The University of Nottingham

# Identification of novel transcripts of CHRNA7 and CHRFAM7A in airway epithelial cells

**Omar Akram Jerjees Ahmad**

**MB,CHB, MSc**

**Thesis submitted to the University of Nottingham for the**

**Degree of Doctor of Philosophy**

# Acknowledgment

I would like to thank Allah for all the gifts he gave me and I might be not aware of them all.

# Dedication

I will never be able to acknowledge enough my family's support and faith in me during my studies. Their help opened all the locked doors, lifting the spirits every day, and making the impossible possible. I will never forget the help of my father and mother and they will always stay in my heart. Special thanks to my children who supported me, as they could, and helped me keep smiling. No words would be enough thanks to my beloved wife, who kept on supporting me even when I lost faith in myself. I am so lucky to have her by my side.

## Declaration

I declare that this thesis is the result of my own work and has not in this, or any other form, been presented to this, or any other university in support of an application for any degree other than for which I am now a candidate.


Omar Ahmad

# **Abbreviations**:

ACh                      Acetylcholine

A1AT                     Alpha-1-antitrypsin

α-BTX                    Alpha-bungarotoxin

α7-nAChRs               Alpha-7 nicotinic acetylcholine receptors

C-A                      C-A allele of exon 9b

CHRNA7                  Alpha-7 nicotinic acetylcholine receptor coding gene

CHRFAM7A               cholinergic receptor alpha7-with FAM7A fusion gene

COPD                    Chronic Obstructive Pulmonary Diseases

CSE                     Cigarette smoke extract

C-terminus              Carboxy-terminal end of the protein

Dup α7                  Duplicated alpha7 protein

ESEs                    Exon splice enhancers

ESSs                    Exon splice suppressors

FEV1%                   Forced expiratory volume during the first second (precent)

FEV1/FVC                FEV1/forced vital capacity ratio

G-G                     G-G allele of exon 9b

GOLD                    Global initiative for chronic obstructive lung disease

GWAS                    Genome-wide association study

HMGB1                   high-mobility group box 1 protein

IC loop                 Intracellular loop

IL                      Interleukins

ISEs                    Intron splice enhancers

ISSs                    Intron splice suppressors

| | |
|---|---|
| LD | Linkage disequilibrium |
| LPS | lipopolysaccharide |
| nAChRs | nicotinic acetylcholine receptors |
| NMD | Nonsense mediated decay |
| N-terminus | amino-terminal end of the protein |
| 3`SS | 3` splice site |
| 5`SS | 5` splice site |
| SNP | Single nucleotide polymorphism |
| snRNP | small nuclear ribonuclear protein |
| TNF | Tumour necrosis factor |

# Abstract

**Background:** RNA splicing is a crucial process for delivering the appropriate message for protein synthesis. Most genes are affected by alternative splicing, and among these is CHRNA7. This gene encodes for the nicotinic acetylcholine α7 receptor subunit that is involved in the cholinergic anti-inflammatory pathway. This anti-inflammatory pathway is considered an important part of the human body's defence line against tissue injury or infection and causative mechanism in COPD.

**Aim:** The aim of the present study was to investigate the role of alternative splicing on the nature of the transcripts generated by CHRNA7 gene and its partial duplicate, CHRFAM7A.

**Methods:** Airway epithelial cell lines, A549 and BEAS2B, were mainly used as targets for testing alternative splicing. RT-PCR, TA cloning and gel extraction methods were used for testing CHRNA7 and CHRFAM7A transcripts. Following RT-PCR, the resulting product band intensities were analysed using densitometric analysis tools. This was followed by the use of several bioinformatics analysis tools to predict the protein structure for the resulting transcripts. For one of the detected transcripts, minigene methods were used to test for the source of expression.

**Results:** A novel transcript missing exon 9 is reported for the first time. Both genes showed the expression of full length and the novel transcripts (missing exon 9) at similar ratios (~2:1). These results could be detected in immortalised cell lines from human alveolar and bronchial epithelial cells (A549 and BEAS2B, respectively) and in BE (2)-c cells (neuroblastoma cells with bone

marrow metastasis). The same results were shown when primary human peripheral blood monocytes cells (PBMC) were tested. This means that the effect of missing exon 9 is not tissue-specific, and is not only found in cancerous cells, indicating that it could be a common feature of splicing for these two genes. Furthermore, another novel transcript was detected which is inserted exon 9b. The initial RT-PCR experiments seemed to suggest that this was derived from CHRFAM7A only. The use of minigene methods showed that this transcript could be expressed from both genes, CHRNA7 and CHRFAM7A, but a single nucleotide base within the inserted sequence (at position 77 from the 5` end) could play a role in enhancing of exon 9b in the mRNA transcripts. This base is C allele in CHRFAM7A sequence of exon 9b, while its corresponding base in CHRNA7 is G allele that has less prominent effect on exon 9b inclusion.

**Conclusion:** CHRNA7 and CHRFAM7A express novel transcripts in different human cells that are missing exon 9. This could be due to inactive splicing factors that are required for recognition of exon 9 as a constitutive exon. For exon 9b transcripts, these lie within the common sequence of CHRNA7 and CHRFAM7A, and it seems that the presence of C allele at position 77 could enhance the inclusion of exon 9b in CHRFAM7A more than the presence of G allele in CHRNA7 sequences. The results shown in this study implicate a possible regulatory role of the transcripts detected on the control mechanism exerted by CHRFAM7A on CHRNA7. These results help to suggest a possible role of in the development of COPD in the form of inflammatory/anti-inflammatory control imbalance.

# Thesis related publication:

## A- Presentations and conference communications:

1. EMU (East Midlands Universities Postgraduate Student Research) conference at University of Nottingham/Sutton Bonington campus, 5th July, 2012.

2. Winter meeting of Pathologic Society of Great Britain and Ireland, 7th-9th January, 2013.

http://onlinelibrary.wiley.com/doi/10.1002/path.4190/pdf (page 13)

## B- Awards and prizes:

Runner up of best poster presentation at the School of Molecular Medical Sciences postgraduate day, UoN, UK, 2010.

# Contents:

# List of Tables:

# LIST OF FIGURES:

# 1   The role of CHRNA7 and CHRFAM7A in mediating the anti-inflammatory response

## 1.1   Chronic Obstructive Pulmonary Disease (COPD):

COPD is a chronic inflammatory disease interrupted by episodes of acute exacerbation. A major feature of the disease is the partial loss of reversibility of airflow obstruction. Although smoking is the major risk factor contributing to the disease, there is strong evidence for the role of genetic factors, which are not understood completely. Chronic inflammation is a hallmark of COPD, a process that may be controlled via the cholinergic anti-inflammatory pathway. The α7 nicotinic acetylcholine receptor subunit (nAChR) has a central role in this anti-inflammatory pathway. This chapter will cover a short introduction to COPD and the potential role of CHRNA7 in this disease.

### 1.1.1   Definition

COPD is a preventable and treatable disease with pulmonary and extra pulmonary components. The pulmonary component is characterized by airflow limitation that is not fully reversible, progressive, and associated with an abnormal inflammatory response of the lungs to noxious particles or gases (GOLD, 2013).

#### 1.1.1.1 Chronic bronchitis and emphysema

The term COPD is becoming more frequently used for describing patients having chronic bronchitis and emphysema. Chronic bronchitis is defined (symptomatically) as chronic productive cough for three successive months per year for more than one year, while emphysema is defined (pathologically) as abnormal permanent dilatation of airways distal to the terminal bronchioles (Kumar, c2012).

**1.1.2   Classification**

A simple spirometric classification of COPD based on lung function is recommended by GOLD (Table 1-1).

**Table 1-1: Spirometric Classification of COPD Severity Based on Post-Bronchodilator FEV1 % predicted.**

| COPD stage | FEV1/FVC | FEV1% predicted |
|---|---|---|
| **1: Mild** | < 0.70 | $\geq$ 80 % |
| **2: Moderate** | < 0.70 | 50% $\leq$ and < 80% |
| **3: Severe** | < 0.70 | 30% $\leq$ and < 50% |
| **4: Very severe** | < 0.70 | < 30 % |
| FEV1: forced expiratory volume during the first second, FVC: forced vital capacity, Adopted from (GOLD, 2013). | | |

For simple classification, spirometry is used. This method includes the use of short-acting bronchodilators to test for reversibility and to minimize variation in the results between different individuals tested. Following the use of a bronchodilator, the individual is asked to apply forceful expiration using the spirometer tool. This will help assessing the forced expiratory volume during the first second as a percentage of the predicted value ($FEV_1$% predicted).

**1.1.3   The burden of COPD:**

COPD is a known cause of worldwide morbidity and mortality and it is expected to be the third leading cause of worldwide mortality and the fifth leading cause of morbidity by the year 2020 (Buist et al., 2007). This worldwide recognized burden on health and treatment costs is increasing due to the increased prevalence of the aging population which is contributing to the increasing prevalence of the disease in older age groups. The estimated

population prevalence in adults is about 10% (Lopez et al., 2006, GOLD, 2013), but this may be an underestimate as it will take more than a decade for a COPD to be established and it is difficult to diagnose COPD in the early stages (Lokke et al., 2006). Recent prospective cohort studies revealed that after 25 years of smoking, more than 20% of smokers will develop COPD, two-thirds of which will develop severe forms of this disease (stages 3 and 4 according to GOLD classification) (Mannino et al., 2003, Lokke et al., 2006).

### 1.1.4   Pathogenesis

COPD is a disease of complex etiology, and there are several factors that play a role in its pathogenesis. Early pathological findings in chronic bronchitis include mucus glands hypertrophy with resultant increased mucus secretion. If the condition continued to progress, the bronchi would be inflamed, and inflammatory cells would infiltrate the bronchial walls. This may end up with fibrosis and narrowing of the airways. On the other hand, the changes in emphysema includes dilatation of the small bronchioles leading to air trapping and limitation of airflow, which may progress and cause loss of the elastic recoil of the lung tissue (Barnes et al., 2003) (Figure 1-1).
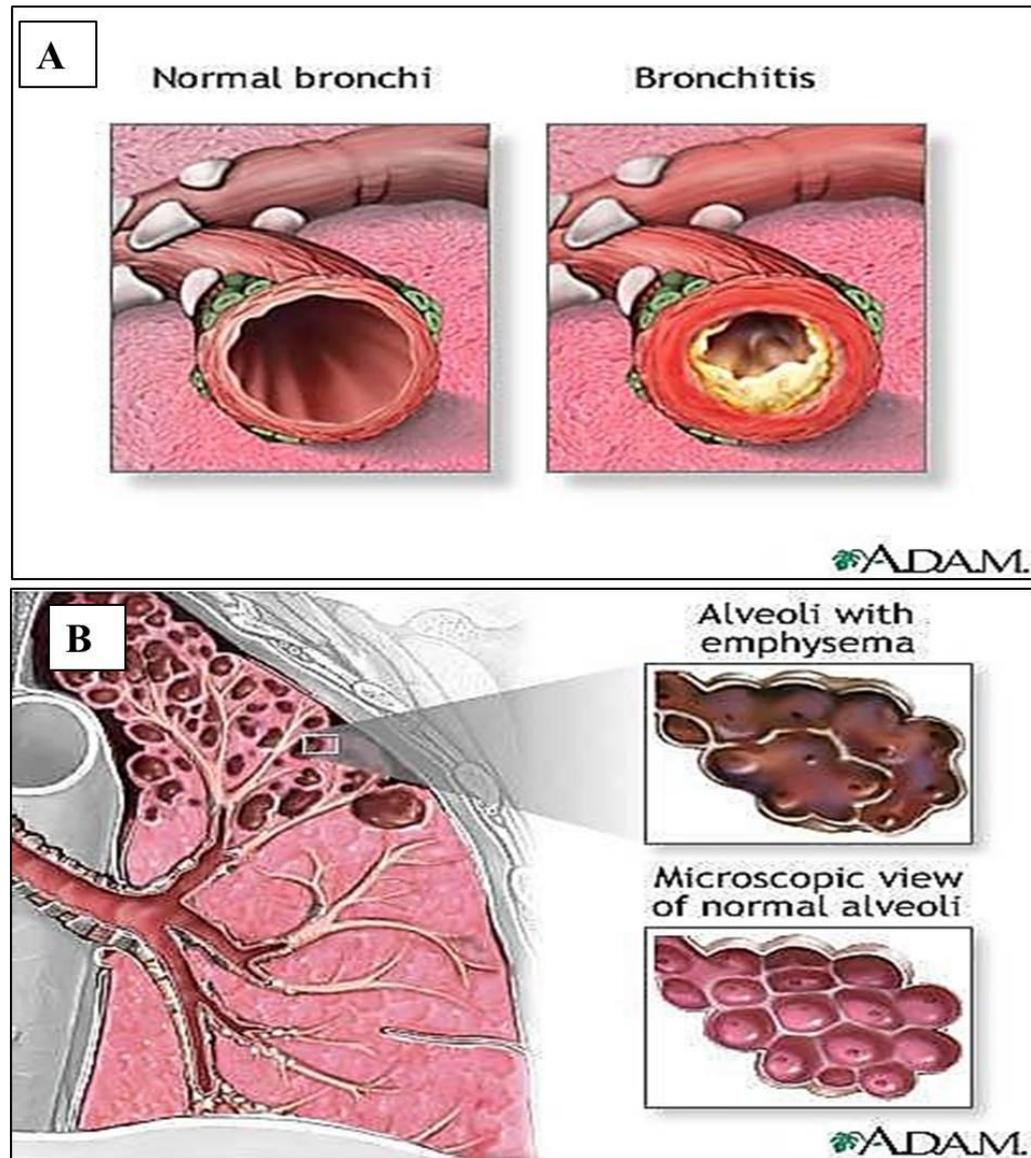
**Figure 1-1: Pathological features of COPD.**

Chronic bronchitis and emphysema, the two major component parts of COPD, are compared with healthy lung tissue. *A:* Chronic bronchitis showing the inflamed bronchi precipitated with the inflammatory products and excess mucus thus decreasing the lumen size of the bronchi (as compared to the healthy bronchi). B: Emphysema showing alveoli with loss of the alveolar membranes and thus decreasing the gas-exchange surface area (as compared to the healthy alveoli).

Taken from (nurseslabs, 2012).

**1.1.5   Development of COPD:**

There are a number of factors that can contribute individually or collectively to the development of COPD (Barnes et al., 2003) (Table 1-2).

**Table 1-2: Risk factors for COPD.**

| |
|---|
| Smoking |
| Protease/anti-protease imbalance |
| Genetic factors |
| Oxidative stress |
| Other factors (occupational exposure, outdoor pollution, age, infections, inflammation and other factors). |

**1.1.5.1 Smoking**

Although there are several risk factors for COPD, cigarette smoking remains to be the predominant one (GOLD, 2013). There are many factors that determine the decline in lung function and subsequent development of COPD amongst smokers, like age at starting and pack-years of smoking. Although cessation of smoking is an effective strategy to slowdown this decline in experimental animals (Li et al., 2007), it requires a long period to resolve the established inflammatory changes in the human lungs (Barnes et al, 2003; Rutgers et al, 2000)(Lokke et al., 2006). Despite the significant role of smoking in developing COPD, only about 20 % of smokers develop COPD and the disease also occurs in nonsmokers (Mannino et al., 2003, Lokke et al., 2006, Celli et al., 2005). However, when using the new criteria based on the GOLD classification of COPD, it seems that up to half of smokers develop COPD (Lundback et al., 2003).

## 1.1.5.1.1  How does smoking cause COPD?

In response to prolonged inhalation of cigarette smoke and other noxious particles, inflammation can start in the lung tissue. In some smokers, this response is exaggerated, leading to chronic inflammation and damage to the lung parenchyma (with subsequent emphysema).  In addition, small airway fibrosis may result from disturbed immune mechanisms. Both effects cause air trapping and airflow limitation leading to the development of COPD (Barnes et al., 2003). These effects are mediated by activation of inflammatory cells (tissue macrophages and neutrophils in the lung). Similarly, smoking leads to increased oxidants released from these cells, and may cause decreased levels of antioxidants in COPD patients, resulting in what is called 'Oxidative stress'. Smoking can affect the protease/antiprotease balance, an effect that is mediated by inflammatory cells, leading to dominant protease effects and subsequent damage to the lung tissue (Figure 1-2).
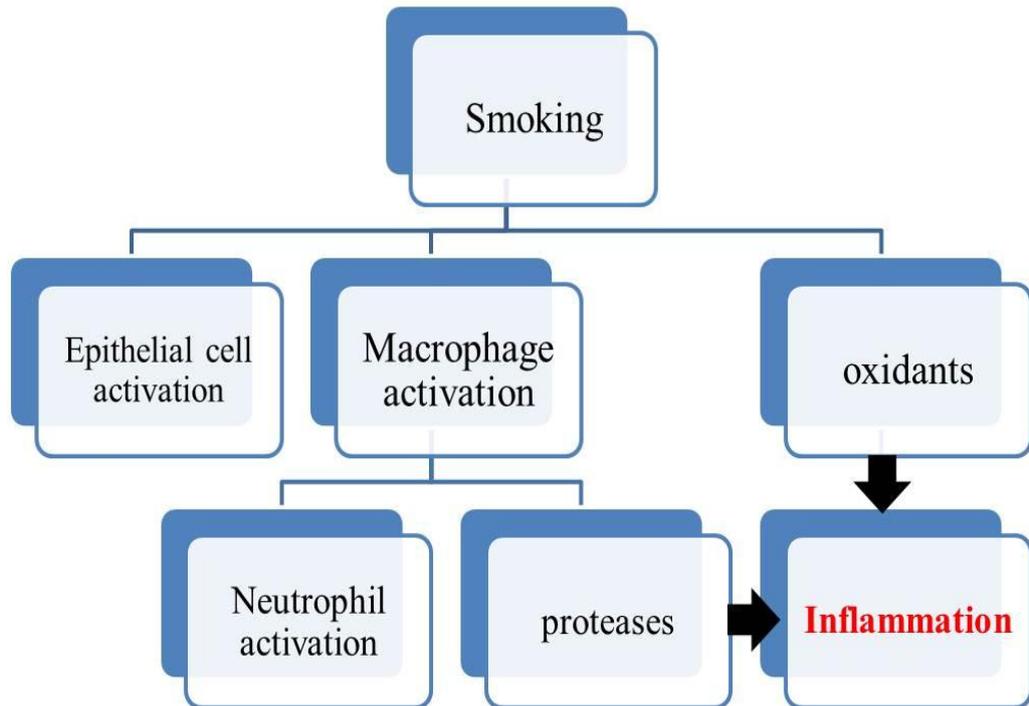
**Figure 1-2: The role of smoking in COPD pathogenesis.**
Smoking can lead to the activation of inflammatory cells like macrophages and epithelial cells. The activation of tissue macrophages can cause the start of inflammatory-like conditions in the lungs, which serve to recruit neutrophils to the inflamed area. This can further be complicated by the release of protease enzymes from the inflammatory cells (macrophages, neutrophils, and epithelial cells). On the other hand, another consequence of smoking is the imbalance of oxidants and anti-oxidants in the lungs leading to oxidative stress. Collectively, all these factors can cause lung inflammation.

**1.1.5.2 Genetic factors**

*Can genetic factors predispose to COPD?*

COPD is an example of interaction between genetic and environmental factors (Kalsheker and Chappell, 2008). Although smoking has been shown to be the most important environmental factor, family based studies suggest a role for genetic factors in the development of COPD (Silverman et al., 1998, GOLD, 2013). The best recognized genetic factor that contributes to the development of COPD is alpha-1-antitrypsin (A1AT) deficiency. The fact that COPD tends to cluster in families was the first evidence of association of genetic factors with the development of COPD that interplay with environmental factors.

**1.1.5.2.1   Familial aggregation of COPD:**

Family studies provided evidence for genetic factors contributing to COPD (COPD tends to be clustered in families). Silverman and coworkers used spirometric measurements to compare lung function between families of early onset COPD smoker probands and a control group. They showed that in families of COPD probands the FEV1% and FEV1/FVC levels were both significantly lower than those in the healthy controls. Similarly, the former group had higher risk for developing chronic bronchitis than the control group, highlighting the importance of genetic factors predisposing to the development of COPD in certain families (Silverman et al., 1998). Higher risk (by two to three times) for developing COPD in siblings of patients with COPD was shown by other studies indicating the importance of familial clustering of COPD within the smoking relatives of COPD probands (Kueppers et al., 1977, McCloskey et al., 2001).

Further evidence of familial aggregation of COPD was evident from twin studies, showing that monozygotic twins are at high risk for developing COPD if one of them develops the disease. This concordance was not complete, and it shared almost equally with environmental factors for developing COPD in these patients (Hubert et al., 1982, Redline et al., 1987, McClearn et al., 1994).

### 1.1.5.2.2 Alpha-1-antitrypsin (A1AT) deficiency:

A1AT is a serine protease inhibitor. It irreversibly inhibits serine proteases secreted by neutrophils (neutrophil elastase (NE), proteinase 3 and cathepsin G), with NE being its principal substrate (Barnes et al., 2003). This action protects the lung against proteolytic damage. In 1963, Laurell and Eriksson were the first to note the absence of the α1-band on protein electrophoresis in samples from patients with early-onset emphysema (Laurell and Eriksson, 1963). The A1AT deficiency accounts for only 1-2% of all COPD cases (Ganrot et al., 1967, Silverman et al., 1998, Kalsheker and Chappell, 2008), and the disease affects individuals from different racial groups worldwide (de Serres, 2002).

### 1.1.5.2.3 Candidate gene studies:

Determination of genetic susceptibility for COPD remains a big challenge. There are many genetic associations with COPD identified by candidate genes. These have been inconsistently replicated (Castaldi et al., 2010). The only exception for this is the SERPINA1 gene (that encodes for A1AT), that has been proven to affect susceptibility of COPD (Stoller and Aboussouan, 2005).

Candidate gene studies showed no association with COPD in many genes. This was basically due to ethnic group selection and the number of participants (Celedon et al., 2004, Hersh et al., 2005, Castaldi et al., 2010). In a recent meta-analysis of candidate gene studies for COPD, it was suggested that the genes involved with COPD development can be classified into three main pathways: the inflammatory, protease/antiprotease and the oxidative stress pathways (Smolonska et al., 2009). However, among the variations studied, only SERPINA1 variations showed a well recognised association with COPD.

### 1.1.5.2.4  GWAS studies for COPD

These are non-hypothesis driven studies, designed to identify association of genetic variants with common diseases in a given population. This type of study uses somewhat evenly distributed variation across the genome which helps to identify specific areas of the gene that are significantly associated with the disease phenotypes (Pearson and Manolio, 2008). It was suggested that GWAS for COPD could be a valuable tools to reveal further genetic association to this disease other than the A1AT deficiency (Kalsheker and Chappell, 2008).

