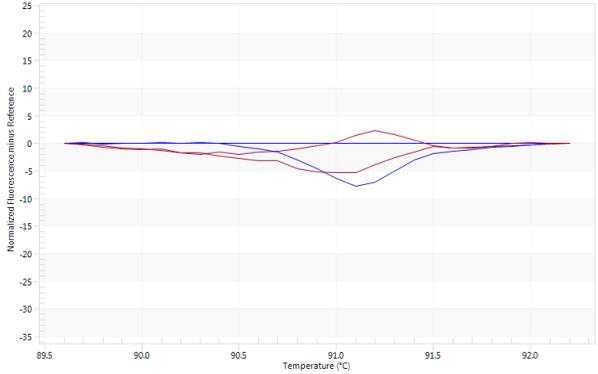
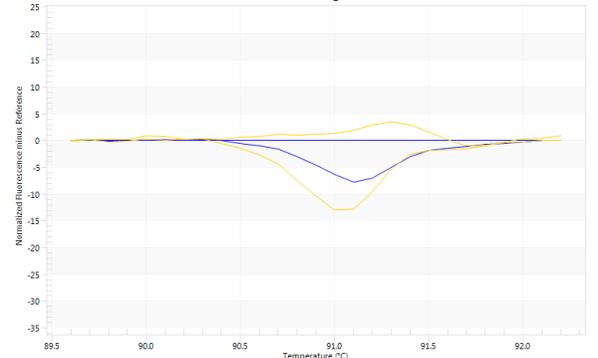
Number of G and C is less for all the positive samples and negative samples, compared with the original exon 3, which may be due to incomplete sanger due to complex secondary structures in PCR products

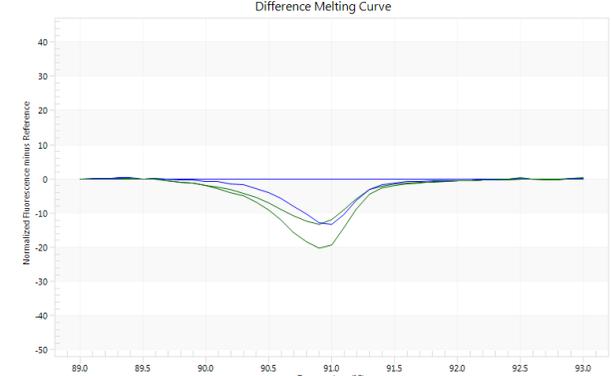
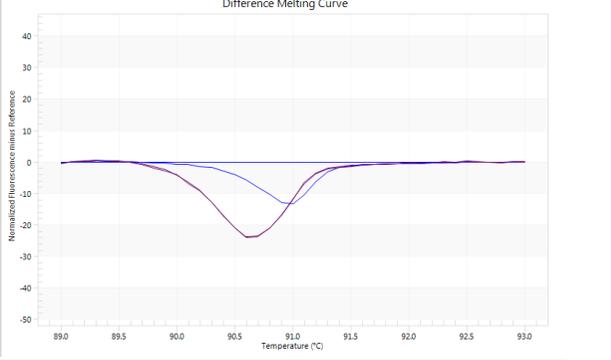
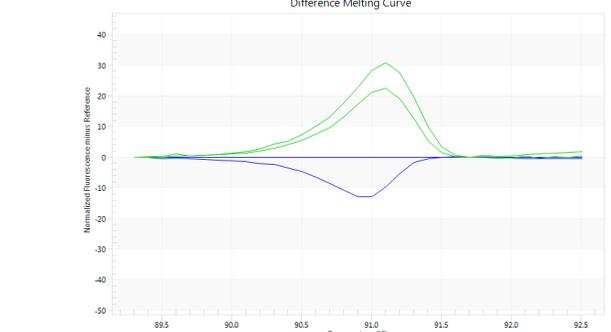
% GC content may be inaccurate as the total base count isn't the same for all samples. The range for % GC spreads over a range of 66 to 67.8