The first GWAS for COPD reported by Pillai and colleagues in 2009 identified a strong association of COPD with two α-nicotinic acetylcholine receptor (nAChRs) loci on chromosome 15 (rs1051730 and rs8034191 on CHRNA3/5 gene locus). This locus was suggested to represent a major susceptibility for COPD in the general population (Pillai et al., 2009). Furthermore, the same alleles of this locus have been reported recently to be associated with lung cancer and nicotine addiction (Amos et al., 2008, Hung et al., 2008,

Thorgeirsson et al., 2008). Another GWAS of pulmonary function in the Framingham Heart Study identified a SNP on chromosome 4q31 (rs13147758) that was strongly associated with FEV1/FVC ratio (Wilk et al., 2009). Recently, a study by Repapi and colleagues in 2010 reported a meta-analysis of GWAS of pulmonary function in individuals of European ancestry (Repapi et al., 2010). This study confirmed the previously reported association of pulmonary function with SNPs detected at 4q31 locus by the Wilk et al. study, and showed that five novel, previously unreported, loci were robustly associated with pulmonary function that represent good candidate genes for lung development or renewal pathways.

It is worth mentioning that the results for the two SNPs at CHRNA3/5 locus reported by Pillai et al could not be replicated by the other two GWAS studies of Wilk et al and Repapi et al, possibly due to investigating different phenotypes of COPD. Interestingly, in a recent meta-analysis study to test for susceptibility to COPD, SNP (rs1051730) (in the CHRNA3/5 locus) was found to have a strong correlation with COPD development, showing an association with lower FEV1 % predicted values and increased the risks of developing emphysema (Zhang et al., 2011). Although α7 nAChR gene (CHRNA7) was not in LD with the CHRNA3/5 locus, it is still functionally related to the disease (as will be discussed later in this chapter).

### 1.1.5.3 Inflammation

Inflammation is one of the defense mechanisms used by the immune system to protect the human body against pathogens. Normally this is well orchestrated

by the immune system in a precise and rapid manner. The process involves many cells, such as tissue macrophages, neutrophils and epithelial cells. These cells will release several inflammatory mediators (like Tumor necrosis factor (TNF), Interleukin-1β (IL-1β), IL6, IL8, IL18, and High-mobility group box1 (HMGB1)), causing the start and progression of inflammation (Tracy, 2002; Pavlov and Tracy, 2004; Wang et al, 2004).

It is well known that TNF (previously known as TNF-α) is one of the earliest pro-inflammatory cytokines released during inflammation. It mainly functions to control invasion and help promote healing (Tracey, 2002). TNF was identified by Beutler and colleagues in 1985 as a hormone causing cachexia, who later on that same year identified its key role in mediating septic shock in response to  infection (Beutler et al., 1985a, Beutler et al., 1985b). TNF is produced by many cell types, like macrophages, monocytes and neutrophils. TNF activates and prolongs the inflammatory response by activating other cells to release cytokines (such as HMGB1) (Wang et al., 1999, Wang et al., 2004). In the typical and successful inflammatory response, the TNF release is kept to the minimally effective level and period of secretion without progression to systemic TNF release. After stimulation with gram negative bacterial endotoxin/lipopolysaccharide (LPS) (0.1 mg/kg), the levels of TNF in experimental animals were elevated in the serum, the liver and the spleen, confirming that TNF is secreted by components of the mononuclear phagocyte system (formerly known as the reticuloendothelial tissue) (Wang et al., 2003).

Cytokines and interleukins (IL) are small molecules with a profound activity exerted at very low concentrations (Picomolar or femtomolar) (Ulloa and Tracey, 2005).

The Central nervous System (CNS) on the other hand, will limit this response from spreading via stimulating the vagus nerve (Borovikova et al., 2000) (this part will be discussed further in the anti-inflammatory pathway).

### 1.1.5.3.1 COPD and inflammation:

*What role does inflammation plays in COPD?*

The bronchial epithelium is one of the important lines of defense against infections, and is enriched for inflammatory mediators. COPD was suggested to be a chronic inflammatory disease (Nathan, 2002, Barnes et al., 2003). This was supported by the presence of large numbers of alveolar macrophages and neutrophils within the lung tissues and secretions. This was shown to be concomitant to increased release of inflammatory mediators, such as TNF, IL8 and other inflammatory cytokines. However, the release of these mediators may be modified by cigarette smoking which may be the case in COPD. The alveolar macrophages seem to have a central role in the start of inflammatory changes of the lungs which may subsequently precipitate COPD. These changes include recruiting neutrophils and other inflammatory cells, release of pro-inflammatory cytokines and release of elastase enzyme. The latter can affect the lung parenchyma and alveolar cell walls leading to the development of emphysema. This effect together with the presence of chronic inflammation of the bronchi (chronic bronchitis) will form the components of COPD (Barnes et al., 2003). Thus, several pulmonary cell types (the bronchial epithelial cells, alveolar epithelial cells, neutrophils and the pulmonary macrophages) seem to contribute to the development of COPD. It is well established that the chronic inflammatory condition in COPD is affecting the respiratory epithelial cell

remodeling and regeneration which can further complicate the tissue texture (Maouche et al., 2009). These inflammatory changes in COPD function via a pathway that involves the nicotinic acetylcholine receptors (as mentioned below).

## 1.2   Nicotinic acetyl choline subunits: a family of receptors

The nicotinic acetyl choline receptors (nAChRs) are a group of receptors that form part of the superfamily of ligand-gated ion channels. The nAChRs are characterised by ion selectivity to cations (Seguela et al., 1993).

In humans, the nAChRs constitutes 16 members (discovered so far) and these are expressed in different tissues. These include α-subunits (1-10, except α8), β-subunits (1-4) in addition to γ, δ and ε subunits (Albuquerque et al., 2009). The α-subunits differ from the other subunits in having two cysteine residues in close proximity within the amino terminal end. This will form the cysteine loop which is part of the agonist binding site helping to establish the functional role of the receptor (Boyd, 1997, Albuquerque et al., 2009). Generally, these subunits bind to acetylcholine (ACh) and to nicotine as agonists. The nAChR subtypes α2-6 bind with high affinity to nicotine and requires β-subunits 2-4 for forming heteropentameric receptors. On the other hand, α7-10 has lower affinity to bind nicotine but have high affinity to bind alphabungarotoxin (αBTX), and tend to form homopentameric receptors. Of this group, α7 nAChR is the most abundant in the mammalian brain tissue, while α8 was only recognised in chickens and α9 & 10 were shown to have limited expression in a range of human tissues, such as outer hair cells of the inner ear (Ulloa, 2005, Albuquerque et al., 2009).

This receptor family (nAChRs) was identified in mammals, including Humans. However, it was suggested that this family developed in Humans from the α7-nAChRs which involved gene duplications and mutations over time to yield this group of receptors (Le Novere and Changeux, 1995, Hurst et al., 2013).

### 1.2.1   General structure for nAChR subunits:

All these members share a common structure formed of three major domains: the extracellular domain, the intracytoplasmic domain and the transmembrane domain (Tsetlin et al., 2011). Generally, the extracellular domain is composed of a long amino terminal end (N-terminus) of more than 200 amino acid chain and a short carboxy terminal end (C-terminus). The transmembrane domain is composed of several transmembrane domains (1-4) and the intracellular cytoplasmic domain is composed of the relatively long (more than 100 amino acids) intracytoplasmic loop between transmembrane domains 3 and 4 (Figure 1-3) (Brejc et al., 2001).

**A**                                                                **B**
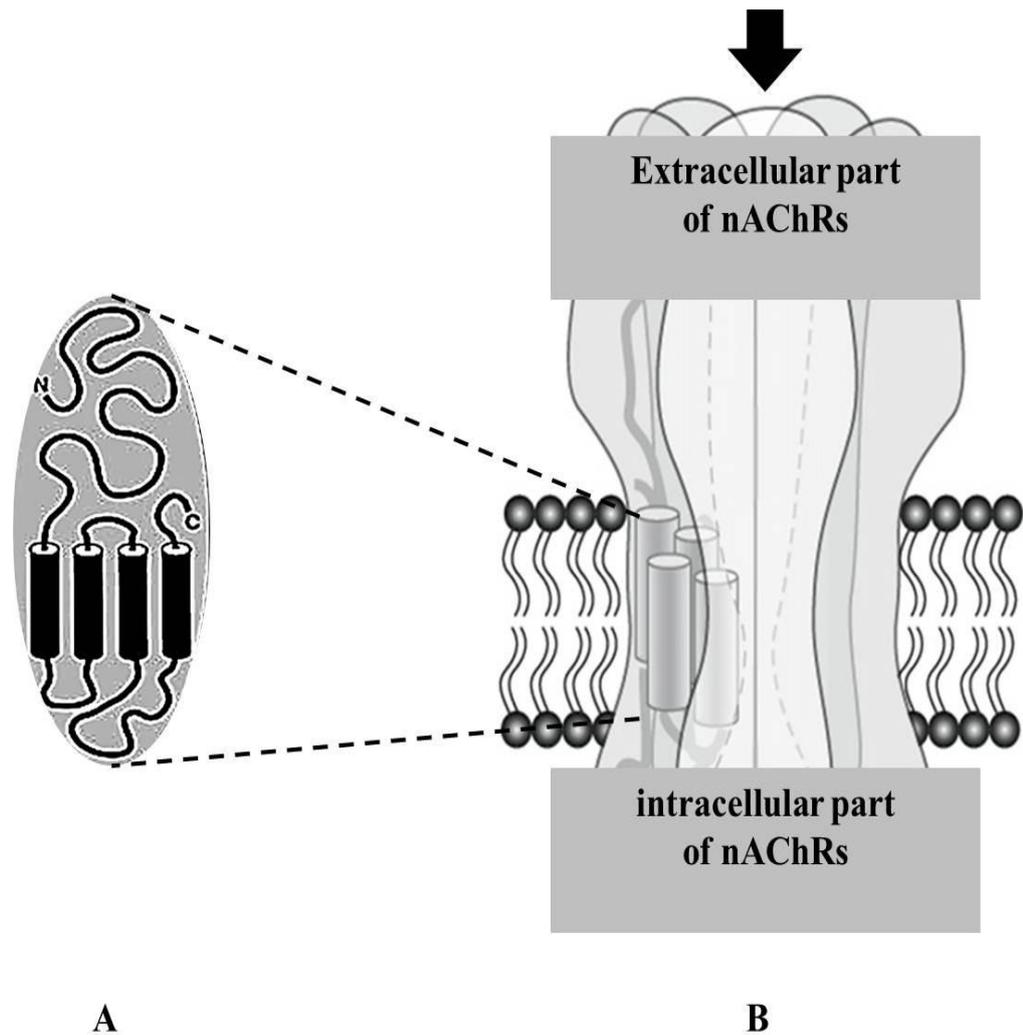
**Figure 1-3: General structure for nAChRs.**

A: single subunit of nAChRs showing composition of three main parts: extracellular part (including N-terminus (N) and C-terminus (C)), transmembrane parts (black cylinders) and intracytoplasmic loop (between TMD3 and 4, TMD1 is the farthest left).

B: schematic representation of five nAChR subunit assembly enclosing ion channel (the pore of which is indicated above the channel-arrowed-).

Figure was modified from (Millar, 2003, Hurst et al., 2013).

The N-terminus constitutes the major part of the receptor and includes the signal peptide, the agonist binding sites and the glycosylation sites. This part is followed by the transmembrane domains (TMDs 1-4) that help stabilise the receptor into the cell membrane (Unwin, 2005). The TMD2 lines the core of the ion channel, while TMD 1 and 3 separate TMD2 from the lipid layer of the membrane, with TMD 4 being the farthest from the rest (Taly et al., 2009). The intracytoplasmic loop connects between TMD 3 and 4 and includes three phosphorylation sites (Jansen et al., 2008, Kracun et al., 2008).

A summary of the parts forming the nAChRs is shown in (Table 1-3).

**Table 1-3:  The nAChRs domains.**

| nAChR Part | Included sites | aa | Position |
|---|---|---|---|
| **N-terminus** | Contains SP, LB, GLY | 200 | extracellular |
| **TMDs 1-4** | - | 19-24 | transmembrane |
| **Intracellular loop** | 3P | >100 | cytoplasmic |
| **C-terminus** | - | - | extracellular |
| N-terminus: amino terminal end, SP: signal peptide, LB: ligand binding site, GLY: N-glycosylation site, P: phosphorylation site, C-terminus: carboxy terminal end, aa: number of constituting amino acids. | | | |

### 1.2.2   Subunit assembly and posttranslational modifications:

The nAChR subunits can assemble into homomer or heteromer pentameric receptors that surround a ligand-gated ion channel pore. Almost all nAChR subunits can form heteropentamer receptors, while α7 are recognized in Humans to form a homopentamer receptor (Corringer et al., 2000, Millar, 2003, Wu and Lukas, 2011). However, heteropentameric receptors of α7 combined with β2 subunits could be detected from rodent brain (Figure 1-4) (Liu et al., 2009).
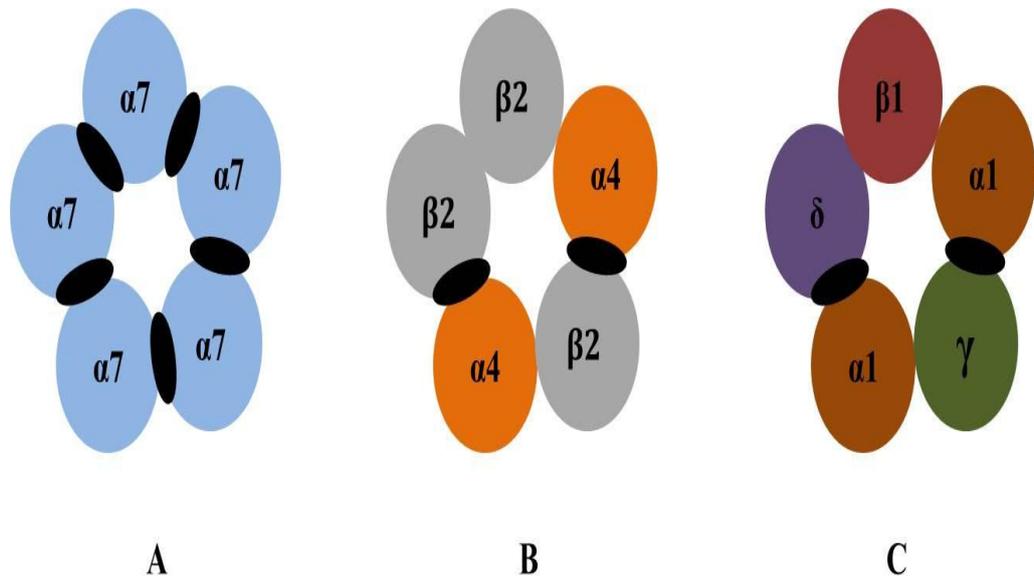
**Figure 1-4: Schematic representation of some of nAChRs pentamers.**
A: α7-nAChR subunits assembled into homopentameric receptor expressed on the cell surface of neuronal and non-neuronal cells. B: α4β2-nAChR subunits assembled into heteropentameric receptor expressed on the cell surface of neuronal cells. C: α1β1δγ nAChR subunits assembled at the neuromuscular junction (muscle type nAChRs). The ACh (agonist) binding sites are shown as black ovals between α-subunits or between α-and nonα-subunits.
Adapted from (Corringer et al., 2000, Albuquerque et al., 2009, Wu and Lukas, 2011).

Although α7-nAChRs are the best described homopentameric receptors, they are not the only ones, and similar homopentamers were described from α9 and α10 subunits (Hurst et al., 2013). However, the fact that α7-nAChRs are widely distributed in human tissues while α9 and α10 nAChRs are less widely expressed in human tissues (α9 and α10 were shown to be expressed by the outer hair cells of the inner ear while only α9 was expressed by neurons involved in pain perception) is determining the α7-nAChRs as the prototype for such a group of receptors, at least so far. In addition, the wide expression of

α7-nAChRs enabled researchers to use this subunit for studies involving the ion channels' function, drug design experiments and signaling pathways testing (Bertrand D., 1995, Bertrand et al., 2008).

Following the translation of nAChR subunits mRNAs, the polypeptide formed will be subjected to a number of posttranslational modifications that take place within the endoplasmic reticulum (ER) till the receptors are expressed on the cell surface (Green and Millar, 1995). First, the polypeptide chains would fold in a way that helps establish the correct secondary structure of nAChR protein. Then, the folded subunits would gather to assemble and form the pentamer receptor. This will include establishing the agonist binding sites at the subunit interfaces (between α subunits for homopentamer receptors, or between both the α- and β- subunits for heteropentamer receptors). This preference would leave heteropentamer receptors with two agonist binding sites compared to five for homopentamer ones (Unwin, 2005, Hurst et al., 2013). Finally, the N-glycosylation sites and the disulfide bonds are added to the N-terminus (Gehle et al., 1997).

Different parts of the nAChRs protein contribute to the assembly and function of the receptor. It has been established that the N-terminus is responsible for the receptor assembly and formation of the agonist binding sites in nAChRs (Verrall and Hall, 1992, Brejc et al., 2001). Similarly, the cytoplasmic loop is involved in subunit folding, assembly, trafficking and efficient membrane insertion of the protein in addition to ensuring the formation of functional receptors  (Jansen et al., 2008, Kracun et al., 2008). This may be partly due to the effect of three phosphorylation sites that are essential for stimulating the subunit assembly and for the control of receptor expression, desensitization and

function of the receptor (Green et al., 1991, Guo and Wecker, 2002).  Although

the C-terminus is relatively short and represents a small part of the protein, its

presence within the nAChRs protein is essential for enhancing the protein

transport to the surface and subsequent binding site function (Pons et al.,

2004).

The expression of α7-nAChR proteins was found to be not always correlated

with their corresponding mRNA levels (Blumenthal et al., 1997). This effect

was shown to affect some of the normal and cancerous respiratory cells and

cell lines that might explain a defective translation, assembly or the relatively

scarce protein concentrations (Carlisle et al., 2004, Plummer et al., 2005,

Carlisle et al., 2007).


### 1.2.3   Agonist binding and channel opening:

The nAChR receptors function mainly as ligand-gated ion channels, lining

cation channels that permit calcium, and to some extent sodium. This results in

the influx of calcium (or sodium) inside the cell which will help upregulate

certain genes (Dani and Bertrand, 2007). Acetylcholine (ACh) is well known

for acting as an endogenous agonist for nAChRs. During the resting condition

of these receptors, the ion channel is closed and no ions are permitted to enter

the cell. Once ACh or nicotine binds to the agonist binding site of the receptor,

a conformational change of the receptor results in opening the channel and ion

influx inside the cell (Wu and Lukas, 2011). This effect tends to be short-lived

(for few milliseconds) and the receptor will return to the resting state (Dani and

Bertrand, 2007, Taly et al., 2009). However, a prolonged stimulation of

nAChRs with ACh or nicotine will cause 'desensitization' of the channel (non-

responsive state of the channel) (Figure 1-5) (Galzi et al., 1992, Hurst et al.,

2013).
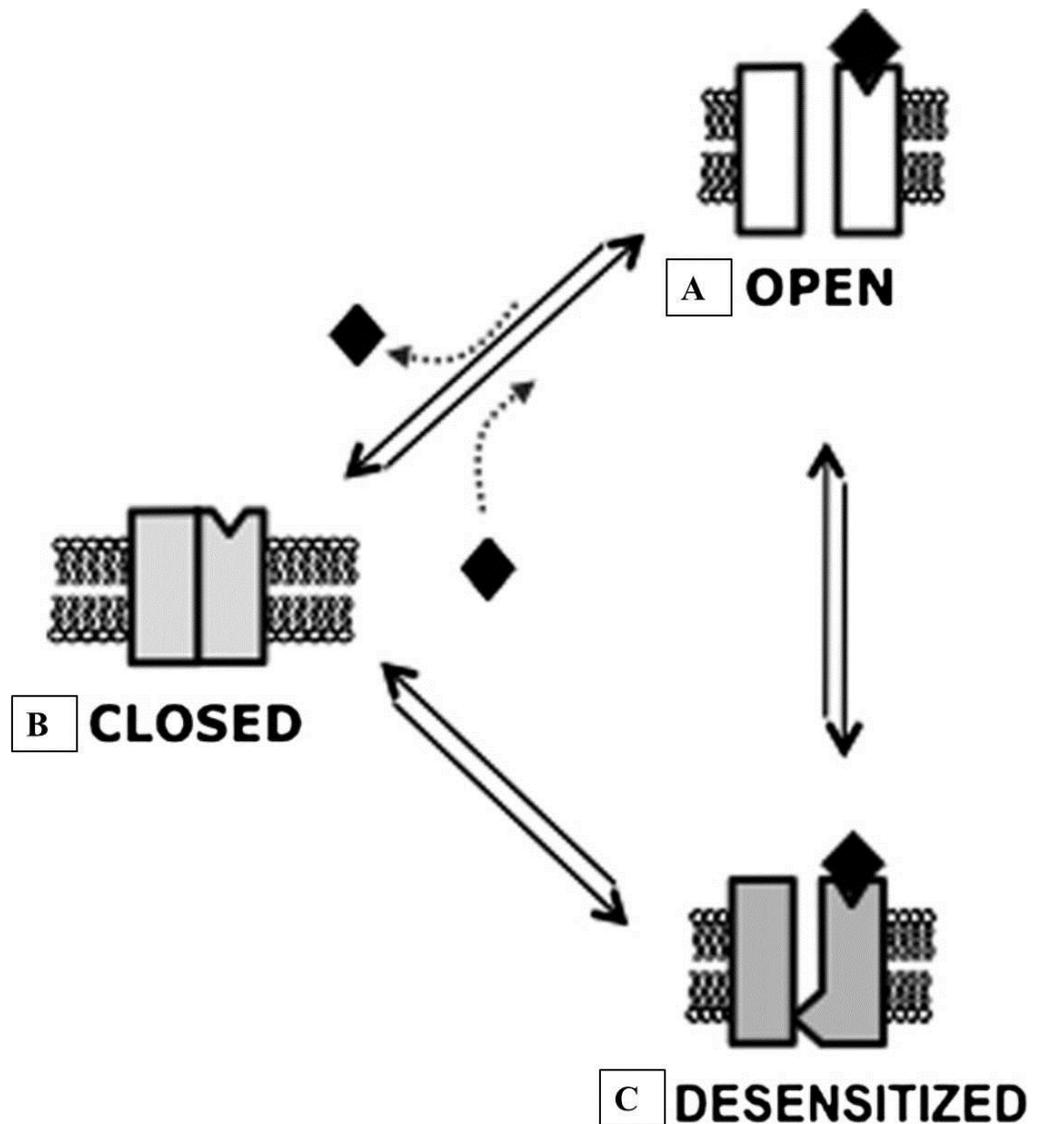


**Figure 1-5: A Diagram showing the different phases of nAChR functions.**
After agonist binding, the channel opens and the ions influx inside the cell (A).

During the resting phase, no agonist is bound to the receptor and the channel is

closed (B). This will cause the start of a cascade of events ending up with

upregulation of a gene or stimulation of a second messenger-signalling

pathway. This stimulatory effect will last only for short time, and will be

followed subsequently by a desensitisation phase of the channel (the state of

non-responsiveness) (C). Adapted from (Hurst et al., 2013).

The nAChRs channels function in a similar mode by allowing cation infusion inside the cell and lead to the upregulation of certain genes or cascades of pathways (Albuquerque et al., 2009). The ion selectivity of nAChRs is dependent on few amino acids that form part of the TMD 2 and TMD 2-TMD 3 loop and changes in such amino acids could change the channel into an anion selective channel (Galzi et al., 1992, Lummis et al., 2005). In contrast, an important competitive antagonist for nAChRs is the alpha bungarotoxin (α-BGT). This chemical compound is a component of snake venom that binds in an irreversible reaction with nAChRs leading to inhibition of the ion flux inside the cell (Albuquerque et al., 2009). For this reason, α-BGT was used by many studies to precipitate these receptors for further investigations. The nAChRs were classified according to their pattern of binding with this compound into α-BGT sensitive or insensitive receptors, with expression intensities varying from no or mild to moderate or even high expression (Kracun et al., 2008). Nicotine was used experimentally in high concentrations to block α-BGT irreversible binding to nAChRs (Wang et al., 2003, Govind et al., 2009). The need for high nicotine doses was justified by the fact that α7-nAChRs have low affinity for nicotine. Interestingly, nicotine seems to have a role in activating these receptors via several methods, including competitive agonist binding and blocking factors that suppress α7-nAChRs expression (Araud et al., 2011).

### 1.2.4   Association with human diseases:

The major association of nAChRs with human diseases was via nicotine addiction. Nicotine is a known agonist to nAChRs, but a prolonged exposure to nicotine, as the case for chronic cigarette smokers, can lead to increased

receptor desensitization to nicotine over time. This will result in the up-regulation of receptors with high affinity to nicotine such as α4β2 nAChRs in an attempt to increase receptor sensitization (Govind et al., 2009). Similarly, chronic exposure to nicotine leads to upregulation of the α7-nAChRs (Peng et al., 1999, Wang et al., 2001, Fu et al., 2009).

It was documented that more than two thirds of the nAChRs are retained within intracellular pools, and that only one third was allowed to be transported to the cell surface for expression and functioning (Albuquerque et al., 2009). During this intracellular pooling, many subunits will be degraded, a mechanism that is counteracted by nicotine. In cases of increased nicotine concentrations, fewer subunits are subjected to degradation, and thus more subunits are allowed to be expressed on the surface allowing for the receptor upregulation. Thus, the effect of chronic nicotine exposure, as the case for chronic smoking, can interfere with the proportion of expressed receptors, and leading to the development of certain diseases (Albuquerque et al., 2009) (This part will be discussed in the final discussion).

It is worth mentioning that the stimulatory effect of nicotine on α7- nAChRs is more than that exerted by ACh making the former a favorable drug target for stimulating this subset of receptors (Ulloa, 2005).

## 1.3   α7-nAChR subunit:

This is the only nAChR that is known to assemble into homopentameric receptors, with the earlier studies reporting this in chicken brain tissues (Couturier et al., 1990). The comparison of the chicken and rat α7-nAChRs showed high similarity with human α7-nAChRs at mRNA and protein levels

(Lynn Doucette-Stamm, 1993, Chini et al., 1994, Peng et al., 1994). This receptor was studied extensively due to its unique homopentameric composition in human and animal tissues. It was speculated that due to this property, and due to the high degree of similarity of α7-nAChRs with other mammalian nAChRs, that this receptor forms the ancestor of all nAChRs and during evolution, the other types of nAChRs developed by duplication, inversion and deletion (Albuquerque et al., 2009, Le Novere and Changeux, 1995). It is important to know that RIC-3 (resistance to inhibitors of choline esterases) protein is an essential chaperone for α7-nAChRs to enhance their functional receptor expression. This effect is lost with other heteropentameric receptors such as α4β2 or α3β4 (Halevi et al., 2002).

### 1.3.1   CHRNA7:

This gene codes for the α7-nAChR subunits in human tissues. It is located at chromosome 15q13 region and constitutes ten exons embedded within more than 140 kb of intronic sequences (Gault et al., 1998). The promoter region of this gene was shown to be about 77% GC-rich sequence which is lacking the TATA box and includes binding sites for transcription factors, such as SP1, AP-2, CREB and Egr-1 (Carrasco-Serrano et al., 1998). Not all these transcription factors seem to contribute to the upregulation of α7-nAChR. A good example is the transcription factor AP-2α was found to be negatively controlling the α7-nAChRs mRNA expression (Finlay-Schultz et al., 2011). The CHRNA7 exons 1-10 are relatively short exons spanning around 100 bp or less in size except exon 10 (Figure 1-6).
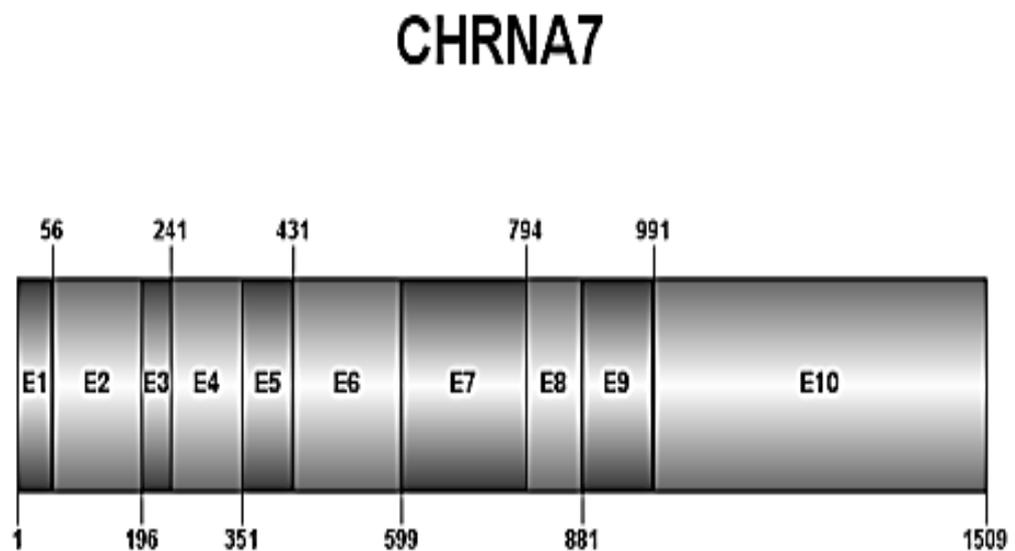


**Figure 1-6: CHRNA7 exons.**
Exons (E) 1-10 demonstrated with the nucleotide sequence (at the start of each exon) shown alternating on the upper and lower panels.

However, the introns connecting these exons vary widely in size from relatively short introns of 300 bps up to large introns of more than 25 kb size (Table 1-4).

**Table 1-4: CHRNA7 exons and introns nucleotide number and cDNA position.**

| | Exon details | | Downstream intron details | |
|---|---|---|---|---|
| **Exon no.** | Exon size (bp) | cDNA position | Intron no. | Intron size (bp) |
| **1** | 55 | 1-55 | 1 | 248 |
| **2** | 140 | 56-195 | 2 | 70,265 |
| **3** | 45 | 196-240 | 3 | 10,440 |
| **4** | 110 | 241-350 | 4 | 42,007 |
| **5** | 80 | 351-430 | 5 | 3,621 |
| **6** | 168 | 431-598 | 6 | 636 |
| **7** | 195 | 599-793 | 7 | 962 |
| **8** | 87 | 794-880 | 8 | 3,570 |
| **9** | 110 | 881-990 | 9 | 4,604 |
| **10** | 519 | 991-1509 | - | - |
| **CHRNA7 accession number NM_00074, NCBI and from (Gault et al., 1998).** | | | | |

**1.3.2    α7-protein:**

This protein showed similar structure to other nAChRs proteins. However, unlike most heteropentameric nAChRs, α7-nAChRs are characterized by their low affinity to ACh and nicotine but with high affinity to bind αBTX (Orr-Urtreger et al., 1997, Ulloa, 2005). This characteristic was used extensively to precipitate α7-nAChRs to study the physical properties of the receptor (Couturier et al., 1990, Rangwala et al., 1997, Drisdel and Green, 2000). However, using nicotine in high concentrations prior to adding αBTX can block the latter binding to these receptors (Wang et al., 2003).

**1.3.2.1 The significance of specific amino acid sequence forming α7-nAChRs:**

The amino acids of the α7-nAChR are arranged in a complex conformation which directs the cation selectivity of the channel. Besides cation selectivity properties, α7-nAChRs are much more permeable to calcium ions than to sodium and potassium ions depending on the unique amino acid structure (Seguela et al., 1993). This high permeability to calcium ions is a unique property of α7-nAChRs and it was correlated with cellular protection. Changes that can affect these receptors can affect the calcium ion balance that is necessary for vital cell processes and thus affecting the cell survival (Uteshev, 2012).

This architecture of α7-nAChRs is important for maintaining the receptor function and changes to certain amino acids with others of different properties can change the ion channel properties (Bertrand et al., 1993, Dani and Bertrand, 2007). The use of site-directed mutagenesis made it easier for

predictions of the effect of amino acid changes on the protein structure, assembly and channel function. Consequently, it was shown that the change of a single amino acid near the extracellular space from the TMD2 can alter the ion selectivity of the channel from cationic into anionic (Galzi et al., 1992). Similar results were detected when substituting charged amino acids at the amino terminal end of TMD 2 of the receptor with a hydrophobic side-chain amino acid decreased the ion channel permeability to calcium ions (Bertrand et al., 1993). Furthermore, the substitution of two amino acids of the same group, both with hydroxyl side chains, led to gain-of-function effect on the α7-nAChRs (Placzek et al., 2005). On the other hand, a loss-of-function of the receptor was detected when substituting an amino acid within the intracytoplasmic loop with an uncharged amino acid. This change caused prolonged and reversible desensitization of the mutant α7-nAChRs in response to protein kinase C activators (Tsuneki et al., 2007).

**1.3.2.2 Posttranslational modification of α7-nAChR:**

There are a number of posttranslational modifications that affect the assembly and functional expression of the α7-nAChRs (Blumenthal et al., 1997, Rakhilin et al., 1999). Among these is RIC-3 protein encoded by the gene hric-3 in humans (Halevi et al., 2003). This chaperone protein plays an important role by enhancing the α7-nAChRs folding, assembly and functional expression (Halevi et al., 2002, Halevi et al., 2003, Lansdell et al., 2005, Williams et al., 2005). Some α7-nAChRs might escape this necessity and form receptors on the cell surface in the absence of RIC-3. Such receptors were proved to be nonfunctioning, a condition that was concomitant to increased intracellular

accumulation of nAChRs. This effect was shown either in cells expressing low levels of RIC-3 or in cells with mutated nonfunctioning form of the protein (Cooper and Millar, 1997, Rakhilin et al., 1999, Halevi et al., 2002, Williams et al., 2005).

Interestingly, the N-glycosylation sites within the N-terminus of the α7-nAChR subunits were shown to affect the α7-nAChRs function without interfering with the receptor surface expression in spite of being located away from the ligand binding site and thus are not directly affecting the agonist binding process (Chen et al., 1998, Avramopoulou et al., 2004). The folding of α7-nAChRs may also play a role in the receptor assembly and function. As far as the disulfide bond is concerned, receptors assembled into single conformation were shown to form non-functional receptors compared to the functional ones formed by the double conformation receptors (Rakhilin et al., 1999, Drisdel and Green, 2000). Moreover, the phosphorylation of certain tyrosine residues within the intracytoplasmic loop can lead to inactivation of the α7-nAChRs and thus affecting the ion channel function negatively. This is supported by the findings of increased α7-nAChRs activity after dephosphorylation (Charpantier et al., 2005, Fu et al., 2009). It is worth noting that another posttranslational process, the palmitoylation, can enhance the α7-nAChRs assembly and the ligand binding site formation, leading to the formation of functional nAChRs (Drisdel et al., 2004, Alexander et al., 2010).

By comparing the results obtained for CHRNA7 in humans to that in chicken and rats, it was shown that α7-nAChR is similar to other nAChRs in general structure (Table 1-5), (Figure 1-7).

**Table 1-5: The parts of the Human α7-nAChR.**

| Exon | Protein part encoded | aa |
|---|---|---|
| **5' UTR** | 77 % GC rich, lack TATA box, contains consensus SP1, AP2, Egr1 and CREB transcription binding sites | - |
| **1** | SP (aa 1-22) | 18 |
| **2** | last part of signal peptide, GLY (aa 46) | 46 |
| **3** | - | 15 |
| **4** | GLY (aa 90), LB (aa 107-117) | 36 |
| **5** | GLY (aa 133) | 26 |
| **6** | LB (aa 169-175) | 56 |
| **7** | LB (aa 207-219), TMD1(aa 231-254) | 65 |
| **8** | TMD2 (aa 261-283) | 29 |
| **9** | TMD3 (aa 293-315) | 36 |
| **10** | TMD4 (aa 472-494), 3 P (aa 367, 417, 444), intracytoplasmic loop (aa 318-469), C-terminus (aa 495-502) | 173 |
| N-terminus: amino terminal end, SP: signal peptide, LB: ligand binding site, GLY: N-glycosylation site, P: phosphorylation site, C-terminus: carboxy terminal end. Two disulfide bonds present between aa at positions: 150-164 and 212-213. aa: number of constituting amino acids. Adapted from (Gault et al., 1998, de Jonge and Ulloa, 2007). | | |
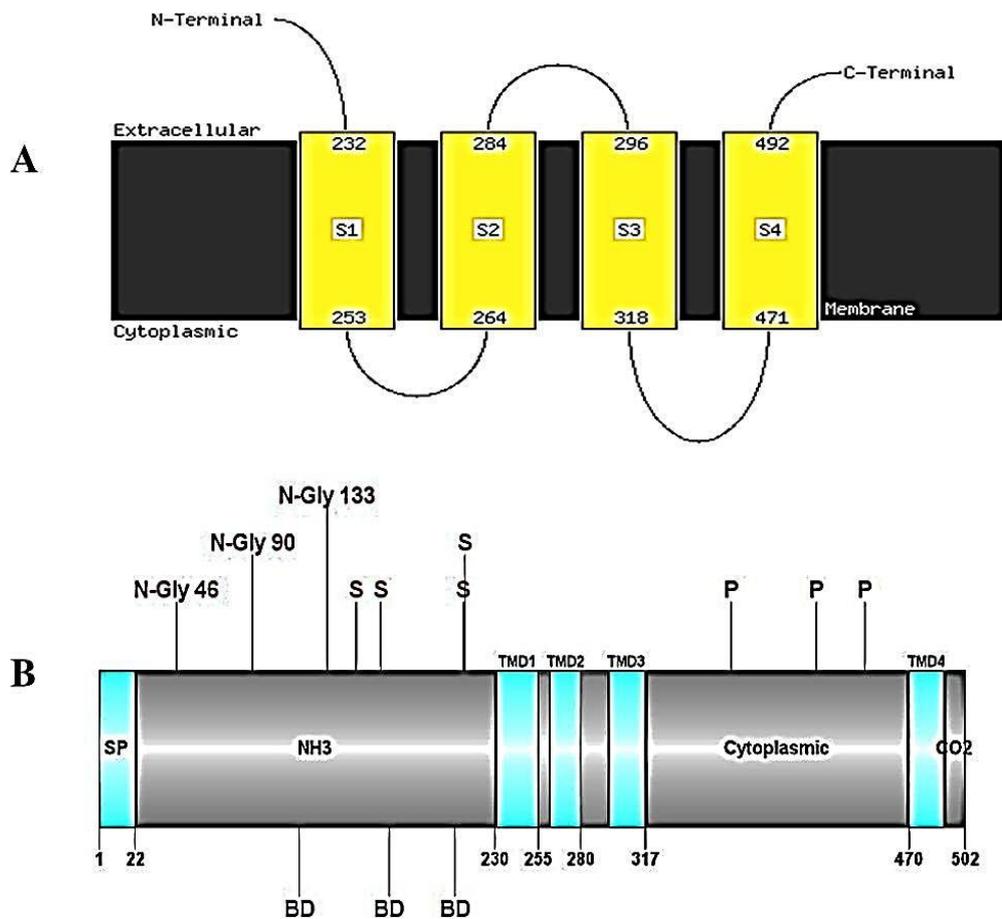
**Figure 1-7: schematic representation of α7-nAChR protein domains.**
A: showing secondary structure for nAChRs using PSIPRED software. The major parts of the protein are indicated (N-terminal and C-terminal ends extracellularly, TMDs 1-4 (S1-4), and intracytoplasmic loop between TMD 3 and 4). The intracytoplasmic loop between TMD 1 and 2 is small and has a less important role on the receptor's assembly and function than the loop between TMD 3 and 4 and the term intracytoplasmic loop refers usually to the latter.

B: schematic representation of nAChR protein structure using CSS palm 3.0 software program. The terms NH3 and CO2 are referring to the N-terminus and C-terminus, respectively. TMDs1-4 are indicated above the figure. SP: signal peptide, cytoplasmic: intracytoplasmic loop, BD: binding domains (corresponding to agonist binding sites), N-Gly: N-glycosylation sites (with amino acid sequence within the protein indicated), S: refers to sulphur atom within cysteine amino acids, P: refers to phosphorylation sites. The amino acid positions of the corresponding parts are indicated on the lower panel of the figure.

Adapted from (Gault et al., 1998) and (Millar, 2003).

**1.3.2.3 The α7-nAChRs expression in different cells:**

There were several studies that tested for the expression of α7-nAChRs at mRNA and/or protein level in different tissues. The homo-oligomeric nature of this receptor helped to shed light on this receptor as a model for heterogeneous expression in cells that do not express α7-nAChRs. The first successful attempt was made by using human recombinant α7-nAChR DNA transfection into cells that do not express nAChRs (Peng et al., 1999). This was followed by many studies to express α7-nAChRs in different types of human and animal tissues. Among the human tissues that expressed α7-nAChRs (mRNA and protein) are SH-SY5Y (human neuroblastoma cell line) (Peng et al., 1994), human brain tissue (Gault et al., 1998, Severance and Yolken, 2008) human bronchial epithelial and aortic epithelial cells (Wang et al., 2001), human fibroblast cells (Carlisle et al., 2004), and different human lung cancer cell lines, including A549 cell line (Carlisle et al., 2004, Plummer et al., 2005, Carlisle et al., 2007, Lam et al., 2007, Paleari et al., 2008). Recently, human mature adipocytes were shown to express functional α7-nAChRs (Cancello et al., 2012).

Similar work on animal tissues showed that it was expressed in different tissues, such as rat phaeochromocytoma cells from adrenal medulla (Rangwala et al., 1997), bovine adrenal chromaffin cells (García-Guzmán et al., 1995, El-Hajj et al., 2007), mouse brain tissue (Saragoza et al., 2003) and rat autonomic neurons (Severance et al., 2004). Most of the work done on α7-nAChRs could detect homopentameric receptors. However, heteropentameric α7-nAChRs were recently identified in animal tissues such as bovine chromaffin cells (El-Hajj et al., 2007) or rodent brain cells (Liu et al., 2009). These results are similar to earlier trials that involved expression of α7-nAChRs and β2-nAChRs

in Xenopus oocytes or human epithelial kidney cell line (tsA201) that yielded

heteropentameric functional receptors (Khiroug et al., 2002).

### 1.3.3   CHRFAM7A:

This gene represents partially duplicated CHRNA7 on chromosome 15,

(upstream of CHRNA7) with exons 5-10 of CHRNA7 duplicated in the same

or opposite orientation to the CHRNA7 (as will be described below). The

duplication forms a new gene called the hybrid α7 (or the fusion gene or the

cholinergic receptor family with sequence similarity 7A (CHRFAM7A)) after

combination with four novel exons (D to A) upstream of exons 5-10, so that

the orientation of the duplicated gene is in the following order: exons D, C, B,

A, then exons 5-10 (Gault et al., 1998) (Figure 1-8) (Table 1-6).
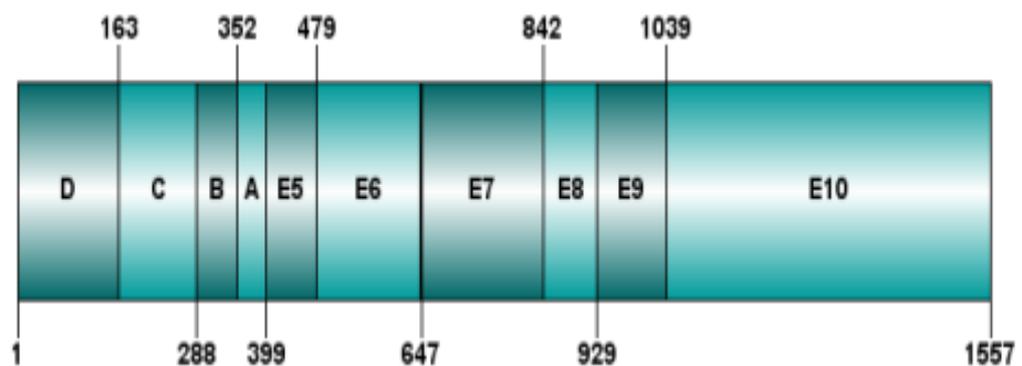


**Figure 1-8: CHRFAM7A exons.**
Exons (E) D-10 demonstrated with the nucleotide sequence (at the start of each exon)
shown alternating on the upper and lower panels.

**Table 1-6: CHRAFM7A exons and introns nucleotide number and cDNA position.**

| Exon no. | Exon details | | Downstream intron details | |
|----------|--------------|---|---------------------------|---|
| | Exon size (bp) | cDNA position | Intron ID | Intron size (bp) |
| **D** | 297 | 1-162 | D | 6,264 |
| **C** | 125 | 163-287 | C | 3,480 |
| **B** | 64 | 288-351 | B | 2,919 |
| **A** | 47 | 352-398 | A | 3,243 |
| **5** | 80 | 399-478 | 5 | 3,884 |
| **6** | 168 | 479-646 | 6 | 636 |
| **7** | 195 | 647-841 | 7 | 962 |
| **8** | 87 | 842-928 | 8 | 3,570 |
| **9** | 110 | 929-1038 | 9 | 4,606 |
| **10** | 519 | 1039-1557 | - | - |

CHRFAM7A accession number NM_139320.1, NCBI and from (**Gault et al., 1998, Araud et al., 2011**).