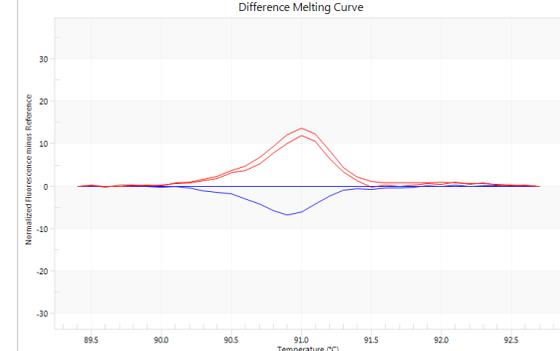
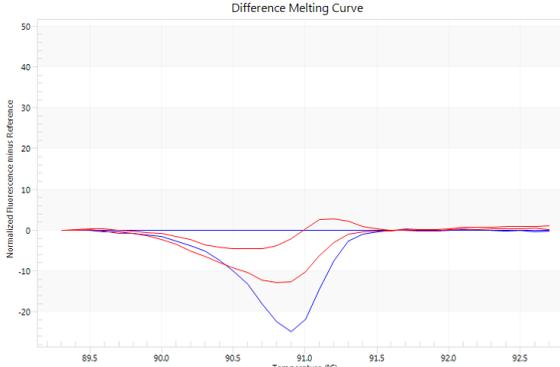
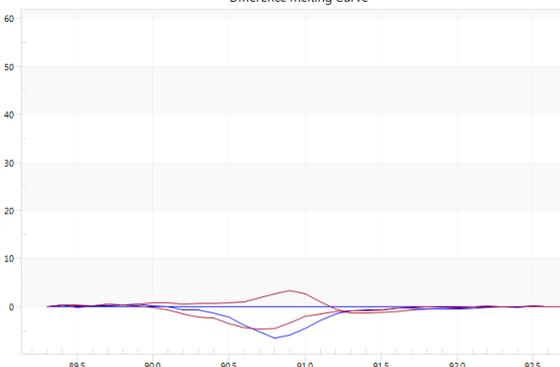
7) Full HRM trend compiled in one Table for 19 negative samples.

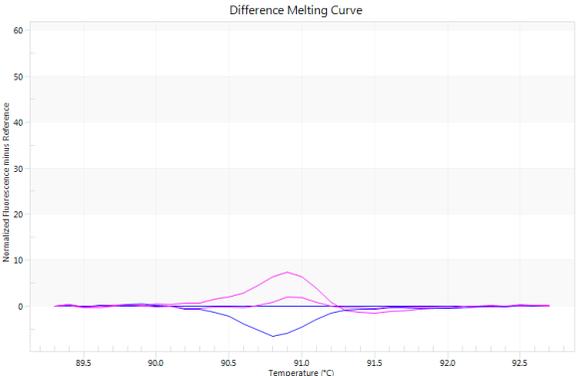
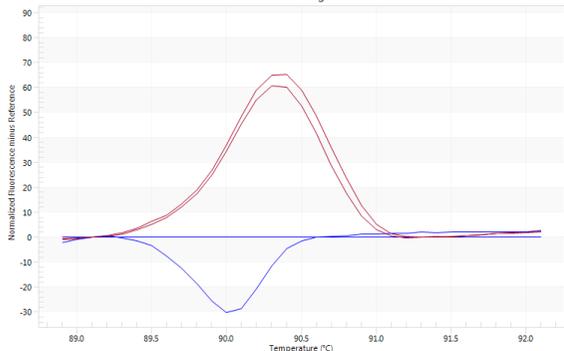
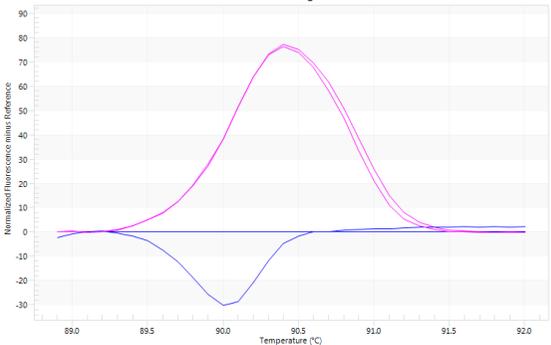
Table - Compilation of all the different variables used to identify a trend in the HRM results for the 19 negative ADR samples. Sanger sequencing results from Table 4.11 were also added into 2 columns based on their databases. HRM difference melt curves and ADRs observed were shown in the last two columns.