This gene was suggested to originate at some point during evolution due to a number of duplication and deletion steps that involved several genes at this locus (15q13-14) (Zody et al., 2006, Makoff and Flomen, 2007). Among these steps, duplication of CHRNA7 and associated or subsequent deletion of its first four exons (1-4) gave rise to the partial duplication segment of exons 5-10 (Riley, 2002). This partial duplication step was suggested to be located following another partial duplication of ULK4 gene on chromosome 15. The ULK4 gene original copy is located on chromosome 3p22, and several copies were detected partially duplicated (exons C-A) on chromosome 15. In all the partial duplication sites, another exon of unknown origin termed "exon D" was located upstream of the duplication giving rise to the composition of exons D-A (Riley, 2002, Araud et al., 2011). This "fusion" of genes combined with partial CHRNA7 duplication was suggested to aid the formation of the fusion or hybrid gene (the CHRFAM7A). This fusion was shown clearly when investigations on the break point that joins the two genes (CHRNA7 duplication and the "fusion" exons D-A) where part of intron 4 (of CHRNA7) and part of intron A (of ULK4) were found merged to form a hybrid intron connecting exons A and 5 within CHRFAM7A sequence. Most individuals seem to have the CHRFAM7A gene but some might miss one or both copies of the gene (Gault et al., 1998, Riley, 2002, Gault et al., 2003). It is not known if this geneotype difference of CHRFAM7A is associated with phenotypes of psychiatric disorders as the case is with CHRNA7 (Figure 1-9).

CHRNA7 (RefSeq NM_000746.3)

CHRFAM7A (RefSeq NM_139320.1)
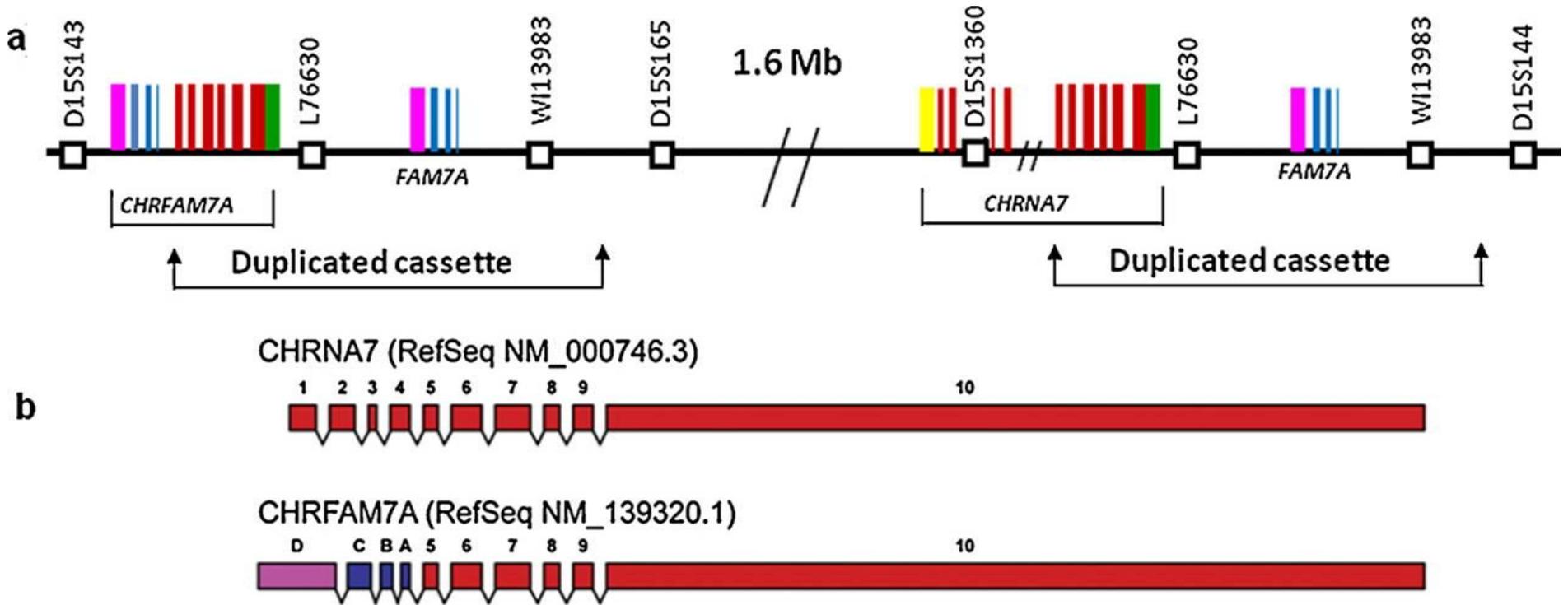
**Figure 1-9: chromosome 15q13-14 locus genetic composition.**

a) The location of CHRNA7 and CHRFAM7A on chromosome 15. Exons 1-10 of CHRNA7 shown as red bars, ULK4 exons C-A shown as blue bars, exon D is shown as pink bar and CHRNA7 promoter region is shown as a yellow bar while the 3`UTR of CHRNA7 is shown as a green bar. Exons D-A are duplicated on chromosome 15 several times, one of which is just upstream of partial duplication of CHRNA7 (CHRFAM7A). CHRNA7 located on the right represent the original copy of the gene and is located within a duplicated cassette on chromosome 15, about 1.6 Mb downstream of the other duplicated cassette that harbours CHRFAM7A; both cassettes are about 300 kb in size. Note that the locus D15S1360 is unique to CHRNA7 within intron 2, locus L76630 is located downstream of 3` end of CHRNA7 exons 5-10, locus WI13983 is located after two of exons D-A duplications while locus D15S143 is located just upstream of these exons in CHRFMA7 gene locus only.

b) The exon composition of CHRNA7 (exons 1-10) and of CHRFAM7A (exons D-A, then exons 5-10).

Taken from (Araud et al., 2011), and information adapted from (Gault et al., 1998, Riley, 2002, Araud et al., 2011, de Lucas-Cerrillo AM, 2011).

Interestingly, this fusion gene is unique to humans and could not be detected in animals (Locke et al., 2003).

### 1.3.4    CHRNA7 and CHRFAM7A: similarities and differences

**1.3.4.1 At gene level:**

Both genes, CHRFAM7A and CHRNA7, are located on chromosome 15 (15q13 and 15q14, respectively) with a 1.6 Mb intervening segment. Both genes have exons 5-10 within their sequence with almost 100% similarity (Gault et al., 1998). Gault et al showed that the orientation of the CHRNA7 is in the sense direction (Gault et al., 1998). This was the first study to show the CHRFAM7A duplication and constitutive exons and suggested that the orientation of the duplicated gene is in the same direction of CHRNA7. A subsequent study by Riley and co-workers conflicted with these results and showed that both genes were in the opposite direction (Riley, 2002). Subsequent studies showed that both results were correct and that the presence of a common polymorphism (2 bp deletion in exon 6 (Gault et al., 1998)) could be used as a marker for CHRFAM7A inversion (opposite to CHRNA7 orientation) (Flomen et al., 2006, Flomen et al., 2008). The identification of the 2bp deletion leads to categorisation of two major alleles of CHRFAM7A (with and without the deletion) that were suggested to have almost equal expression within populations expressing the gene. This means that CHRNA7 and CHRFAM7A are in the same or opposite orientation in almost equal proportions of the population harbouring them.

The copy number variants of both genes, CHRNA7 and CHRFAM7A, gave rise to different allelic constitution among different populations. Most populations have both copies of CHRNA7, but missing one or both copies can

lead to neuropsychiatric disorders (will be discussed shortly in the final discussion) (Shinawi et al., 2009, Endris et al., 2010, Masurel-Paulet et al., 2010, Spielmann et al., 2011, Hoppman-Chaney et al., 2013). Again, most populations have both copies of CHRFAM7A, with around 10% missing one copy and rarely missing both copies (Araud et al., 2011, de Lucas-Cerrillo AM, 2011). Missing both copies of CHRFAM7A was suggested to represent an ancestral genomic structure that persisted over time and escaped the duplication process (Makoff and Flomen, 2007). Interestingly, both genes originate from the sense strand (Severance and Yolken, 2008).

### 1.3.4.2 At genomic sequence level:

The reference sequences for both genes were arranged on separate files for further analysis. The exons, introns and SNP sites were allocated using three databases: *NCBI*, *UCSC* genome browser and *Ensembl*. It was documented that the similarities between both genes (between exons 5 and 10) are exceeding 99% (Gault et al., 1998). Our results were similar to those, indicating only three nucleotide differences within exons 7, 9 and 10, respectively (Table 1-7).

**Table 1-7: The three nucleotide base difference within exons 5-10 sequence between the two genes.**

| Gene | Exon7 (56)* | Exon 9 (53)* | Exon 10 (126)* |
|---|---|---|---|
| **CHRNA7** | C | A | C |
| **CHRFAM7A** | T | G | T |
| **\*:** the nucleotide position on the sense strand is shown between brackets representing the position within each exon (starting from the 5` end of the exon). All the sequence alignments were applied using *ClustalW* software program. | | | |

Using the NCBI database shows that there are three main transcripts deposited for CHRNA7. These include full length, inserted 87 bp of intron 1 and missing exons 4 & 5. However, for CHRFAM7A, only two transcripts were deposited which include the full length and missing exon B transcripts (Table 1-8). There are other mRNA transcripts of CHRNA7 and CHRFAM7A that were reported in Human and animal tissues but were not deposited in the NCBI database. These transcripts will be discussed in detail in chapter 2.

**Table 1-8: Summary for the main transcripts deposited in the NCBI database for CHRNA7 and CHRFAM7A.**

| Gene | Transcript | NCBI Accession | mRNA accession | mRNA size (bp) | Size of coding sequence (bp) | Protein accession | Protein size (aa) | comments |
|---|---|---|---|---|---|---|---|---|
| **CHRNA7** | Variant 1 | - | NM_000746 | 3356 bp | 1509 | NP_000737 | 502 | Full length |
| | Variant 2 | - | NM_001190455.2 | 3443 bp | 1596 | NP_001177384.1 | 531 | 87 bp of intron 1 inserted |
| | Variant 3 | NR_046324 | - | 3166 bp | 1319 | non-coding RNA | | Missing exons 4&5 |
| **CHRFAM7A** | Variant 1 | - | NM_139320.1 | 2858 bp | 1557 | NP_647536.1 | 412 | Full length |
| | Variant 2 | - | NM_148911.1 | 2794 bp | 1493 | NP_683709 | 321 | Missing exon B |
| **Bp: base pairs, aa: amino acids, NCBI database accession numbers: NM (for protein coding mRNA), NP(for protein), NR (for noncoding RNA)** | | | | | | | | |

**1.3.4.3 At translation level:**

Based on the differences in the first four exons of CHRFAM7A, the corresponding protein (dupα7 protein) would have some differences from the α7 protein. These include lack of the signal peptide encoded by exon 1, and lack of part of the agonist binding domain encoded by exons 2-4 (de Lucas-Cerrillo AM, 2011, Araud et al., 2011). However, there still some similarities between both proteins in the TMDs, IC loop and C-terminus parts as the exons coding for them (exons 7-10) are shared between the two genes (Figure 1-10).
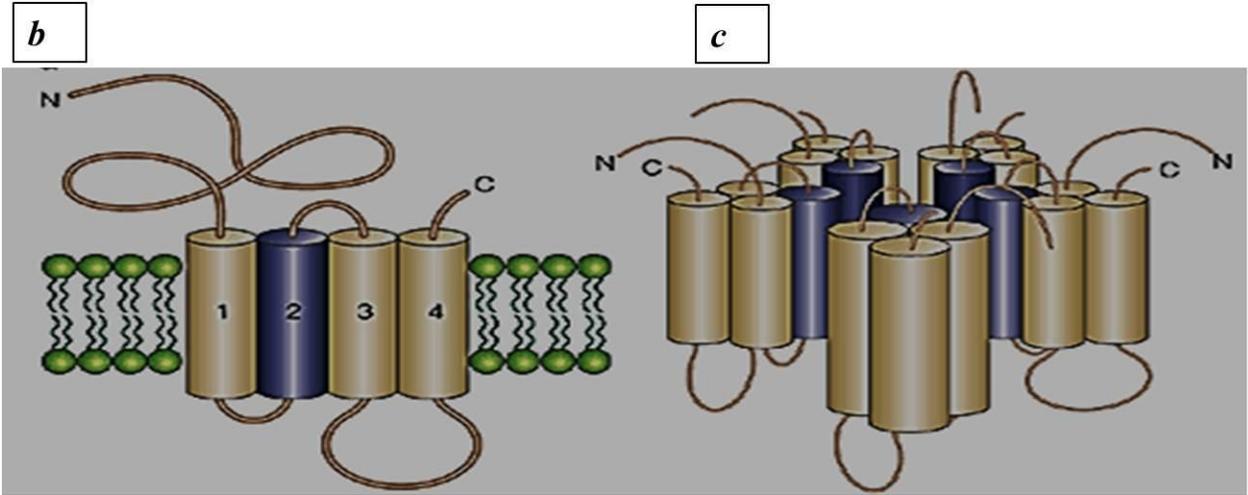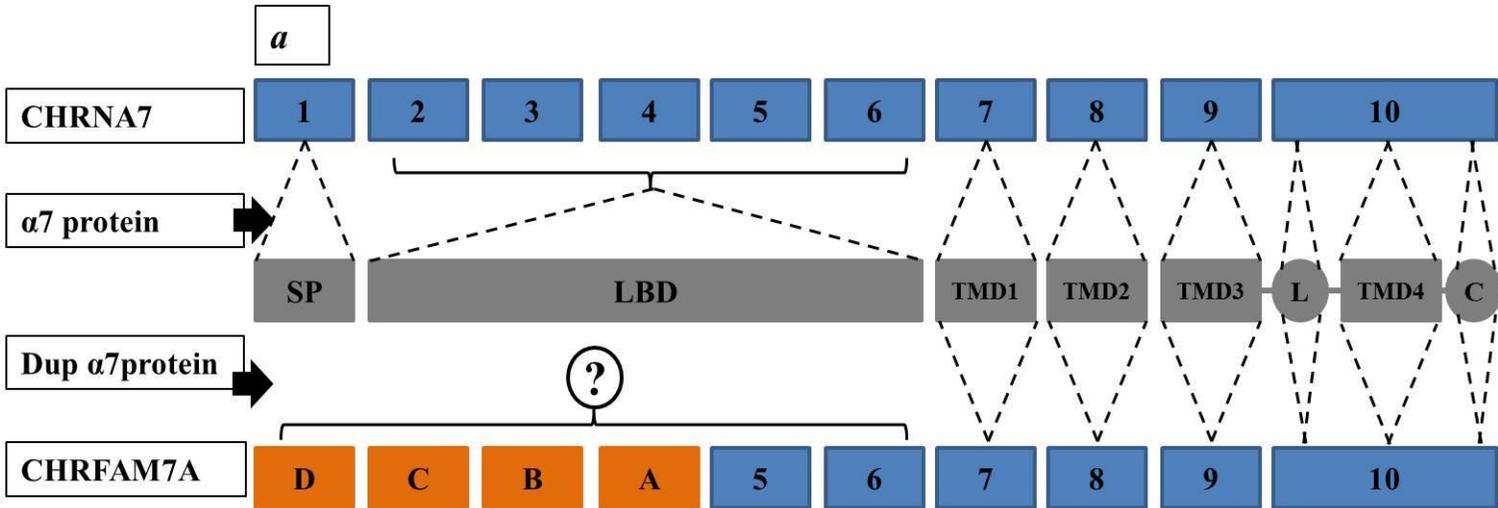
**Figure 1-10: The α-7 nAChRs structure.**

a) CHRNA7 and CHRFAM7A genes are presented with their corresponding protein parts. *Upper panel:* CHRNA7 is composed of ten exons (blue boxes) that codes for 502 amino acids α-7 nAChRs protein. The first exon of CHRNA7 codes for 18 of 22 amino acids that form the signal peptide (SP) part of the protein, exons 2-6 code for the ligand binding part (LBD) and exons 7-10 code for TMDs 1-4, respectively. In addition, exon 10 codes for IC loop (L) and C-terminus (C) parts of the protein. *Middle panel:* main parts of α7 protein. The protein parts are linked with exons coding for them from both genes. *Lower panel:* CHRFAM7A is composed of ten exons, exons D-A (orange boxes) and exons 5-10 are almost the same of CHRNA7 (blue boxes). Note that the dupα7 protein is lacking the SP and LBD but still conserves the TMDs, IC loop, and C-terminus.

b) Schematic representation of the nAChRs (and of the α7-nAChRs) structure. Single subunit parts and their arrangement on the cell membrane are shown.

c) Schematic representation of the nAChRs (and of the α7-nAChRs) structure. Five subunit parts and their arrangement on the cell membrane are shown to represent the pentameric receptor.

Note: for b) and c), N: refers to N-terminus, C: refers to C-terminus and numbers 1-4 referring to TMDs 1-4. TheTMD 2 that is forming the ion channel pore is coloured in blue. The IC loop is shown between TMDs 3 and 4.

Adopted from (Verrall and Hall, 1992, Rakhilin et al., 1999, Moss and Smart, 2001).

In addition, the sequence alignment of proteins of both genes showed a 100% sequence similarities for the part encoded by exons 5-10, which means that these three base differences between the two genes does not affect the amino acids encoded (synonymous nucleotide changes).

The differential expression of both proteins in different human cells and cell lines was reported in many studies. Human tissue macrophages were shown to have higher sensitivity for stimulation by cholinergic agonists than PBMCs (Borovikova et al., 2000). Another study showed that human tissue macrophages can express mRNA and functional proteins on their surface from both genes (CHRNA7 and CHRFAM7A) (Wang et al., 2003). These results were conflicted partially by the findings of a subsequent study showing that PMBCs can express mRNA from the CHRFAM7A only (Severance et al., 2009). Other studies tested for the expression of same (monocytes and PBMCs) and other similar types of cells (human leukemia cell lines) and proved that all these cells express only CHRFAM7A mRNA and protein, although CHRNA7 mRNA was detected from bone marrow cells of the same individuals (Villiger et al., 2002). These results were supported later by the findings of another study on the same cells and cell line (Benfante et al., 2011). However, the latter study showed that human macrophages and monocytes can express low amounts of CHRNA7 mRNA and protein. Taking the latter two study results, it seems that when the cells are in the bone marrow they express both genes, but with different differentiation involved later, with CHRNA7 expression seeming to be lowered according to certain mechanisms. This discrepancy in gene expression within the same cell type can also be attributed to different methods of collection and separation of the cells from the whole blood samples, health

status of the selected individuals (schizophrenia patients or healthy volunteers) or the use of different methods with different detection sensitivity.

When comparing both proteins, it is important to remember the fact that dupα7 lacks the signal peptide forming part. This would subject the protein to subcellular localization and partial retention within the ER (Gault et al., 1998, de Lucas-Cerrillo AM, 2011). Furthermore, it is expected to have a truncated N-terminus that is lacking the important agonist binding site. This can give rise to a difference in the function from that of the α7 protein. This was confirmed when testing PBMCs with dup α7 protein detected the protein but with nonfunctional receptors formed (Villiger et al., 2002).

## 1.4 The regulatory role of CHRFAM7A: negative regulation of CHRNA7

Until recently, the role of the dup α7 protein was only a matter of speculation. Research on CHRNA7 and CHRFAM7A during the last decade revealed what was described as non-functional α7-nAChRs expressed on human PBMCs (Villiger et al., 2002). Further research work revealed that these `receptors` were actually dup α7 protein subunits assembled together. This fact raised the following question: if dup α7 protein can act like other nAChRs and form pentamers (similar to α7-nAChRs), is there is a possibility of production of heteropentameric receptors from α7 and dup α7 proteins?

Subsequent studies aimed at answering this question and to test some of the speculations about dup α7 protein: First, the subcellular localization of dup α7 protein (as a result of lacking the signal peptide part) was tested (de Lucas-Cerrillo AM, 2011). This was concomitant with confirming the expression of

α7 nAChRs on the cell surface. The second finding included the fact that the dup α7 protein can co-assemble with the α7-nAChRs forming heteropentameric receptors with less affinity for α-BTX and less functional activity than homopentameric α7-nAChRs (Benfante et al., 2011, de Lucas-Cerrillo AM, 2011). It is worth noting that similar results were obtained when comparing for dup α7 heteropentamers originating from missing or including 2bp in exon 6 (Araud et al., 2011). Accordingly, this highlights a regulatory role for dup α7 protein on α7-nAChRs (or for CHRFAM7A on CHRNA7). This role was evident when testing for the effect on inflammation using the inflammatory model cells (such as primary human monocytes and macrophages). In such a model, it was shown that by using LPS for stimulating pro-inflammatory cytokine release and start of inflammation, the CHRFAM7A was down regulated at transcriptional and translational levels, while CHRNA7 was transcriptionally upregulated (Benfante et al., 2011, de Lucas-Cerrillo AM, 2011). It is well established that once inflammation has started, in response to LPS or other pro-inflammatory mediators, it will be localized and limited to the affected area via many mechanisms. Among these is the cholinergic anti-inflammatory pathway, which functions principally through activation of the α7-nAChRs (Borovikova et al., 2000, Wang et al., 2003). However, it was not clear until recently how the body controls this effective and possibly damaging anti-inflammatory mechanism. It is evident now that under resting conditions, the CHRFAM7A and its dup α7 protein keep the CHRNA7 and its α7 protein in a minimal expression level thus allowing for inflammation to proceed in response to tissue invasion or injury. When inflammation starts, inflammatory mediators act to suppress CHRFAM7A role and thus allowing for CHRNA7

expression and activation. Among the anti-inflammatory roles of CHRNA7 is the suppression of LPS and other pro-inflammatory mediators. Again, this reverses the CHRFAM7A blocking, paving the way for CHRNA7 down regulatory role, and helping the tissue to go back to the resting stage (Borovikova et al., 2000, Wang et al., 2003, Benfante et al., 2011, de Lucas-Cerrillo AM, 2011) (Figure 1-11).
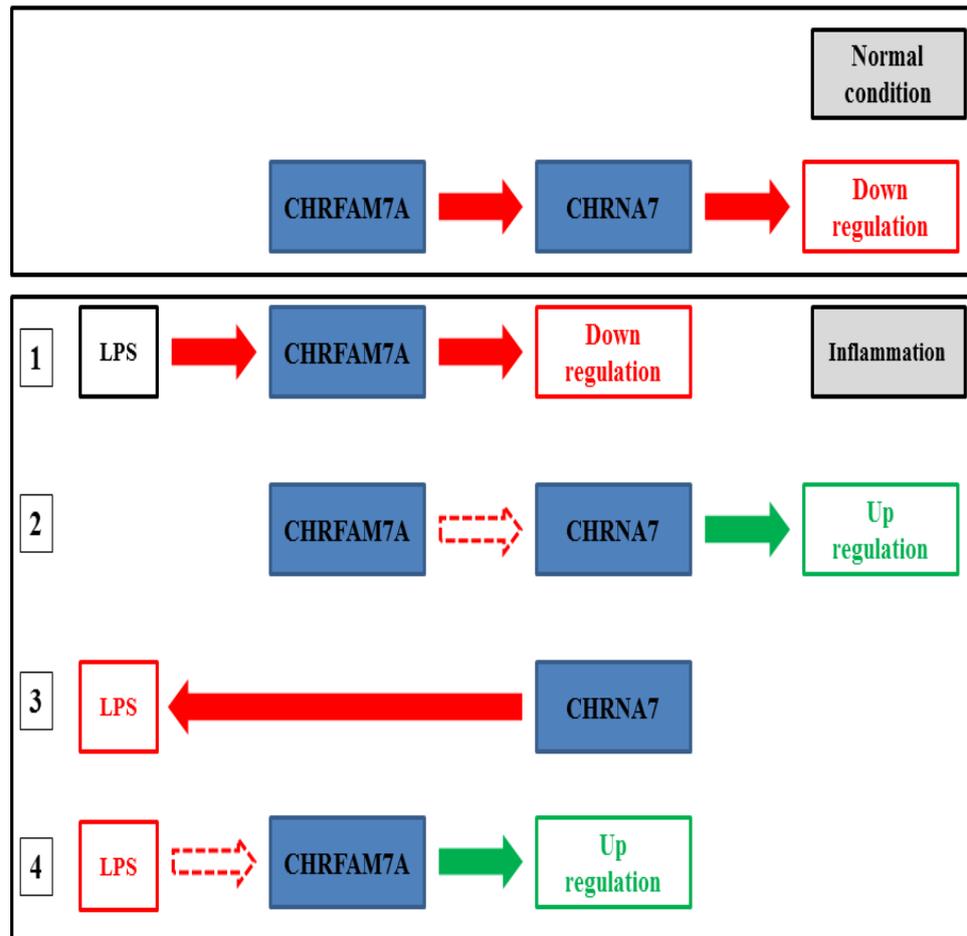
**Figure 1-11: The control mechanism for CHRFMA7A on CHRNA7.**
*Upper panel:* in normal condition, CHRFAM7A is down regulating CHRNA7 expression. *Lower panel:* during inflammation. *1:* LPS and other inflammatory mediators cause CHRFAM7A suppression. *2:* CHRFAM7A effect on CHRNA7 is diminished, causing the latter's upregulation. *3:* CHRNA7 activation leads to down regulating LPS expression. 4: CHRFAM7A is upregulated again, indicating the resolution of inflammatory phase and return to the normal condition (as in the upper panel).

Red arrows and boxes: are showing the down regulation effects. Green arrows and boxes are showing the upregulation effects. Interrupted arrows: diminishing effects.

Adopted from (Borovikova et al., 2000, Tracey, 2002, Wang et al., 2003, Benfante et al., 2011, de Lucas-Cerrillo AM, 2011).

It is worth noting that this relationship between the two genes is different in different human tissues. The relative ratio of mRNA expression of CHRNA7:CHRFAM7A was 5:1 in cerebral cortex while the opposite is true for macrophages (de Lucas-Cerrillo AM, 2011). This might highlight the different roles this receptor is playing in different human tissues, where in the brain it was correlated with neuronal transmission that have an effect on cognition and memory, while in the peripheral tissues, such as tissue macrophages, it seem to have a role in the anti-cholinergic pathway.

## 1.5 The role of α7-nAChR in cholinergic anti-inflammatory pathway:

Human health and disease is balanced by control mechanisms. Human tissues survive repeated insults from infections, inflammatory conditions, injury and cancer. These conditions are dealt with continuously via the immune system. This is fundamentally achieved, in part, through the production of pro-inflammatory mediators (cytokines) that encounter the invading pathogens or particles in a controlled manner (Ulloa and Tracey, 2005). Inflammation is controlled, in part, via the stimulation of the anti-inflammatory pathways, and once the anti-inflammatory pathway is diminished or lost it can give rise to the progression of inflammatory conditions ultimately resulting in human morbidity and mortality (de Jonge and Ulloa, 2007). It is well established that the human body keeps the pro-inflammatory and the anti-inflammatory pathways balanced. These pathways interchange their roles during tissue injury and/or infection helping to keep the tissue damage to a minimum (Tracey, 2002). Disturbances to this balance can lead to anti-inflammatory part

dominance causing infections, sepsis and even cancer. On the other hand, dominance of the inflammatory pathway can precipitate a wide range of chronic inflammatory conditions (Pavlov et al., 2003). This might be, in some cases, more damaging than the original insult. Besides the humoral anti-inflammatory pathway (that forms one of the immune barriers of the human body), another pathway was described that is functioning mainly via the efferent vagus nerve (cholinergic nerve that functions by ACh release). This pathway was termed "The cholinergic anti-inflammatory pathway" (Borovikova et al., 2000) (Figure 1-12).
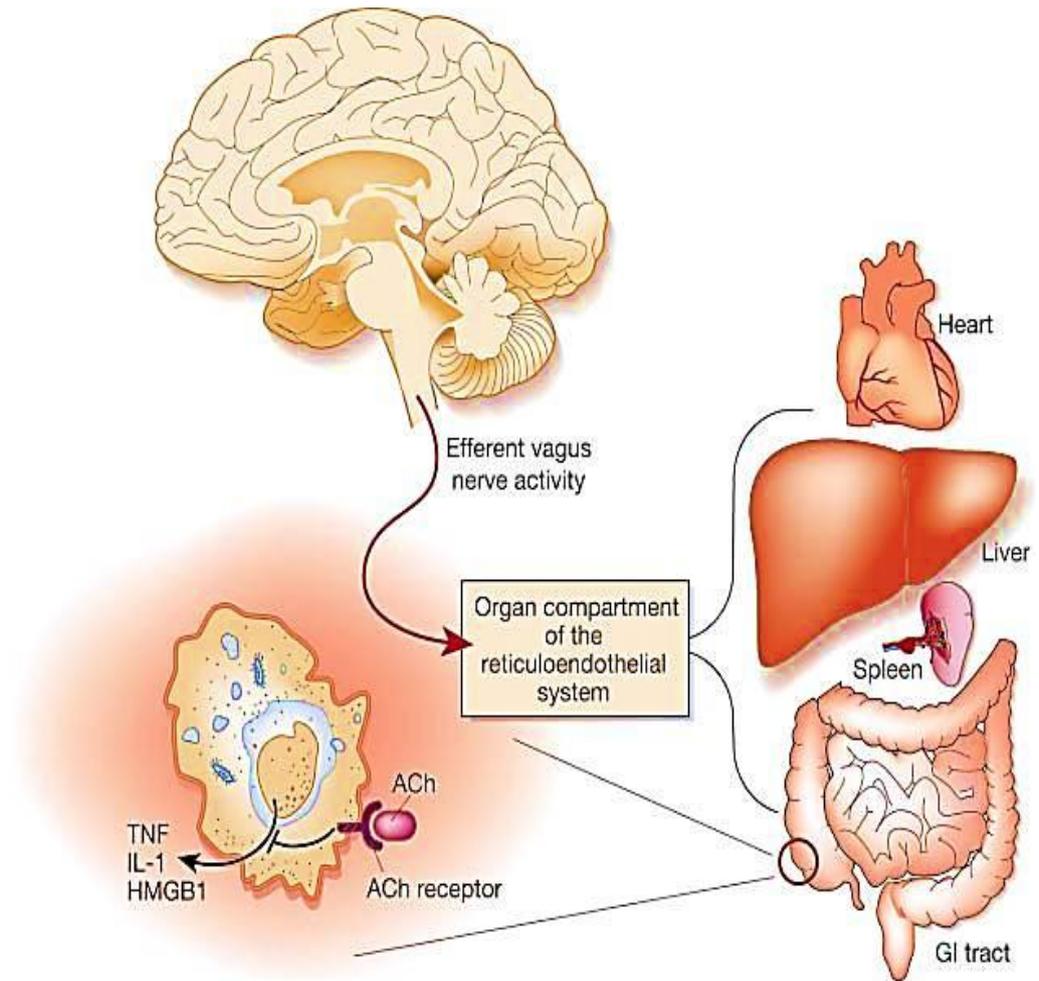
**Figure 1-12: The cholinergic anti-inflammatory pathway.**
Efferent vagus nerve stimulation leads to the release of ACh in organs of the mononuclear phagocyte system (the previously known as the reticuloendothelial system), such as the heart, liver, spleen and gastrointestinal (GI) tract. The released ACh stimulates the α7-nAChRs on the surface of tissue macrophages leading to inhibition of TNF, IL1β and other inflammatory cytokines.

Taken from (Tracey, 2002).

This pathway was shown to be mediated (in experimental animal and human macrophages as models of endotoxemia) through α7-nAChR activation and could suppress the release of pro-inflammatory mediators, such as TNF, IL1β, IL6, IL18 and HMGB1 in response to LPS stimulation (Borovikova et al., 2000, Wang et al., 2003, Wang et al., 2004). Similar results were shown in experimental animal models of acute lung injury (Su et al., 2007) (Figure 1-13).
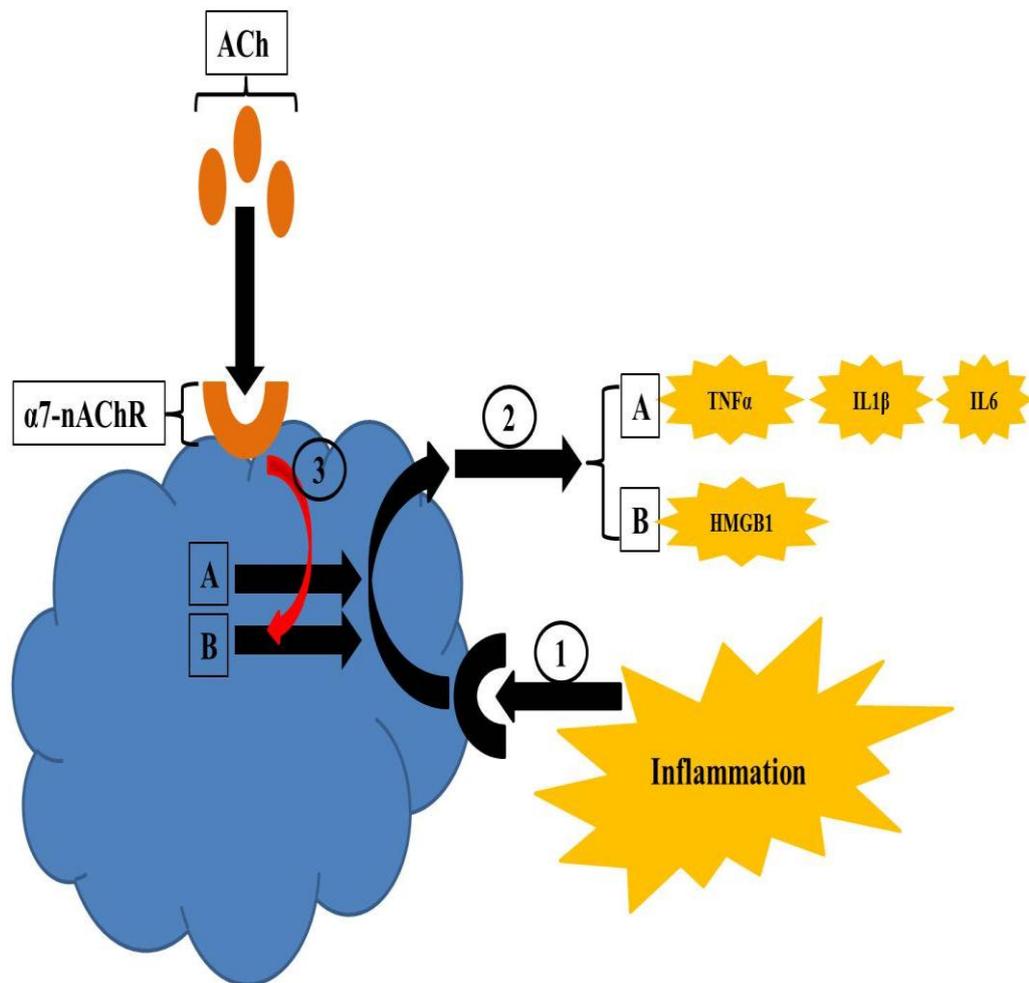
**Figure 1-13: The role of α7-nAChRs in inflammation.**

The inflammation starts with tissue injury or infection (1). This causes the release of pro-inflammatory cytokines from the inflammatory cells (2), such as the tissue macrophages, via translational (A) (for cytokines TNFα, IL1β and IL6) or post-translational mechanisms (B) (for HMGB1). This effect is supposed to be localized to the site of inflammation and with limited inflammatory pro-cytokine release. To ensure this limit is not exceeded, the cholinergic anti-inflammatory pathway starts performing its effect via Ach stimulation of the α7-nAChRs. This pathway will eventually inhibit the cytokine release at both translational and post-translational levels (3).

Black arrows indicate stimulatory effect while red arrows indicate inhibitory effect.

Adapted from (Borovikova et al., 2000, Wang et al., 2003, Wang et al., 2004, Ulloa, 2005).

However, this anti-inflammatory pathway function extends far beyond suppressing pro-inflammatory mediator release into the suppression of endothelial cell activation during inflammation. This suppression will prevent leukocyte migration to the inflamed tissue. The sum of these effects causes further suppression of the inflammatory chemokines used for recruiting other inflammatory cells (Saeed et al., 2005). Furthermore, the spleen was found to be crucial for activation of the cholinergic anti-inflammatory pathway and α7-nAChRs mediate this pathway through suppression of TNF release from the macrophages within the spleen (Huston et al., 2006, Rosas-Ballina et al., 2008). Surprisingly, this effect was found to be mediated via memory T-cell population that can synthesize and release ACh in response to vagus nerve stimulation (Rosas-Ballina et al., 2011). This role for the spleen proved to be crucial for the anti-inflammatory pathway as most of the TNF released during endotoxemia in experimental animals was originated in this organ (Wang et al., 2003, Huston et al., 2006, Olofsson et al., 2012).

Although the cholinergic anti-inflammatory pathway depends on the vagus nerve, the α7-nAChRs are considered as a vital part of the pathway, and using α7-nAChRs antagonists or knockdown techniques was enough to render the pathway ineffective in counteracting inflammation. The term "the inflammatory reflex" was described for the first time by Tracey in 2002. It showed a relay pathway that involves an afferent pathway for informing the central nervous system (CNS) that an injury or infection is starting at certain parts of the human body (Pavlov and Tracey, 2004, Pavlov and Tracey, 2006). This afferent signaling is mediated mainly via the vagus nerve in response to the release of inflammatory cytokines (Olofsson et al., 2012). This represents

the beginning of inflammation. The opposite (cholinergic) anti-inflammatory pathway is represented by the efferent vagus nerve that will stimulate α7-nAChRs and suppress inflammation (Pavlov et al., 2003, Olofsson et al., 2012). This 'reflex' recruits immune cells that can defend the body against invading pathogens and keep an imprint of such a pathogen for possible repeated invasion (Figure 1-14).
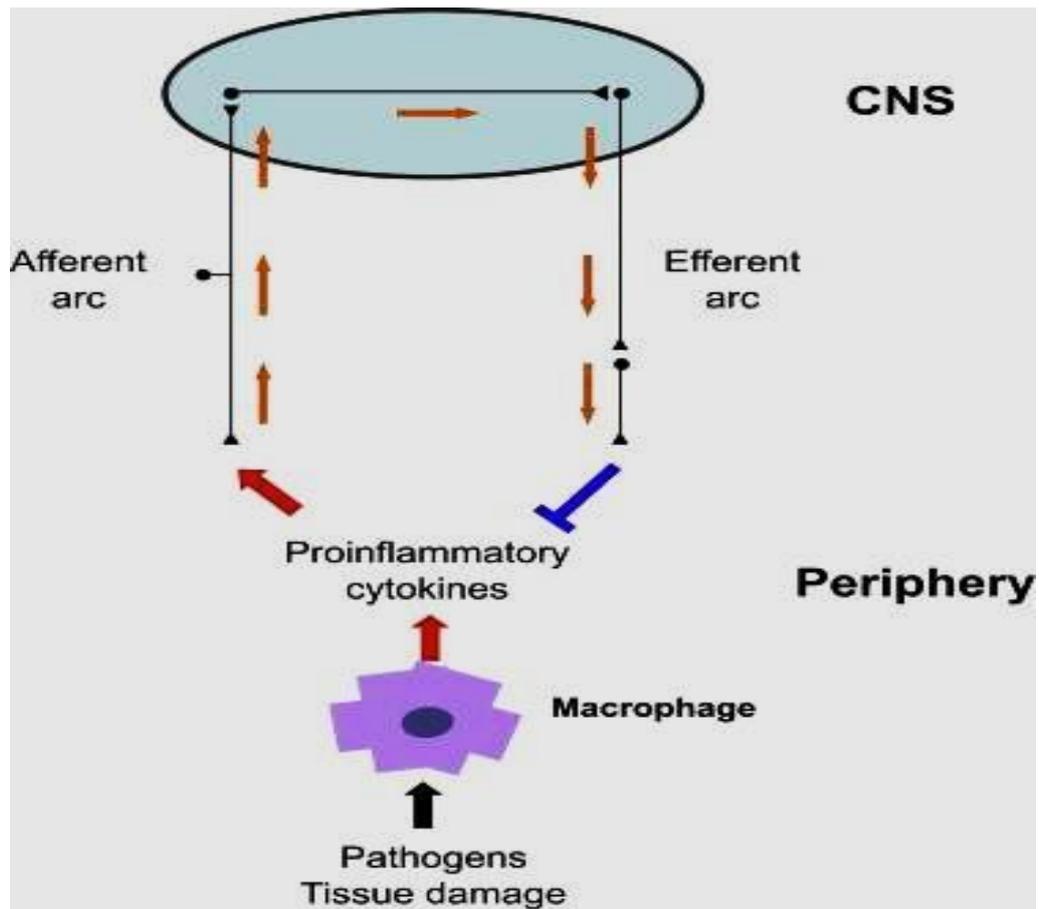
**Figure 1-14: The inflammatory reflex.**

The inflammatory reflex represents a bi-directional communication between the central (CNS) and peripheral (PNS) nervous systems to accommodate inflammation. During inflammation, the tissue macrophages start releasing pro-inflammatory cytokines. This is followed by the stimulation of a number of pathways that inform the CNS (such as the afferent vagus nerve). The CNS response comes through the activation of the neuronal anti-inflammatory pathway represented by the cholinergic anti-inflammatory pathway (via the efferent vagus nerve stimulation of the α7-nAChRs on the surface of tissue macrophages and other inflammatory cells).

Taken from (Tracey, 2002, Rosas-Ballina and Tracey, 2009).

It is worth noting that both the neuronal and immune cells release ACh and thus play a common role in controlling inflammation by delivering the vagus nerve signal to activate the cholinergic pathway. The use of acetylcholine esterase inhibitors helped prevent inflammatory cytokine release and could be used as a potential drug target for treating inflammatory diseases (Ulloa, 2005, Pavlov et al., 2009).

Results from nicotine exposure experiments suggested that chronic nicotine exposure can inactivate α7-nAChRs which may be the basis for defective inflammation suppression in chronic smokers and in COPD (Kawashima et al., 2012). However, stimulation of this pathway might not always be beneficial. For instance, COPD is a chronic inflammatory disease and the chronic exposure of nicotine in these patients suppresses inflammation, but helps precipitation of infection in such patients (de Jonge and Ulloa, 2007). Interestingly, α7-nAChRs were shown to have a critical role in suppressing inflammatory mediators in an inflammatory lung model in humans (Kox et al., 2011).

Although the dup α7 protein, encoded by the CHRFAM7A gene, plays a role in controlling the α7-nAChR expression, its role in the anti-inflammatory pathway requires further exploration (Araud et al., 2011). It was shown that the CHRFAM7A mRNA transcript (but not protein) was down regulated in response to inflammatory cytokine stimulation. This suggested a possible negative regulatory role for this gene upon α7-nAChRs expression during resting condition which can be abolished once inflammation has started and α7-nAChRs expression is required (Benfante et al., 2011, de Lucas-Cerrillo AM, 2011).

## 1.5.1 The role of α7-nAChR agonists in treatment of chronic inflammatory diseases:

More than a decade passed after the first trials of using nicotine for treatment of chronic inflammatory diseases. The earlier attempts for using nicotine treatment was employing nicotine transdermal patches for the treatment of ulcerative colitis (Guslandi, 1999). However, the effect of nicotine toxicity and the less than expected clinical improvement had a role on limiting the use of such treatment. It was shown in many studies that ACh and nicotine can be used as agonists for the α7-nAChRs in suppressing inflammation. However, nicotine proved to have a much more powerful effect than ACh in such sets of experiments (Wang et al., 2003, Wang et al., 2004). Furthermore, α7-nAChR agonists were shown to suppress inflammatory pathways in patients with rheumatoid arthritis (van Maanen et al., 2009).

## 1.5.2 Other regulatory roles for α7-nAChRs:

It was well documented that α7-nAChRs have a major role on respiratory cell growth and subsequent development of lung cancer. This effect is thought to be mediated via the prolonged exposure of α7-nAChRs to nicotine, a condition that is similar to chronic smoking. Prolonged exposure of human cell lines to nicotine can lead to α7-nAChR upregulation (Peng et al., 1999). It is now clearly evident that α7-nAChR controls the regeneration and proliferation of the airway epithelial cells (Maouche et al., 2009). This role of the receptor helps the respiratory epithelial cells to regenerate following wear and tear process or after wound healing. The chronic exposure of α7-nAChRs to nicotine, as the case for COPD patients, was suggested to inactivate these receptors and thus alter the epithelial remodeling. It is worth considering that

nicotine exposure can first cause receptor upregulation but also can cause a reciprocal inactivation due to desensitization effect (Maouche et al., 2009). It was postulated that this effect of nicotine can extend to more than α7-nAChR upregulation and subsequently leads to lung cancer in animal models (Carlisle et al., 2007, Lam et al., 2007, Paleari et al., 2008). It is worth noting that among the mechanisms postulated to explain this upregulation is the repressive effect of nicotine on transcription of factors that suppress α7-nAChRs. Such a repressive effect proved enough to enhance α7-nAChR transcription and subsequent receptor expression (Finlay-Schultz et al., 2011). It is known that nicotine in cigarette smoke is a contributing factor to develop COPD. This effect was explored in GWAS studies to find a possible link between smoking and lung function and COPD.