Sample	Temperature change (degrees)	Possible base change + associated allele	NCBI BLAST	IMGT/HLA BLAST	HRM difference melt curve	ADRs manifested
HLAG12	0.3	-	B*53:01, 58:74, 35:01, 35:10, 35:20, 35:28	All B*58 alleles	 <p>Difference Melting Curve</p> <p>Normalized Fluorescence minus Reference</p> <p>Temperature (°C)</p>	Vasculitits rash
HLAG13	0.1	TTTAAA (SS) or G (B*40) AG (B*48)	B*40, B*48	B*40, C*04, B*48,58:27	 <p>Difference Melting Curve</p> <p>Normalized Fluorescence minus Reference</p> <p>Temperature (°C)</p>	Mild allergy-itchiness

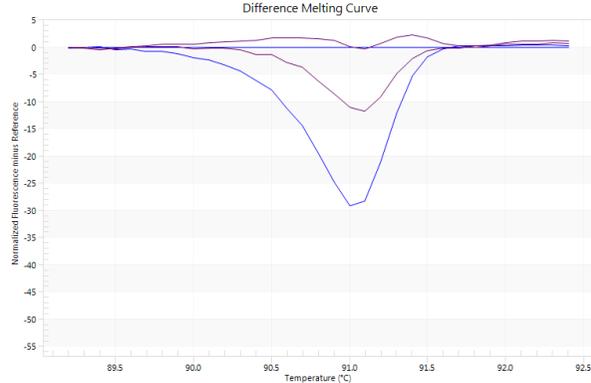
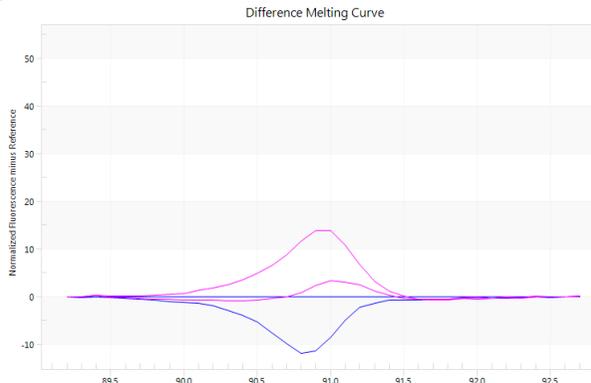
HLAG18	0.2	AAAA (SS) or G (C*04)	C*04:01:01:01	C*04	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus Reference Temperature (°C)</p>	Mild allergy- rash
HLAG19	0*	AA (SS) or CCCC (C*07)	C*07	C*07	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus Reference Temperature (°C)</p>	Mild allergy- itchiness
HLAG28	0*	GGAA (B*44:02) or CTC (SS)	B*44	B*44	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus Reference Temperature (°C)</p>	Itchiness

HLAG38	0.15	ACC (SS), GGGG (B*15:01), GGGA (B*15:02)	B*15:01/15:02	B*15	 <p>Difference Melting Curve</p> <p>Normalized Fluorescence minus Reference</p> <p>Temperature (°C)</p>	SJS
HLAG39	0.25	TT (SS), GC (B*12:02:02)	C*12	C*12	 <p>Difference Melting Curve</p> <p>Normalized Fluorescence minus Reference</p> <p>Temperature (°C)</p>	SJS
HLAG48	0.2	T (B*35:03:01) or TA (SS)	B*35, B*35:03	B*58:34, B*35	 <p>Difference Melting Curve</p> <p>Normalized Fluorescence minus Reference</p> <p>Temperature (°C)</p>	Tophi reuptured, allergy to allo, changed to feb. no mention of specific ADRs

HLAG53	-0.05	TA (SS) or T (B*35:03:01)or AA(B*35:02:02)	B*35, B*35:03:01, B*35:02:02	B*35, B*58:34	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus reference vs Temperature (°C)</p>	Rashes
HLAG64	0*	TCA (SS), GGGG (B*15:01), GGGA (B*15:02)	B*15	B*15, B*15:02	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus reference vs Temperature (°C)</p>	Rashes
HLAG68	0*	TCA (SS), GGGG (B*15:01), GGGA (B*15:02)	B*15	B*15, B*15:02	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus reference vs Temperature (°C)</p>	Allergy to allo- Anaphylaxis reaction

HLAG69	-0.05	CTTTT(SS) G (B*40) AG (B*48)	B*40, B*48	B*40, B*48	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus reference</p> <p>Temperature (°C)</p>	Allergy to allo
HLAP2	-0.6	AAA(SS), G (C*04)	C*04	C*04	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus reference</p> <p>Temperature (°C)</p>	Generalized rash
HLAP3	-0.95	GGA(44:03:01), TCC (SS), GGAA (B*44:02)	B*44:03:01	B*44	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus reference</p> <p>Temperature (°C)</p>	Allergy to allopurinol-generalised rash + face swelling. Febuxostat allergy- rash

HLAP37	0.35	TTA (B*56), TTTA (SS)	B*56	B*56, B*58:73		Generalized rash
HLAP51	0.3	TCA (SS), GGGA (B*15:02)	B*15, B*15:02	B*15, B*57:41, B*57:22		Allopurinol- exfoliative dermatitis + febuxostat allergy
HLAP52	0.15	AA (SS) GC (B*12:02:02)	C*12	C*12		Allergy to allopurinol- transaminitis- no rash seen. Another type of allergy?

HLAP56	-0.05	GGA(44:03:01), TCCA (SS)	B*44:03:01, B*44	B*44	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus Reference</p> <p>Temperature (°C)</p>	Generalized rash+ skin eruptions (raised and flat eruptions), macular popular rash
HLAP64	-0.1	G (B*40), TT (SS), AG (B*48)	B*48, B*40	B*48, B*40	 <p>Difference Melting Curve</p> <p>Normalized fluorescence minus Reference</p> <p>Temperature (°C)</p>	Severe allergy to allopurinol-angioedema + eye swelling

## 8) Calculation of cost of HRM run for 1 sample in our laboratory

1 HRM run= 15 unknown samples + 1 positive control

Reagents used- Sensifast ready mix

Platform- eco plate and seals needed

Price for 5ml of sensifast readymix= RM954

Price for 50 eco plates and seals= RM 779.1

1 sample needs 30ul of sensifast ready mix= RM 5.724

1 HRM run (16 samples) sensifast readymix needed= RM91.58

1 eco plate and seal =RM 15.58

Therefore, reagents and eco plate and seals for 1 HRM run (16 samples)=RM107.17

Thus 1 sample HRM run, all inclusive= RM 6.70

## Appendix H- Next Generation Sequencing (NGS)

### 1) NGS PCR with HLA-B allele

#### Analysis of HLA-B PCR products by gel electrophoresis:

Gel electrophoresis is carried out on a 0.7% agarose gel with SYBR Safe dye and run at 90V for 1hr. The gel image below shows an example of a gel run at UMBI with 5ul of the PCR product loaded into the well. The correct band at around 4000bp is shown, but is however faint, even after a full set of optimization has been done twice. This is attributed to gene's characteristics, being highly polymorphic and of a length of 4kb plus. A slight change in any of the optimized conditions results in no results at all.