Many GWAS studies were conducted to explore possible susceptibility loci for COPD development. Among the loci tested, only one locus (CHRNA3/5) on chromosome 15 (15p25) showed correlation with the disease (Zhang et al., 2011). It is important to know that CHRNA7 locus on chromosome 15 (15q13) was not included in any of these studies probably due to the relatively large distance from the CHRNA3/5 locus (more than 40 Mb distance) (Pillai et al., 2009, Wilk et al., 2009, Repapi et al., 2010).

## 1.6   Conclusion:

Inflammation is considered as a control mechanism that helps the body to overcome infections and injuries. This mechanism requires a limiting counter regulatory system that helps to localize the body response and prevent inflammation from changing to a chronic condition. When such a regulatory mechanism (such as the cholinergic anti-inflammatory pathway) is defective, inflammation is expected to be chronic and harmful. The role of α7-nAChR is crucial in the anti-inflammatory pathway to limit the effect of inflammation in certain tissues. The fact that chronic obstructive pulmonary disease is among the examples of chronic diseases might help suggesting a role for α7-nAChR in the disease development or progression. Furthermore, in addition to CHRNA7 coding for this receptor, the expression of the latter is under further control by another gene, the CHRFAM7A (the fusion gene). The fusion gene is down regulated once inflammation starts (via inflammatory mediators) allowing for CHRNA7 up regulation and subsequent α7-nAChR expression. For this purpose I proposed to study the expression of these two genes and their splice variants in airway epithelial cells to help understand possible mechanisms of control between the two genes.

# 2 The characterisation of CHRNA7 (and its duplicate gene) transcripts in a range of human cell lines.

This chapter will discuss the processes of RNA splicing and alternative splicing and the possible effects of splicing elements (enhancers and silencers) on these processes. In addition, the role of alternative transcripts of the two genes controlling the expression of α7-nAChRs (CHRNA7 & CHRFAM7A) is discussed. I present my results on alternative splicing of these two genes in several human cell lines.

## 2.1   NORMAL RNA SPLICING:

Proteins constitute the essential components of living cells and their synthesis involves a multi-step process. For protein synthesis, an intermediate step is required involving passing the genetic code from the gene in the nucleus (DNA) coding for that specific protein via a message (messenger RNA-mRNA). First, the DNA is transcribed into 'pre-mRNA', which includes a replicate of the DNA template with all its exons and introns. The introns are then excised to form the mature 'mRNA', leaving only the exonic sequence to be transferred to the cytoplasm for subsequent translation into the required protein. This process, referred to as 'RNA splicing', takes place in the nucleus (Pierce, 2006, Strachan, 2010). This process requires first the recognition of exons and introns to help a group of proteins (known as the splicing complex or splicing machinery) to assemble on an intron, forming the spliceosome. The next step would be removing that intron and joining its flanking exons (Robberson et al., 1990) (a detailed discussion about splicing will follow shortly).

### 2.1.1   Transcription and splicing:

For protein formation, the translation of mRNA is a fundamental step. However, in addition to transcription there are other essential steps to form mRNA in the nucleus include adding a 5` cap, removing intervening introns and adding the 3`poly (A) tail. These three major steps help to produce a mature and stable mRNA that can ultimately pass to the cytoplasm and be recognized by the ribosomes for subsequent decoding (translation) to form the required protein.

It has been established that RNA transcription and splicing are two different processes but proceed simultaneously. This feature allows for shortening of the time required for splicing the pre-mRNA, thus allowing for only consensus splice sites to be recognized (Dujardin et al., 2012, Kosti et al., 2012). First the production of pre-mRNA from the template DNA starts, followed by RNA capping which includes adding a 7-methylguanosine monophosphate base to the 5` end. This modification enables pre-mRNA recognition for subsequent splicing (Izaurralde et al., 1994) followed by its export to the cytoplasm for translation. The RNA cap is thought to protect the mRNA from exonuclease activity, thus stabilising the mRNA and helping in initiating translation. As transcription proceeds, RNA splicing starts via recognised intron/exon boundaries for subsequent intron removal.

### 2.1.2   Intron/exon definition:

For a successful splicing to proceed, exons require definition for the splicing machinery to recognise their boundaries so enhancing their inclusion in the mRNA. On the other hand, introns require differentiation from exons to be

excluded from mRNA. It is still controversial which part is defined first, exons or introns, but after defining both, splicing starts (Dewey et al., 2006, Hertel, 2008).

There are important parts that constitute an intron, which include the 5` splice site (5`SS) which is formed of GU bases (GT bases in corresponding DNA), the A branch point (18-40 nucleotides upstream of 3`end of the intron), the polypyrimidine tract (PPT) and the 3` splice site (3`SS) (formed of AG bases), (Pierce, 2006, Caporale, 2006 ) (Figure 2-1).
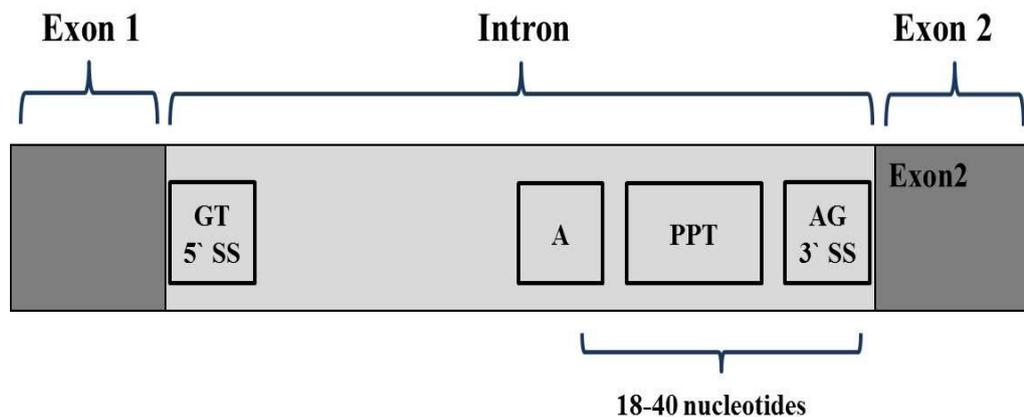


**Figure 2-1: The intron composition (in corresponding DNA sequence).**
The 5` SS (GT) and 3` SS (AG) mark the start and end points of an intron. The Adenine branch point (A) is 18-40 nucleotide bases upstream of the 3` splice site, followed by the polypyrimidine tract (PPT).

The 5`SS and 3` SS sites (with their adjacent sequences) defining an exon are highly conserved and most introns follow the GT-AG rule that defines the start and end points (Smith, 2000   , Strachan, 2010). In addition, the adenine branch point is highly conserved and is required for defining the 3` SS (Kohtz, 1994). Once the splicing machinery identifies these intronic parts, they start to assemble onto an intron. First, the U1 small nuclear ribonuclear protein

(snRNP) binds to the 5`SS. This will be followed by binding of the branch

point protein (BBP) (or splicing factor 1 (SF1)) with the branch point. Another

spliceosome factor termed the U2 auxiliary factor (U2AF) binds to the PPT and

3`SS via its two major parts: 65 and 35 parts, respectively (Matlin et al., 2005,

Strachan, 2010) (Figure 2-2).
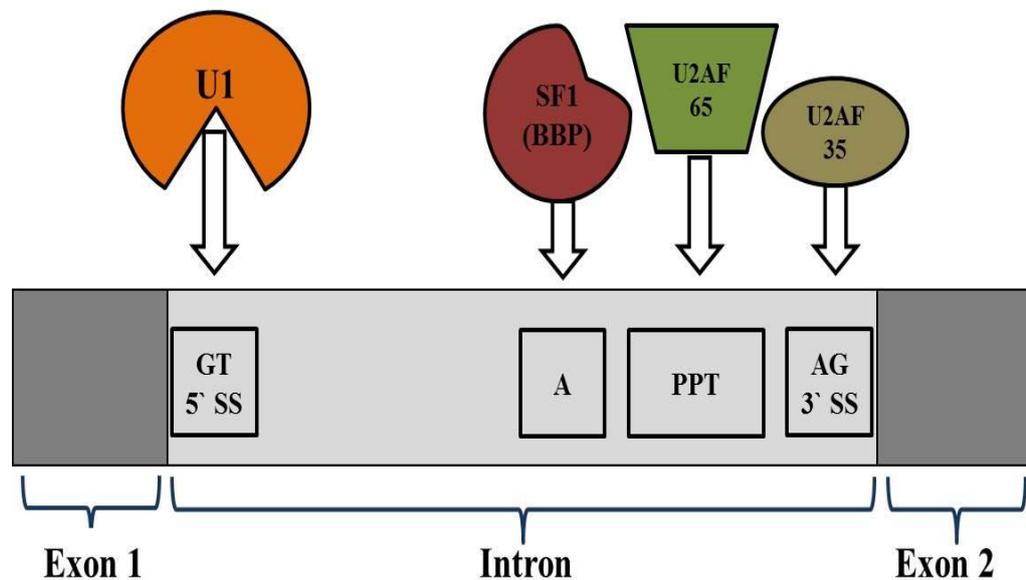


**Figure 2-2: The spliceosome assembly.**
The spliceosome is started by the binding of the spliceosome factors U1, SF1,
U2AF (-65 and -35 parts) to the 5`SS, branch point, PPT and 3`SS,
respectively. Adapted from (Hertel, 2008).

After the spliceosome parts are assembled, the spliceosome machinery starts scanning the sequences downstream of a 3` SS until a consensus 5` SS is identified. This will define an exon in between these two sites, usually within 300 nucleotide bases in length, and the splicing complex removes the identified intron and joins the newly defined exon with the exon upstream of it (Reed and Maniatis, 1986, Robberson et al., 1990).

### 2.1.3    The spliceosome assembly and intron removal:

One of the major factors required for a successful splicing process to take place is the presence of a group of small nuclear ribonuclear RNAs (snRNAs) and proteins (snRNPs), collectively referred to as the 'spliceosome'.

The main functions of spliceosome are first defining exons and introns, and then removing introns to form the mature mRNA (Faustino and Cooper, 2003). The RNA splicing process starts by binding of the first and second snRNPs with the 5`GU and branch point sequences, respectively. The other snRNPs constituting the spliceosome join the reaction, bringing the 5`and 3` splice site together with the branch point close to each other. The complex formed is then cleaved from the exon boundary (at 5` end of the intron) and attaches to the A branch point site, forming a 'Lariat'. The next cleavage will take place at the 3` end of the intron, releasing the intron from the mRNA and allowing the two successive exons to be joined. The process is repeated until all introns are removed and the mature mRNA is formed (Pierce, 2006, Strachan, 2010, Caporale, 2006 ) (Figure 2-3).
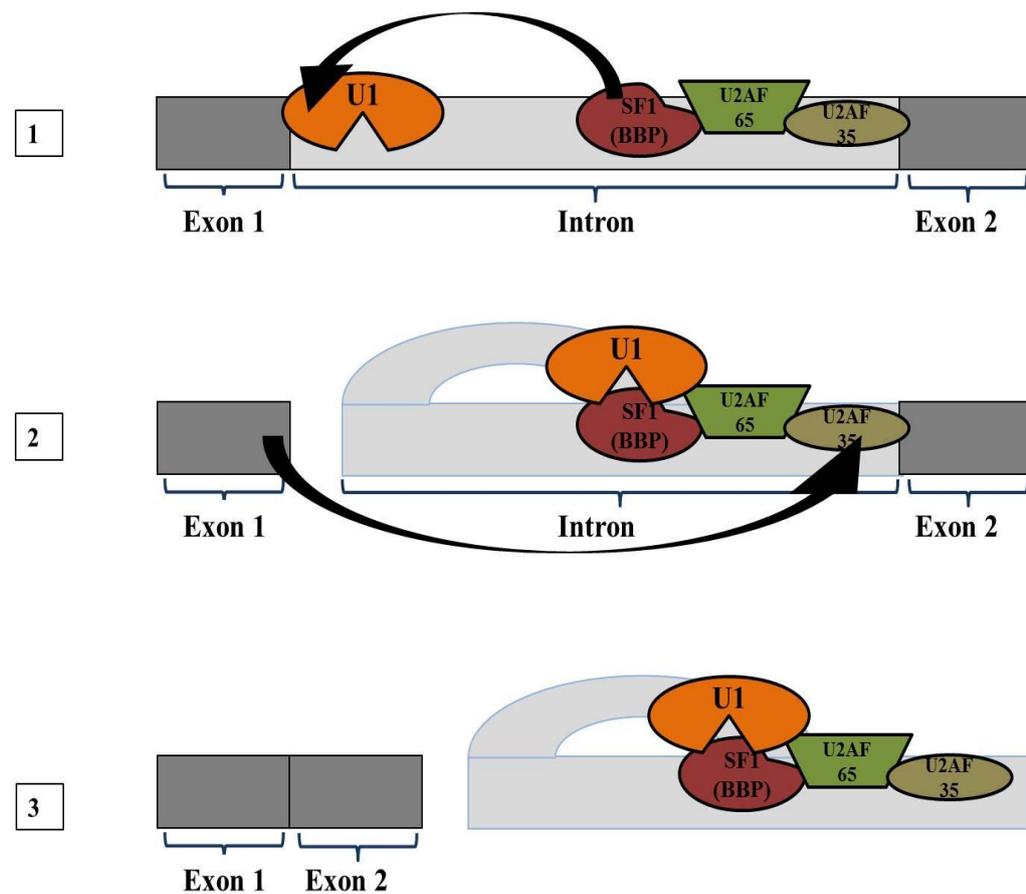
**Figure 2-3: RNA splicing steps.**
*1:* Start of splicing: The Adenine branch point attacks the 5`SS, separating the intronic 5` end from 5` flanking exon. *2:* Lariat formation of the intron. This step starts by joining of the 5`SS to the Adenine branch point to form the intronic lariat. This will be followed by attack of the upstream exon with the 3' end of the intron. *3:* Exon joining and release of the lariat (for subsequent degradation). The spliceosome factors U1, SF1, U2AF (-65 and -35 parts) binding to the 5`SS, branch point, PPT and 3`SS, respectively,are illustrated.

Exons: dark grey boxes, introns: light grey boxes. Adapted from (Caporale, 2006 , Strachan, 2010).

The splicing process requires several factors for successful splicing to be completed, including the spliceosome complex, the splice donor and acceptor sites of the intron, branch point site, Exon Splicing Enhancers (ESEs) and suppressors (ESS), Intron Splicing Enhancers (ISEs) and suppressors (ISS), in addition to ATP (Caporale, 2006 ).

### 2.1.4   Exon sequence motifs controlling splicing:

In addition to the splice sites sequences in the intron, there are a number of exonic sequence motifs that helps control splicing. Different tissues might differ in the use of these sequences, which can alter the results of splicing, giving rise to physiological alternative splicing.

These sequence motifs could be enhancers (ESEs) or suppressors (ESSs). Exon splice enhancers (ESE) are purine-rich conserved sequence motifs that are present in constitutive exons (Black, 2003, Cote J, 2001). They help in exon definition by binding to serine and arginine-rich proteins (SR) that interact with spliceosome parts. The SR proteins are thought to enhance and stabilise the spliceosome complex assembly on the splice sites flanking a specific exon in addition to enhancing propagation on the pre-mRNA. However, changes in ESE sequences might affect exon definition and cause exon skipping. The type of exons affects their inclusion in the subsequent mRNA. Authentic (constitutive) exons are exons that should be included in all mRNA transcripts. These are flanked by constitutive splice acceptor and donor sites, and marked by large number of ESE sequences included within. However, ESE could be found in other exon types, such as pseudo exons. These exons are sequences within the introns and are flanked by weak and/or non-functioning 5` & 3` SS

and branch point (Cote J, 2001). Furthermore, pseudo exons are enriched with silencers (ESS) that interact with suppressor part of the spliceosome and favour their skipping (Black, 2003, Wang, 2004, Zhang, 2004). These ESS act via activating silencer elements (through binding to SR proteins that binds to these silencers), leading to suppressing of the recognition of the flanking splice sites (Wang et al., 2006). Another type of exon is the cassette exons which are not included in all RNA transcripts for a specific gene (Caporale, 2006 ). These might cluster in a close proximity in such a way that only one might be included per each RNA transcript (Smith, 1989, Schmucker, 2000).

Once exon definition starts (by identifying the 3` splice site immediately upstream of the exon) the splicing machinery scans the downstream  sequence for a constitutive 5` splice site. This process is usually active within 300 nucleotide bases (the average exon length), otherwise the search for an alternative (cryptic) 5` splice site starts. If neither 5` splice site was detected, the splicing machinery would skip the 3`splice site-recognised exon and starts scanning downstream sequence for the next constitutive 3` splice site with its immediate downstream exon (Cote J, 2001).

### 2.1.5   Intron sequence motifs controlling splicing:

Intron splice enhancers (ISE) work in a similar pattern to ESE by interacting with SR proteins to enhance the definition of downstream exons (Black, 2003). Similarly, intron splice silencers (ISS) can help in exon skipping. However, the same sequence motif can act as ISE or ISS in different tissues depending on SR

protein recognition of such a motif as an enhancer or suppressor, thus contributing to either exon definition or skipping.

In general, enhancers are thought to enhance exon recognition or oppose a silencer effect allowing constitutive exons to be included in mature mRNA. A change in part or all of enhancers and/or suppressors might affect the selection of flanking exons (Fairbrother and Chasin, 2000) and change the mRNA constitution.

### 2.1.6   RNA polyadenylation:

Following transcription, the RNA polymerase II enzyme scans the pre-mRNA for the polyadenylation signal (usually within less than 30 nucleotides downstream of the stop codon (Strachan, 2010). Once detected, RNA polymerase II starts to cut the pre-mRNA and another enzyme (poly (A) polymerase) starts adding a long sequence of adenines to the 3` end (poly (A) tail). This modification helps in mRNA transport to the cytoplasm and assembly of the ribosome for translation initiation (Pierce, 2006, Strachan, 2010, Jung et al., 2006).

In summary, for most genes the expression in human tissues passes through transcription, 5` capping, splicing, polyadenylation, export to the cytoplasm, and then either into translation or degradation depending on the correct sequence carried in the mRNA.

### 2.1.7   Nonsense-mediated mRNA Decay (NMD):

The accurate translation of mRNA to protein is important for synthesising the correct protein required for a specific function. However, if different transcripts

of mRNA of one gene were all allowed to be translated, different proteins could be produced with different or antagonising functions (Chang, 2007).

Normally, there are a group of mechanisms termed 'mRNA surveillance'. These mechanisms function to prevent, or at least lessen, the effect of alternative mRNA transcripts (splice variants) that might interfere with proper protein translation. One of these mechanisms is Nonsense-mediated mRNA Decay (NMD). This mechanism involves the detection of mRNA with premature stop codons, and subsequent degradation before translation. This is thought to prevent the formation of different (short or truncated) proteins that might possess a different function from the original protein (Chang, 2007). The NMD process depends on detecting the Exon Junction Complex (EJC) (formed of a group of proteins working together) downstream of a detected premature stop codon. Once activated, the NMD activate enzymes that degrade mRNA. The premature stop codon is often found in an exon upstream of the terminal exon that harbours the original stop codon.

In a normal transcript, there are no exon-exon junctions detected downstream of the stop codon, and NMD is not stimulated. Instead, the EJC will enhance translation (Nott et al., 2004). However, in transcripts with premature stop codon, at least one exon-exon junction is detected by the EJC. This leads to stimulation of NMD and subsequent degradation of mRNA (Figure 2-4).
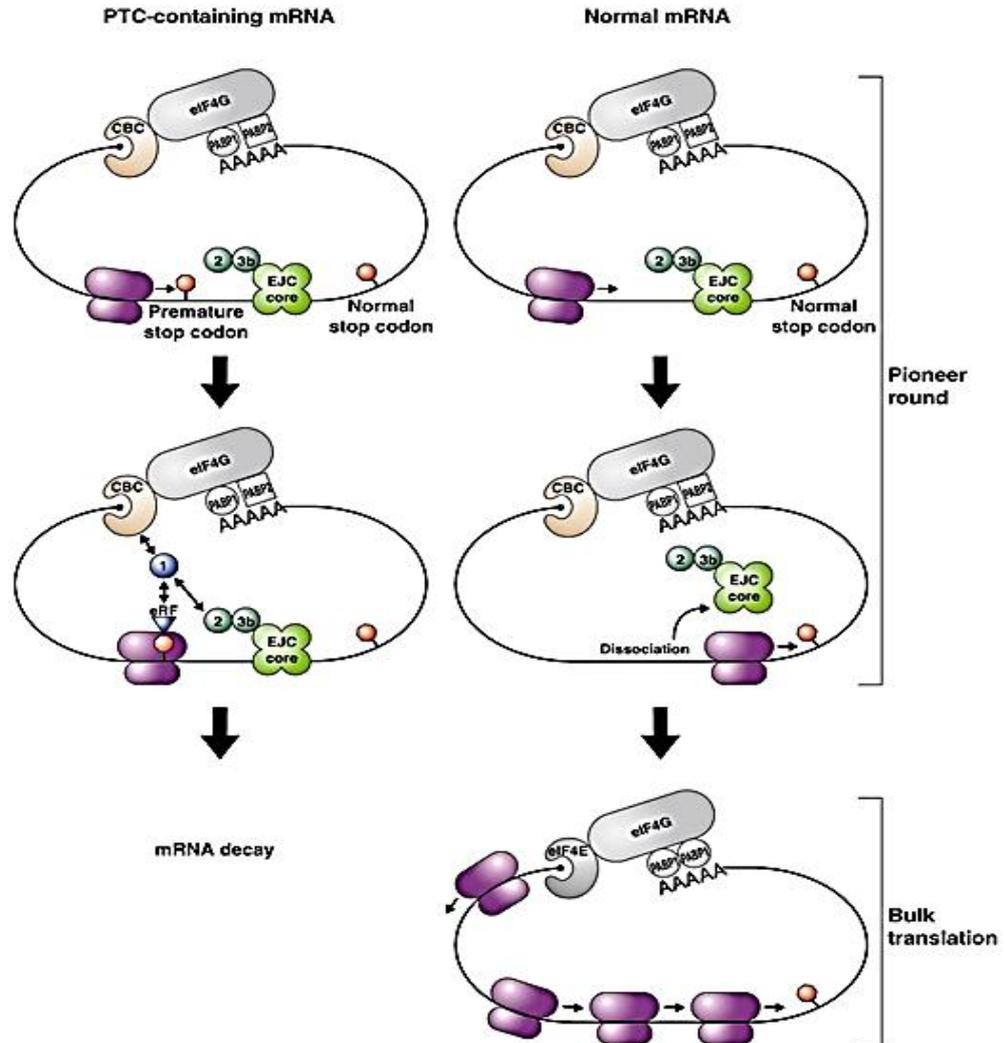
**Figure 2-4: Nonsense-Mediated Decay (NMD).**
The nonsense-mediated mRNA decay (NMD) occurs because of premature stop codon (PTC) recognition during the first round of translation. *Left panel:* When exon junction complex (EJC) represented by (2 & 3b) is detected downstream of PTC, NMD is allowed to degrade RNA. In contrast, a normal transcript (*right panels*) avoids NMD because the last EJC is upstream of the stop codon. CBC: cap-binding protein, eRF: eukaryote release factors.
Taken from (Chang, 2007).