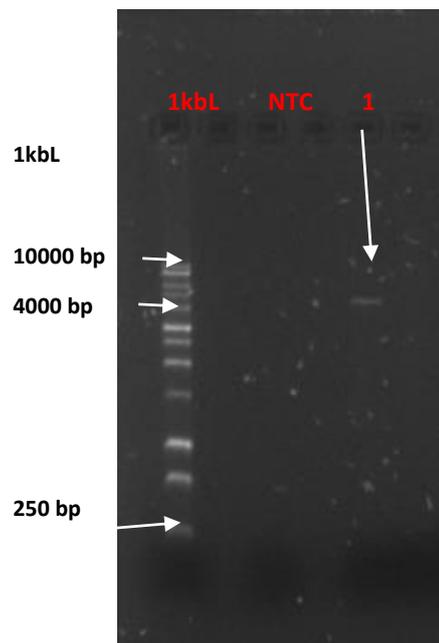


Figure 5- Gel electrophoresis of HLA-B primers for NGS, showing the target around 4000bp when compared to the 1kb Ladder.

## 2) Sanger sequencing for the HLA-B primers product:

Forward sequence:

```
>1st_BASE_2710925_HLAB1_PF
TCAAAAAAATCAGAAAACCCCGTGTGCAGGGGCCCTGGGCGATGTGTGAGCCTCTGTGGTCACAGCT
CCCCTGGACAAGTTTCCGCTGAAGGGACAAGGACAATGGGGCAGTGAAGGTGACCCAACCTGAGGACT
AACCACATAAAGCCCATGATGGACTCAACAACAACCTGGGCACAGGCCCCCGTCCACACTTGGCCCCCA
CAGCCTTCTCCACACCCCACCTGCAACAGACTCAGCACAGCGAACATGCAGATTCTGGAAGGTTCTCA
GGTCTTTATTTGCTCTCTCACATTCCAGGAATTGACTTATTTAATTAATCCATCAACCTCTCATAGCA
AATATTTGAGAAAACAAATTTATATTTCAGATTCTTATTTTCAGTAGGGGAAGTAAGAAGTTGCAGCTCA
GTGCACATAAAGTTGAGACAGAGATGGAGACATCCAGCCCCACCTCTCTGGAACAAGAAAGATGACTG
GGGAGGAAACACAGGTCAGCATGGGAACAGGGGTCACAGTGGACACAAGGGTGGGCTGTCTCTCCACC
TCCTCACATTATGCTAACAGGAACGCAGACACATTCAGGTGCCTTTGCAGAAAGAGATGCCAGAGGCT
CTTGAAGTCACAAAGGGGAGGAGTGAAGAAAATCCTGCATCTCAGTCCCTCACAAAGACAGCTGTCTCAG
GCTACAGAAAACAACAGTCATGAACAAATTCTGGTTAGTCATGGTAAGTGATGACACTCTGAACAGCC
CATCACACACGCGAAACATCCCAATCAAAGAATCTCCATTACCCAGGCCTTTCCCTCTGCCCCCTCC
CCGCCCCCGGCCAATTCAAAACCCCCAAAACCCCAACTTTTTTCAGCCTGTAGAAAAACACTTCAAAC
CCCTGGCAACGGTCTTGGCTGGATTAAAACAAAACCCGGAACCTGGGTCAAACCCCCCAGGAAAACG
TGAACAGGAGGAATTAAGGGGGGGGGGAAACTCCCTCCATACTCCTAACCTCCCAACATTTAACT
GCTAGCCTGGAAAAAAACTTCCCTCCCTTTTCCCCCGGGGGGAAGAAAAATGTCTGGGGAGGGGGA
CTGGTAAGGAAATTAGGTCCATGGAAAACCTAAAAGGAACCTCCCCTGGTCTTGGGACCCCAAAGGG
AATTTTCCCAAAAAGATTGGACTTCCGAACCCCTAGGGGAGTGGATCTGGGAAAACCTTAACGGAAAAG
CCATTGGTTGTGGGTCTGGGAACAAAACAGGCACCCCCCTAAAGGGCTGGGCCCTTTAATAAAGGA
GACCATTACCCCTCGGATTCCCGTGTAAACGCTCGTATAAATTATGGTGTACCCCTCTTTACTTCCCTA
TAACTCCCTCTAAAGGGTGGTATATCACCCCTATTTCCCTTTACTGTTTATCCACTGTATGGGTCTT
GTCTGCAATTTAATAATAACTAATTGGGCT
```