The process of alternative splicing might end up with mRNA that codes for a protein with different structure and physiological function (or sometimes without a specific function).

### 2.1.8  Tissue-Specific Splicing:

Many genes express different mRNA transcripts in different tissues. Alternative splicing was reported to affect more than two thirds of all human genes (Black, 2003) and of these up to 30% was shown to be tissue specific splicing (Xu et al., 2002). House-keeping genes are expressed in all tissues due to an essential requirement for their function. However, other genes might be expressed variably in different tissues, depending on the splicing factors specificities or the particular stage of differentiation of these cells (Strachan, 2010, Grosso, 2008). For a particular gene, this might cause more transcripts in one tissue than other tissues due to alternative splicing. In a recent study, it was shown that alternative promoters can cause recognition of alternative transcription initiation sites, and thus contribute to more than half of the alternative splicing in the human genome (Kimura et al., 2006). In this study, the most common tissues affected by alternative promoters were testis and brain tissue.

## 2.2  Alternative RNA Splicing:

More than three decades ago, it was estimated that alternative splicing of mRNA affected only 5% of the whole human genome. Now with the advances achieved in bioinformatics and genome-wide analysis it has been shown that alternative splicing may be as high as up to 90% in the human genome

(Dujardin et al., 2012, Matlin et al., 2005). This question was first brought into light by the fact that the reported transcripts from the human genome outnumbered the reported human genes by almost five times (Modrek and Lee, 2002, Piva et al., 2012), highlighting that there might be more than one transcript for each gene. The process of alternative splicing was investigated throughout the human genome to uncover the mechanisms and pathways of some diseases. However, alternative splicing can be a part of a regulatory pathway (Pierce, 2006, Caporale, 2006 ). This is best explained in *Drosophila melanogaster* (fruit flies), where alternative splicing plays a major role in sex differentiation (Black, 2003). Female flies have the sex lethal protein (Sxl) which affects the splicing of two genes: transformer (Tra) and Male-specific lethal 2 (Msl2) genes. On the Tra gene, Sxl in female flies binds to a 3`SS downstream of exon 3, enhancing splicing of exon 3 and subsequent protein translation (the reverse is true for male flies, where Sxl is absent causing an inactive form of the protein to be translated). The opposite effect was noticed in the Msl2 gene, where the presence of Sxl protein in female flies causes intron 1 retention and a subsequent inactive form of the protein to be translated, while the male flies that lack Sxl will have an active form of the protein.

There are many types of alternative splicing, including exon skipping, mutually exclusive cassette exons, use of alternative 5` or 3` splice sites, use of alternative promoter or poly (A) tail or intron retention (Caporale, 2006 , Black, 2003) (Figure 2-5).
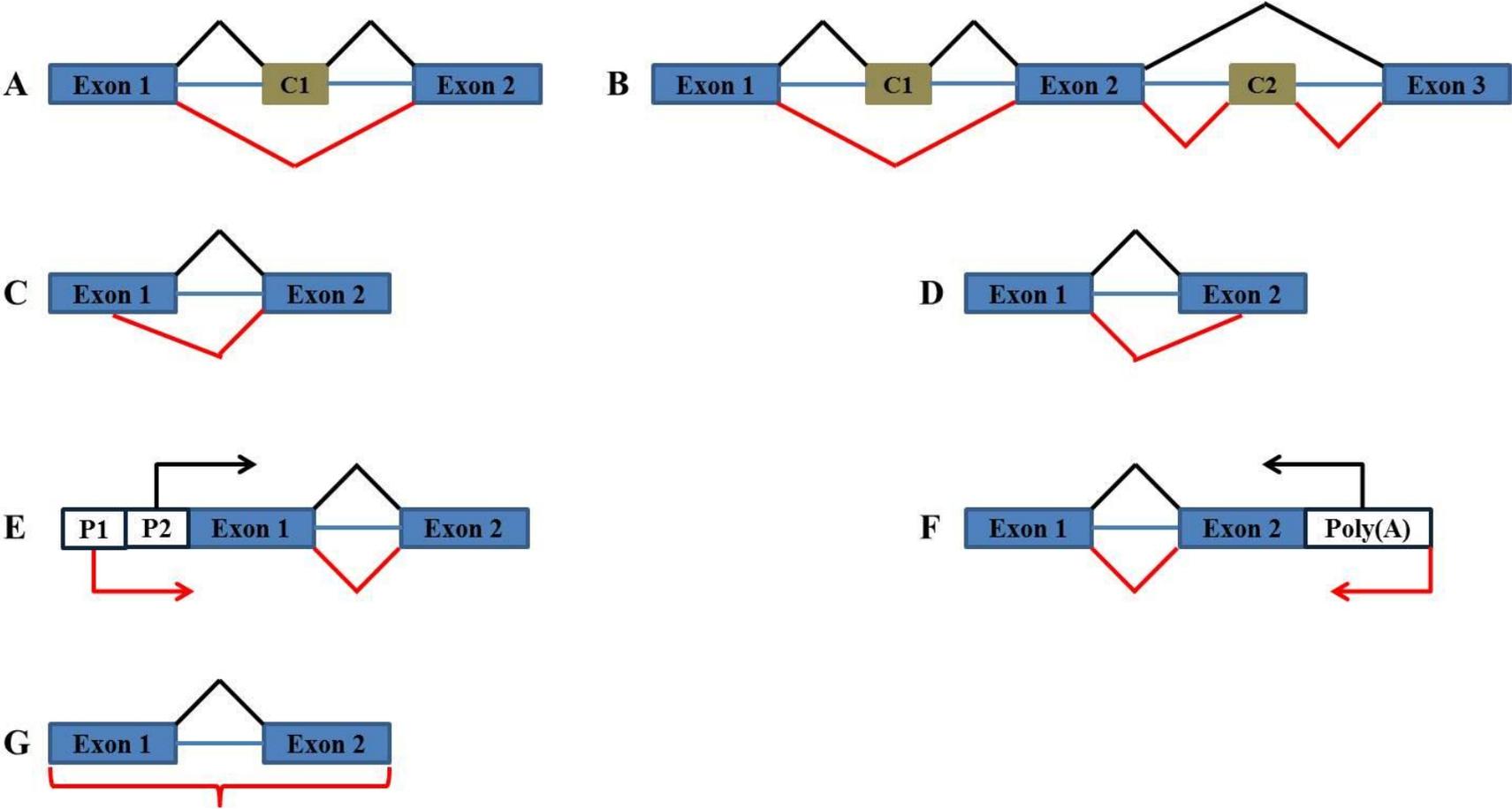
**Figure 2-5: Alternative RNA splicing forms.**

*A:* Exon skipping of cassette exon (C1), *B:* inclusion of mutually exclusive cassette-exons (C1 or 2), *C:* alternative 5` splice site, *D:* alternative 3` splice site, *E:* alternative promoter (P1 or P2) region, *F:* alternative poly (A) tail, *G:* intron inclusion. Blue boxes: exons, blue lines: introns, black lines: normal splicing pathway, red lines: alternative splicing pathway.

The high percentage of alternative splicing could be justified by the fact that the spliceosome assembly is dependent on three short sequence motifs at the splice sites (5` and 3`) and the adenine branch point. Although these three sites are highly conserved (Strachan, 2010) they still represent a dot in a sea of nucleotides that constitute introns, and can easily be mixed with similar or even sometimes identical sequences at cryptic (decoy) splice sites (Cote J, 2001) that might be contained in cassette exons or pseudo exons. Alternative splicing could affect human genes in stress, but sometimes they could lead to diseases (pathological alternative splicing) (Kelemen et al., 2012). Many examples for pathological alternative splicing exist in human diseases, several of which were based on finding the inhibitory effects of alternative transcripts and their translated truncated proteins on the normal transcripts and their proteins (Faustino and Cooper, 2003). Other factors might play a role in alternative splicing of some genes, such as the presence of some SR proteins in abundance in some tissues and not in other tissues. This might affect the ESE that specifically binds to such (target) SR protein, and thus enhancing the inclusion of exons containing such a specific ESE in tissues with abundant (target) SR protein than other tissues (Caporale, 2006 ).

## 2.3   Alternative splice variants of CHRNA7 in human and animal tissues:

Like many other genes, several splice variants for CHRNA7 have been reported in human and animal neuronal tissues. Garcia-Guzman et al showed for the first time an alternative splicing of CHRNA7 in chromaffin cells of

bovine adrenal medulla (García-Guzmán et al., 1995). The mRNA and protein

for the full-length gene were detected, and the assembly of α7 subunits into

functional receptors was demonstrated. The splice variant included a transcript

missing the 87 bp of exon 8 (Figure 2-6, E). This will cause the second

transmembrane domain to be omitted, but still maintains the reading frame.

When this transcript was injected in Xenopus oocytes, no functional ion

channels were detected. However, when tested for the effect on α7 subunit

assembly, this variant proved to have an inhibitory effect when co-expressed

together with the full-length copy of the gene. It was thought that the exon

skipping was related to the short length of the exon and weak flanking splice

sites. This splice variant was suggested to be under the control of species-

specific splicing factors, as it could not be detected in rat brain and adrenal

medulla tissues (tissues that contain α7-nAChR). A few years later, Gault et al

showed for the first time the sequence of the full length (CHRNA7) gene for

the human α-7 nAChR subunit (Gault et al., 1998). They showed in this study

that, in addition to the full-length transcript, there are five splice variants for

the gene (from human brain tissue) that ranged between deletions of exons 3, 4,

3 & 4, 4 & 5 and 3-5. Of these variants, only one (missing exon 3) maintained

the open reading frame of the α-7 coding sequence while the other transcripts

contained premature stop codons (Figure 2-7, A-C, E-G). Later, Saragoza and

colleagues showed that a novel exon exists in intron 9 of CHRNA7 of mouse

brain tissue and termed it exon 9b (Figure 2-6, G) (Saragoza et al., 2003). This

exon included 218 bp inserted from intron 9 sequence, harbouring three in-

frame termination codons leading to NMD. Interestingly, this exon was found

to lie within a 520 bp region, which is conserved in the corresponding part of

the human gene. Similar to transcripts missing exon 8, transcripts with exon 9b insertion were shown to have an inhibitory effect when co-expressed with the full-length copy of the gene. However, this variant was shown to be influenced by tissue-specific splicing factors in mice, where it was detected in mouse brain and dorsal root ganglion but not in retinal tissues. Following that, the first discovery of exon 4a in rat neuronal tissue was made by Severance et al (Severance et al., 2004). This included a novel exon of 87 bp that would preserve the reading frame of the transcript in rat tissue. Several other splice variants for CHRNA7 were also detected in this study. These variants included deletions of exons 3-6 (completely or partially), all of which contained premature stop codons that affected the reading frame and resulted in a protein missing parts that encode for agonist binding sites of the α7 receptor (Figure 2-6, B-D, F). When compared to full-length transcripts, the exon 4a variants showed more sensitivity to bind Ach, and a higher affinity to bind αBTX; however, their desensitisation rate was slower (Severance et al., 2004).

Following their work in rat tissue, Severance and co-workers detected a novel exon between exons 4 and 5 (termed exon 4a) of the human gene (Severance and Yolken, 2008). This was larger than the corresponding exon in rat neuronal tissue detected previously (124-127 bp size). Another six splice variants were detected in this study, two of which were reported earlier by Gault et al (Gault et al., 1998) (missing exon 4 or exons 4 & 5). However, the rest were new variants including, in addition to inserted exon 4a, partially deleted exon 4, exon 5 deletion or a CAG (singlet/doublet) polymorphism at the exon 4-inserted sequence splice junction (Figure 2-7, C-D, F, H-K). The new variant (with the added exon 4a) resulted in addition of 41-42 amino acids that

maintained the open reading frame. Surprisingly, the studies on human CHRNA7 lacked the functional characterisation of the resulting variant, and its effect on translation, assembly, and folding of α7 protein is still unknown.
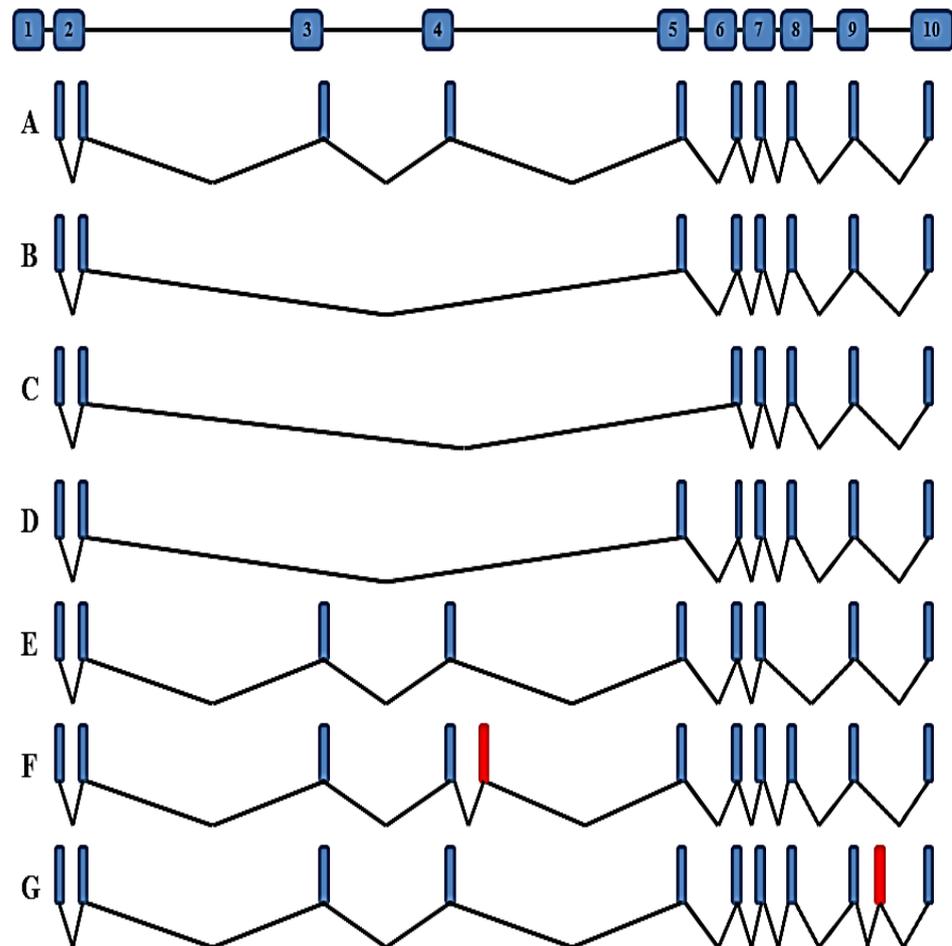


**Figure 2-6: Alternative splicing detected for CHRNA7 in experimental animal neuronal tissues.**
Top panel: exon numbers, *A:* full length CHRNA7 mRNA, *B:* missing exons 3&4, *C:* missing exons 3,4&5, *D:* missing exons 3&4 and splice junction of exons 5-6, *E:* missing exon 8, *F:* exon 4a, *G:* exon 9b. Blue boxes: constitutive exons, red boxes: cassette exons.
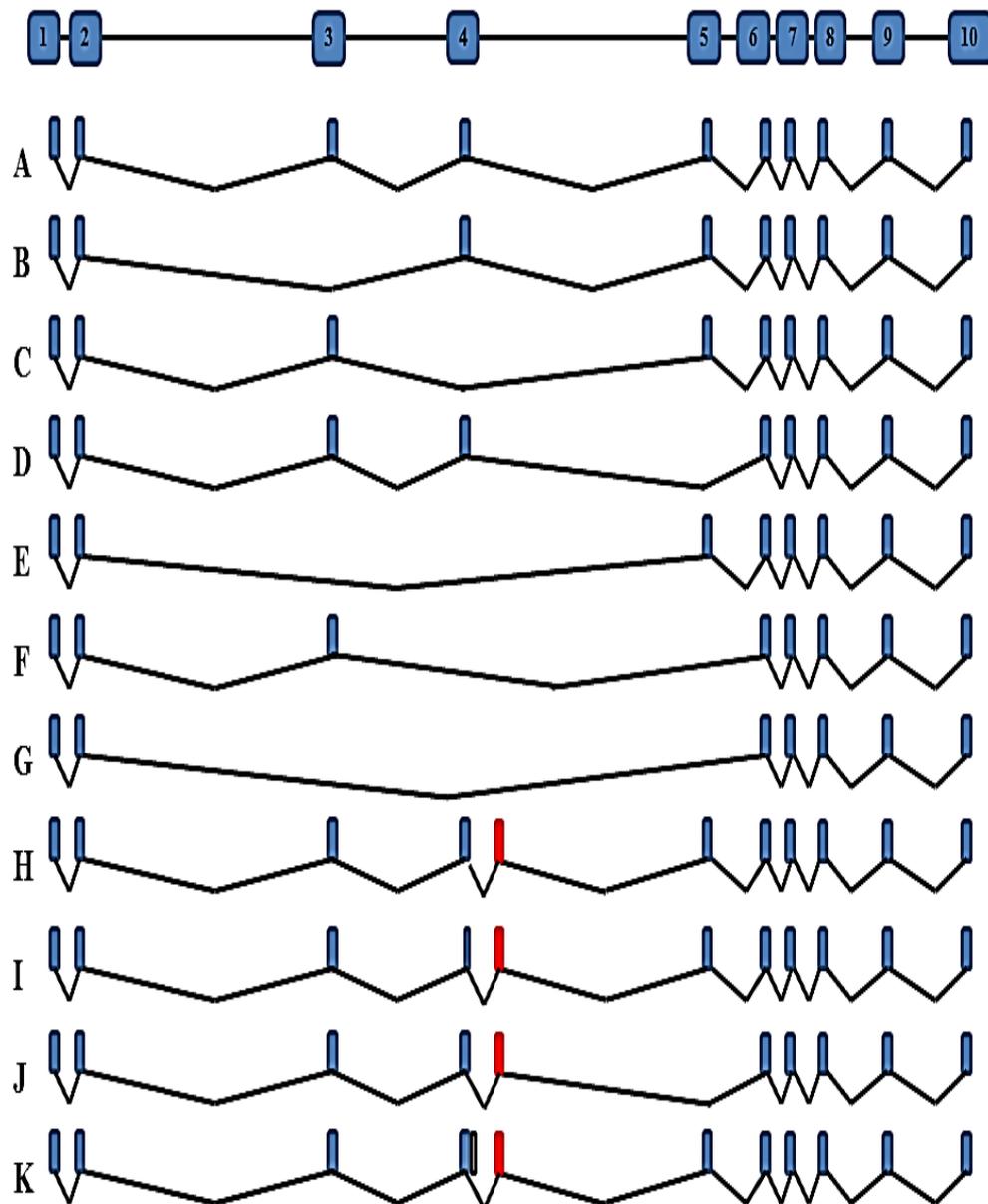
**Figure 2-7: Alternative splicing for CHRNA7 in human brain tissue.**
Top panel: exon numbers, *A:* full length CHRNA7 mRNA, *B:* missing exon 3, *C:* missing exon 4, *D:* missing exon 5, *E:* missing exons 3&4, *F:* missing exons 4&5. *G:* missing exons 3-5, *H-K:* Exon 4a transcripts (*I:* with partial exon 4 deletion, *J:* with exon 5 deletion, *K:* with CAG polymorphism at exon 4 inserted sequence splice junction). Blue boxes: constitutive exons, red boxes: cassette exons.

## 2.4  Alternative splice variants of CHRFAM7A:

Generally, fewer alternative transcripts for CHRFAM7A were reported. These include exon E (between exons A & B), exon D' or missing exon B (Riley, 2002, Gault et al., 2003). However, none of these studies explored the expression of the CHRFAM7A protein or its function (Gault et al., 1998, Riley, 2002) (Figure 2-8).
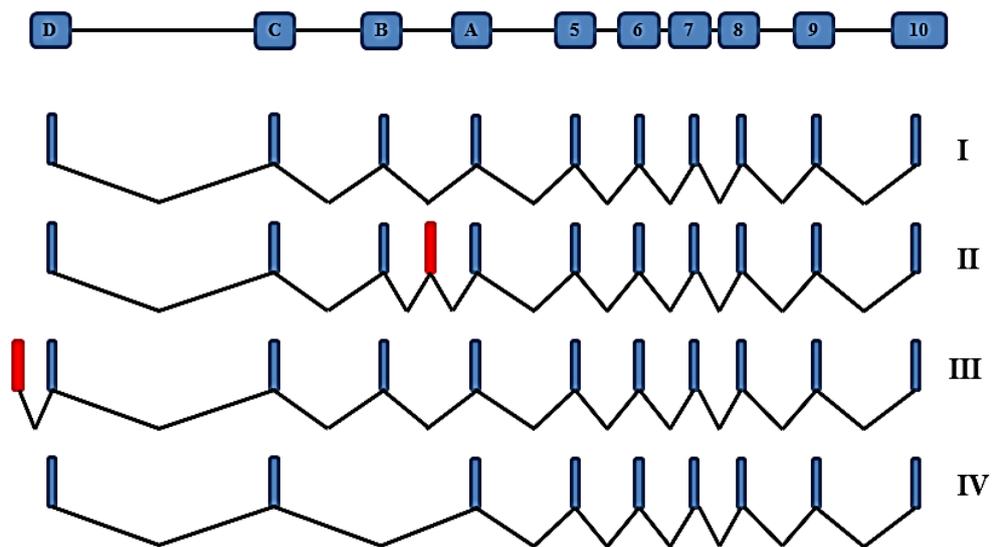


**Figure 2-8: CHRFAMA alternative transcripts.**
Top panel: exon numbers, *I:* full length CHRFAM7A mRNA, *II:* including exon E between exons B & A, *III:* including exon D' upstream of exon D, *IV:* missing exon B. Blue boxes: constitutive exons, red boxes: cassette exons.

## 2.5  Aims:

This chapter had two aims. The first was to characterise the CHRNA7 and CHRFAM7A transcripts in an immortalised airway epithelial cell line (A549) and other cell lines (BEAS2B, PBMC and BE (2)-c). The second aim was to quantify the main transcripts and to predict the protein structure.

## 2.6   Methods:

### 2.6.1   Cell culture:

The A549 cell line (human lung adenocarcinoma epithelial cells) was used as an *in vitro* model for type II pulmonary epithelial cells. Other cell lines were also used for testing alternative transcripts of CHRNA7 and CHRFAM7A (Table 2-1).