Reverse sequence:

```
>1st_BASE_2713835_HLAB1_PR
NNNNNNNNNTTNNNNTTGNATCTATCATTCAGGGATTACCAATATTGTGCTACCTACTGTATTAAC
AAACAAAAGGAACTGGTCTCTATGAGAATCTCTGTGTGGTGCCTTCAGACAAAACCTCGCCAGGTT
TANAGAGAAACCCCTGTCTCTACACCTCCATTCNCANGGCGAGCTCNCCTCTNNCATCAAGTTNTC
GTGNTCNGTTTTCCCTACCANGATCCANNAGGAGAGGTANNGAGNGNAGGCAGGGAGTCCAGTTCAGGN
ACAGGNATTCGGGAGGAAAGTGANGGGAANGGGTGGNGCACNTGNNGTNCCTCCCTGGTTCACAGACNA
TCNNGTTCAGGACTAGGCAANAGTNGACAANGGCGGTGGTGAAGAAGAATCGAAAATCCNGGCCNNG
GAGGCTCTCTCGGCTCAGGCTTNAAGCCTTGTCTGGTTGGGGAGGCGCACCGTTGGGGATTTCCCTC
TCCCAGATTCACTTCTTCTCTCAACCTATGTGCGGTCCCTTCTTCCCGATACTCTCGACACGTCCCCA
TTTTACTCCCATTGGGTGTCNGATATCTANAGAAANNATCAGTGTGCGCGGGGTTCACTTCTCTAGT
CCCCACGCACCCCCCGACTCTCAACTCCTCTCACACCNGATGCGCTGGACATGGCCCCGAANCGTCTT
CCTGCTGCTCTGGGGGGGACTGGGCTGAANAGANNTGTGCCGGTGAGTGCAGCGGTGGGGAGGGAATGG
CCTCTGTGGGGAGGAGCGAGGGGACCGCAGGCGGGGGCGCAGGACCTGAGGAGCCGCGCCGGGAGGAG
GGTCNGG
```

The sequence was inputted in NCBI BLAST and the majority generated the presence of the HLA-B gene. However, as shown by the sanger sequencing raw data, the full 4000bp wasn't sequenced properly due to high polymorphism and presence of secondary and tertiary structures which makes the sequence difficult to decipher by Sanger. This is a validated primer, used for NGS runs in several studies, hence showing its accuracy and specificity.

### 3) NCBI BLAST Results table:

<input checked="" type="checkbox"/> 8	IMGTHLAcds:HLA00163	B*15:01:01:02N 1208 bp <i>Cross-references and related information in:</i> <a href="#">▶ Nucleotide sequences</a> <a href="#">▶ Literature</a>	1208	157.1	95.1	95.1	2.4E-37
<input checked="" type="checkbox"/> 9	IMGTHLAcds:HLA01451	C*04:09N 1197 bp <i>Cross-references and related information in:</i> <a href="#">▶ Nucleotide sequences</a> <a href="#">▶ Literature</a>	1197	153.1	94.3	94.3	3.7E-36
<input checked="" type="checkbox"/> 21	IMGTHLAcds:HLA26055	B*81:09 1089 bp	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 22	IMGTHLAcds:HLA05306	B*81:05 1089 bp <i>Cross-references and related information in:</i> <a href="#">▶ Nucleotide sequences</a> <a href="#">▶ Literature</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 23	IMGTHLAcds:HLA24269	B*81:01:05 1089 bp <i>Cross-references and related information in:</i> <a href="#">▶ Nucleotide sequences</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 24	IMGTHLAcds:HLA23607	B*81:01:04 1089 bp <i>Cross-references and related information in:</i> <a href="#">▶ Nucleotide sequences</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 25	IMGTHLAcds:HLA21285	B*81:01:03 1089 bp <i>Cross-references and related information in:</i> <a href="#">▶ Nucleotide sequences</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 26	IMGTHLAcds:HLA18730	B*81:01:02 1089 bp <i>Cross-references and related information in:</i>	1089	54.0	96.7	96.7	2.5E-6

<input checked="" type="checkbox"/> 26	IMGTHLAcds:HLA18730	B*81:01:02 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 27	IMGTHLAcds:HLA25244	B*81:01:01:02 1089 bp	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 28	IMGTHLAcds:HLA00398	B*81:01:01:01 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 29	IMGTHLAcds:HLA20138	B*78:10 1089 bp	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 30	IMGTHLAcds:HLA13984	B*78:09 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 31	IMGTHLAcds:HLA00990	B*78:04 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 32	IMGTHLAcds:HLA00395	B*78:02:02 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 33	IMGTHLAcds:HLA00394	B*78:02:01 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 34	IMGTHLAcds:HLA27077	B*78:01:01:03 1089 bp	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 35	IMGTHLAcds:HLA16309	B*78:01:01:02 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/> 36	IMGTHLAcds:HLA20912	B*58:09 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	1089	54.0	96.7	96.7	2.5E-6

<input checked="" type="checkbox"/>	37	IMGTHLAcds:HLA20863	B*58:98 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/>	38	IMGTHLAcds:HLA20469	B*58:97 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/>	39	IMGTHLAcds:HLA20453	B*58:96 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/>	40	IMGTHLAcds:HLA19642	B*58:95 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/>	41	IMGTHLAcds:HLA19509	B*58:94N 1089 bp	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/>	42	IMGTHLAcds:HLA18350	B*58:93N 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/>	43	IMGTHLAcds:HLA16426	B*58:86 1012 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	1012	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/>	44	IMGTHLAcds:HLA16243	B*58:84 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/>	45	IMGTHLAcds:HLA15737	B*58:83 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/>	46	IMGTHLAcds:HLA14211	B*58:76 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	1089	54.0	96.7	96.7	2.5E-6

<input checked="" type="checkbox"/>	45	IMGTHLAcds:HLA15737	B*58:83 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/>	46	IMGTHLAcds:HLA14211	B*58:76 1089 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	1089	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/>	47	IMGTHLAcds:HLA12592	B*58:87 1012 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a>	1012	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/>	48	IMGTHLAcds:HLA10773	B*58:58 1012 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	1012	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/>	49	IMGTHLAcds:HLA10772	B*58:57 1012 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	1012	54.0	96.7	96.7	2.5E-6
<input checked="" type="checkbox"/>	50	IMGTHLAcds:HLA10789	B*58:56 1012 bp <i>Cross-references and related information in:</i> ▶ <a href="#">Nucleotide sequences</a> ▶ <a href="#">Literature</a>	1012	54.0	96.7	96.7	2.5E-6

#### 4) Pre-processing data for all 6 NGS samples

SAMP LE CODE	PRE-PROCESSING					Percentage mapped to the Human Genome (%)
	With adaptors		Without adaptors		Percentage left after trimming (%)	
	Total sequence length (bp)	Average FAST Q Score	Total sequence length (bp)	FAST Q Score		
PC580 1	8334	38-40	7411	38-40	88.92488601	94.45
HLAG 12	10196	38-40	8913	38-40	87.41663397	92.93
HLAG 19	9305	38-40	7848	38-40	84.34175175	87.55
HLAG 26	10982	38-40	9251	38-40	84.23784374	91.98
HLAG 28	10435	38-40	7780	38-40	74.55678007	96.36
HLAG 29	10105	38-40	8294	38-40	82.07817912	83.43