**Table 2-1: Human cell lines used for alternative transcript test.**

| Cell line | Description |
|---|---|
| A549 | Alveolar adenocarcinoma cell line |
| BEAS2B | Bronchial epithelial cell line |
| PBMC | Peripheral blood mononuclear cells |
| BE(2)-c | Neuroblastoma cell line –derived from bone marrow metastasis |

The culture medium used for maintaining A549 cells was composed of Dulbecco's Modified Eagle's Medium (DMEM) (SIGMA-ALDRICH, UK), 10% FBS (foetal bovine serum) (SIGMA-ALDRICH, UK), 2mM L-glutamine (SIGMA-ALDRICH, UK), 100U/ml penicillin, 100µg/ml streptomycin (Invitrogen, UK) and 2.5µg/ml amphotericin B (SIGMA-ALDRICH, UK) in 75cm$^2$ flasks. Regarding BE (2)-c cells, the culture medium used was Eagle's minimum essential medium (EMEM) (SIGMA-ALDRICH, UK), with 15% FBS (foetal bovine serum) (SIGMA-ALDRICH, UK), 2mM L-glutamine (SIGMA-ALDRICH, UK), 100U/ml penicillin, and 100µg/ml streptomycin (Invitrogen, UK) and 2.5µg/ml amphotericin B (SIGMA-ALDRICH, UK) and 5% NEAA (non-essential amino acids) in 75cm$^2$ flasks. However, RNA for BEAS2B and PBMC was kindly provided from other PhD students.

For maintenance of A549 or BE (2)-c cells, they were incubated at 37° and 5% $CO_2$ humidity for 24 hrs. For cell splitting, the cell cultures were first examined under the light microscope and were subcultured once they reached ~80% confluency (about $3x10^6$ cells). The medium was removed, followed by washing with PBS (phosphate buffered saline) (Oxoid, England), and cells were dissociated from the flasks using trypsin-EDTA (SIGMA-ALDRICH, UK), followed by incubation at 37° for 5 minutes. Once the cells were successfully detached, DMEM was added and mixture split into two different 75 $cm^2$ flasks and incubated at 37° for 72 hrs.

## 2.6.2 RNA Extraction:

Total RNA was isolated from human cell lines of interest (Table 2-1). When cell cultures were > 80 % confluent, they were used for RNA extraction. First, the cells were harvested by washing the cultures with PBS (5 ml) prior to cell dissociation using trypsin-EDTA. The cell suspension was centrifuged for 5 minutes at 1000 rpm (300g) to pellet the cells. For RNA preparation, RNeasy Mini kit (QIAGEN) was used according to the manufacturer's guidelines. In addition, RNase-free DNase set (QIAGEN) was used according to manufacturer's guidelines. This step was used for removing remnants of genomic DNA that may interfere with downstream steps. The DNA yield was then measured using a Nanodrop spectrophotometer.

## 2.6.3 cDNA Generation:

The cDNA was generated from 2 µg of RNA using 500 ng of Oligomer dT (Stratagene, Germany) in a final reaction volume of 15.7 µl. The mixture was

then incubated for 5 minutes at 65° and then for 10 minutes at room temperature. The following reagents from the Affinity Script kit (Stratagene, Germany) were then added to each tube: 1x buffer, 4 mM dNTP, 1 μl of the reverse-transcriptase multi-temperature enzyme and 1 U/ μl of RNase Block. The reactions were incubated for 10 minutes at 25°, 60 minutes at 42°, 15 minutes at 70° and at 10° hold. In addition, RT-negative samples were used (using 1.5 μl of RNase-free water instead of reverse transcriptase enzyme and RNase block) to enable the detection of products derived from contaminating genomic or plasmid DNA rather than from generated cDNA.

### 2.6.4   PCR:

CHRNA7 is located on chromosome 15, proximal to CHRFAM7A (partially duplicated CHRNA7) gene. To detect possible expression of splice variants for CHRNA7 by A549 cells, several primers were designed to amplify different parts of the mRNA. Most of these were designed in a way to avoid detecting the common exons 5-10 expressed by both genes. Table 2-2 shows the primer sets used to detect CHRNA7 variants. However, for CHRFAM7A, the primers were designed to detect for two main parts of the gene mRNA: FAM7A exons (D to A) and the duplicated CHRNA7 exons (5-10). Table 2-3 shows the primer sets used to detect the main splice variants of CHRFAM7A.

**Table 2-2: Nucleotide sequences of the primers used for CHRNA7 variants detection.**

| Primer ID | Primer sequence (5`→3`) | Position |
|-----------|------------------------|----------|
| A7E1_S | GACTCAACATGCGCTGCTC | Exon 1 |
| A7E2_S | GCCAATGACTCGCAACCACTC | Exon 2 |
| A7E4_S | AGATGGCCAGATTTGGAAAC | Exon 4 |
| A7E4a_S | TTATGACCAACAGCAGGAAGTG | Exon 4a |
| A7E4a_AS | CACAACAGGGAACATAGACGA | Exon 4a |
| A7E6_AS | CCAGCGTACATCGATGTAGCA | Exon 6 |
| A7E7_S | CAGATTCCGGGGAGAAGATT | Exon 7 |
| A7E10_AS | AGCCGATGTACAGCAGGTTC | Exon 10 |
| All primers are from Eurogentec, Belgium | | |

**Table 2-3: Nucleotide sequences of the primers used for CHRFAM7A variants detection.**

| Primer ID | Primer sequence (5`→3`) | Position |
|-----------|------------------------|----------|
| A7EA_S | CCAGCATTTGTGGATAGCTG | Exon A |
| A7ED_S | TCAAGGCCAAACCGAAGTTA | Exon D |
| A7EC_AS | GATTCCAGGTCCTGCTGACT | Exon C |
| A7E10_AS | AGCCGATGTACAGCAGGTTC | Exon 10 |
| All primers are from Eurogentec, Belgium | | |

The RT-PCR was performed using a 30 µl reaction, using 1 µl of cDNA, in addition to the following reagents in a final concentration/reaction: 1x Taq buffer with $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, 0.2 mM dNTPs mix (Fermentas Life Sciences, UK), 0.3 µM of each primer, and 1 U Taq polymerase (Roche, USA). The thermal cycler was used for the PCR amplification (Applied Biosystems) using the conditions shown in Table 2-4.

**Table 2-4: RT-PCR conditions used.**

|  | Stage1 | Stage 2 | | | Stage3 | Hold |
|---|---|---|---|---|---|---|
|  | Initial denaturation | Denaturation | Annealing | extension | Final extension | - |
| **Cycles** | 1 | ←35→ | | | 1 | - |
| **Temperature** | 94° | 94° | 60° | 72° | 72° | 10° |
| **Time\*** | 2:0 | 00:30 | 00:45 | 1:00 | 7:00 | ∞ |
| \*: (minutes : seconds) | | | | | | |

A summary for CHRNA7 and CHRFAM7A transcripts targeted in A549 cells and their predicted sizes is shown in (Figure 2-9) and (Figure 2-10), respectively.
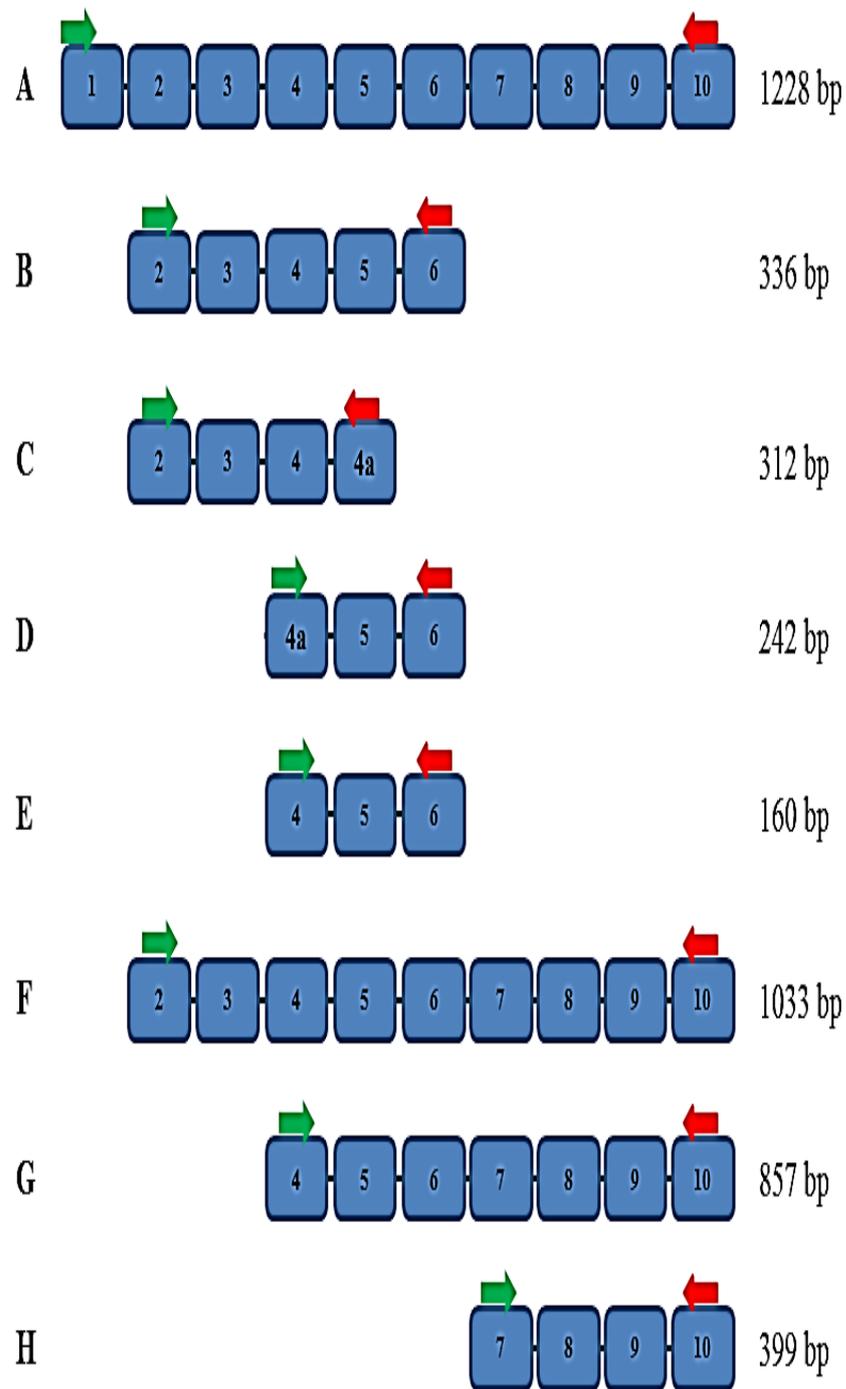
**Figure 2-9: CHRNA7 amplified parts.**
Diagram showing location of primers used to amplify different transcripts of CHRNA7 in A549 RNA. Blue boxes: exons, Green arrow: forward primer, Red arrow: reverse primer. The products are *A:* (exons1-10), *B:* (exons 2-6), *C:* (exons 2-4a), *D:* (exons 4a-6), *E:* (exons 4-6), *F:* (exons 2-10), *G:* (exons 4-10), *H:* (exons 7-10). Only exons included in the PCR amplification were included in the diagram. The expected product size (in bp) is shown on the right.
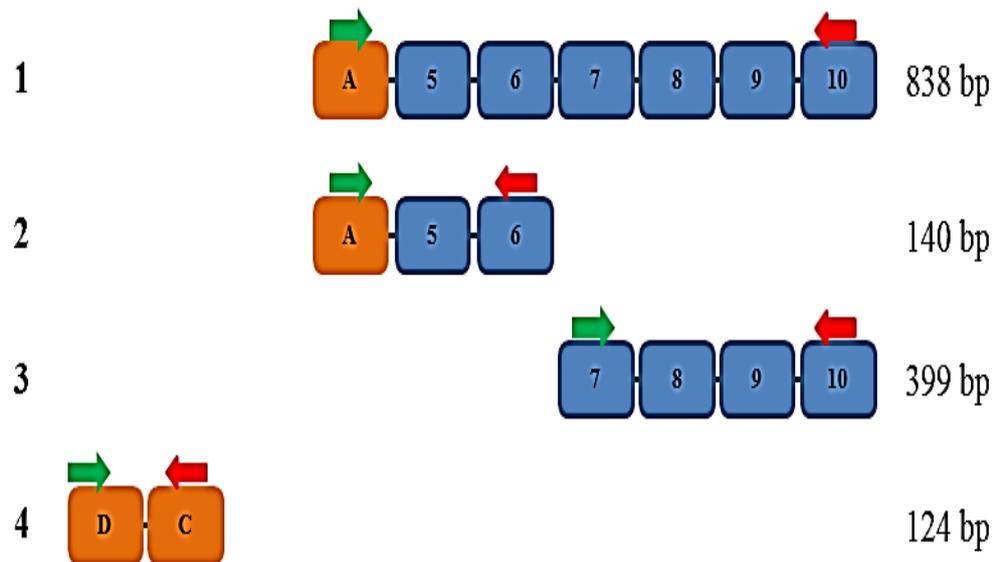
**Figure 2-10: CHRFAM7A amplified parts.**
Diagram showing location of primers used to amplify different transcripts of CHRFAM7A in A549 RNA. Blue boxes: exons, Green arrow: forward primer, Red arrow: reverse primer. The products are *1:* (exons A-10), *2:* (exons A-6), *3:* (exons 7-10), *4:* (exons D-C). Only amplified exons were included in the diagram. The expected product size (in bp) is shown on the right

### 2.6.5 Agarose gel electrophoresis:

PCR products were subjected to agarose gel electrophoresis using 1% agarose gel (Invitrogen, Paisley, UK), with 1 x Tris-acetic acid-EDTA (TAE) running buffer and 5 ng ethidium bromide (SIGMA-ALDRICH, UK) staining/25 ml of the gel. 5µl of 100bp DNA Step Ladder (Fermentas life Sciences, UK) was used as a size marker and the results visualised using UV light reader (Anachem, Luton, UK).

**2.6.6   TA cloning:**

This included PCR purification, insert ligation, transformation and restriction digestion. The purpose of the method is to clone PCR products with multiple amplicons. The method is based on hybridization of the Adenine nucleotide base (added by Taq polymerase enzyme to PCR products) with the Thymine base of a linearized vector with the aid of ligase enzyme. This hybrid DNA fragment is then transformed into *Escherichia coli* bacteria for subsequent amplification. Such a method would help separate multiple amplicons amplified using one primer set to detect different transcripts of the same gene (Holton and Graham, 1991).

**2.6.6.1 PCR purification:**

To improve the ligation efficiency, a clean-up step was performed prior to ligation to remove primer dimers. For this, QIAquick PCR purification kit was used (QIAGEN) according to manufacturer's guidelines. The DNA yield was measured using a Nanodrop spectrophotometer.

**2.6.6.2 Insert ligation:**

The PCR product was cloned into the pSC-A-amp/kan StrataClone vector mix (Stratagene, Germany) (Figure 2-11).

**β-galactosidase α-fragment (lacZ´)** 1–352
**multiple cloning site (MCS)** 57–195
**PCR product insertion site** 123
**kanamycin resistance ORF** 463–1254
**ampicillin resistance (bla) ORF** 1266–2123
**f1 origin** 2315–2621
**<loxP> (nonfunctional)** 2688–2721
**pUC origin** 3262–3929
**lac promoter** 4151–4270
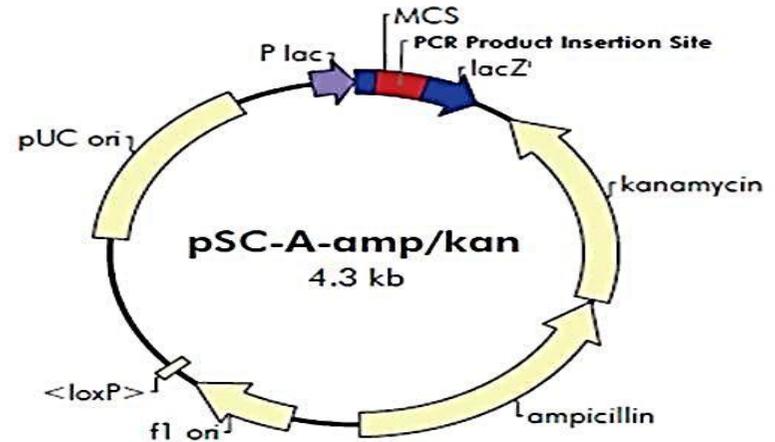


**pSC-A-amp/kan PCR Cloning Vector PCR Product Insertion Site Region**
**(sequence shown 4261–4270, 1–250)**
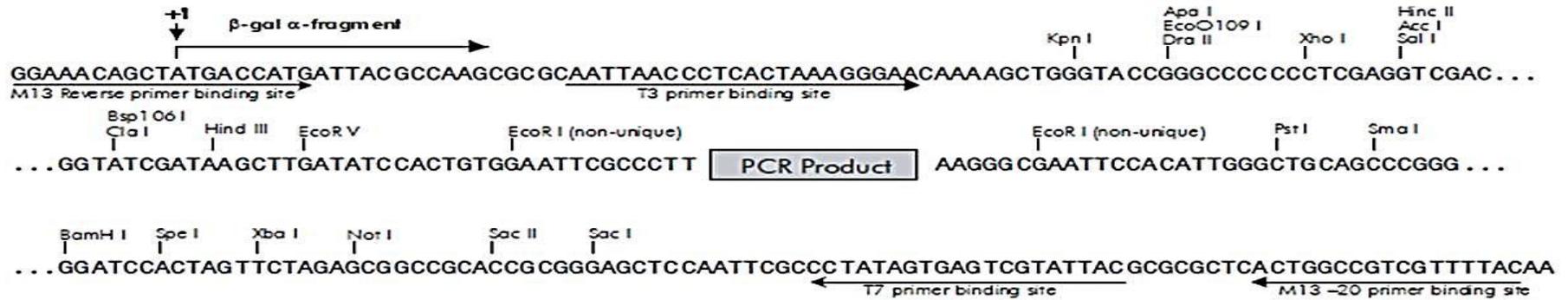
**Figure 2-11: pSC-A-amp/kan PCR cloning vector.**

*Top*: vector map, showing all included sites and features (illustrated). Most important of these is the multiple cloning sites (MCS) that is going to harbour the

PCR product insert site (illustrated). Bottom: PCR-insertion site. This site shows the recognition sites for the restriction enzyme and the primer sequence.

For ligation, a 6 µl reaction, the following reagents were added: 3 µl of StrataClone buffer, 2 µl of purified PCR product (about 25ng) and 1 µl of StrataClone vector mix amp/kan. This mix contains the two vector DNA arms with uridine overhangs that base pair with the adenosine overhangs of the Taq polymerase-amplified PCR products.

The reaction was then incubated at room temperature for 5 minutes.

**2.6.6.3 Transformation to competent cells:**

For transformation, StrataClone SoloPack competent cells (Stratagene, Germany) were used. These cells contain Cre-recombinase enzyme that enhances the recombination of the lox P sites on the two vector arms, thus helping the formation of circular plasmid-insert DNA molecule. Such a DNA can replicate in host cells growing on ampicillin or kanamycin-containing media. The cells were removed from storage at -80˚, and left on ice for thawing for 5 minutes. Then 1µl of ligation reaction mixture was added into the competent cells, and the reaction mixture was mixed gently (by repeated pipetting) and then incubated on ice for 20 minutes. The reaction was then heat-shocked at 42˚ for 45 seconds, and then put on ice for 2 minutes. Then 250µl of room temperature SOC solution (super optimal broth medium with glucose used to enhance the plasmid transformation efficiency) was added to each reaction tube. The reaction mixture was then incubated at 37˚ with horizontal shaking for 60 minutes. Finally, 100 µl of reaction mixture was spread on plates containing pre-warmed CG agar (Circle Grow rich bacterial growth media, Anachem, Luton, UK) (containing 4 g/100 ml CG, 1.5 g/100 ml agar and 50 µg/ml ampicillin and 20% X-gal) and incubated overnight at 37°.

Single colonies were inoculated into 5 ml CG bacterial growth media (including 50 μg/ml ampicillin) and incubated for 16-18 hours at 37° in a shaking incubator.

### 2.6.6.4  Insert-specific PCR:

A screening insert-specific PCR was carried out (using the original amplimers) to show which cultures contained the plasmid and insert. All other reagents were the same used for PCR mentioned in section 2.6.4 while 1 μl of liquid growth was used as the PCR template. The same PCR program shown in table (2-4) was used, except for a difference in the initial denaturation time (4 minutes rather than 2 minutes). The PCR products were subjected to gel electrophoresis for analysis (as mentioned in section 2.6.5).

### 2.6.6.5 Plasmid-insert extraction:

Using Pureyield Plasmid Miniprep System (Promega, USA), plasmid was extracted from bacterial cultures according to manufacturer's guidelines. The DNA yield was then measured using a Nanodrop spectrophotometer. This extraction method depends on the use of silica membrane columns that helps in nucleic acid extraction. The method includes precipitation of the cell debris using centrifugation and lysis buffer. The obtained supernatant (containing the plasmid DNA) can then be added on the silica membrane for subsequent purification using several buffers (such as the endotoxin removal buffer). These purification steps help preparing a plasmid DNA free of proteins, RNA, and endotoxin contaminants.

**2.6.6.6 Restriction digestion:**

In order to identify plasmids containing the insert of interest, plasmid DNA was screened using restriction digestion with EcoRI. This enzyme will recognise the sequence G: AATTC and then generates fragments that contain 5`-cohesive termini. The following was added into 10µl reaction: 600 ng of plasmid DNA, 1µl of 10x cut buffer-H (Roche), 5U of EcoRI restriction enzyme (Roche) and the reaction volume completed by adding nuclease-free water. The reaction mixture was then incubated for 2 hr at 37˚. For detection of the size of predicted inserts, the digested DNA was run on 1% agarose gel (Invitrogen, Paisley, UK), with 1 x TAE running buffer and 5ng of ethidium bromide (SIGMA-ALDRICH). 5µl of 100bp DNA stepladder (Fermentas life sciences, UK) was used as a size marker, and the results visualised using UV light reader (Anachem, Luton, UK). Equivalent samples from the undigested DNA were run on the same gel to compare results.

**2.6.7    Gel extraction:**

The QIaquick gel extraction kit (QIAGEN) was used for extraction of PCR products from the agarose gel. The PCR products with multiple and relatively separable bands were chosen for gel extraction. First, the PCR product of interest was run on a range of 1-3.5% agarose gel (as mentioned in section 2.6.5) to visualise the bands of interest. The gels obtained were put in new 1.5ml micro centrifuge tubes, and weighed. An empty 1.5ml micro centrifuge tube weight was subtracted from the gel-containing tubes weights. The yielded weight (in mg) was considered as an equivalent one volume in µl. Three gel volumes of QG buffer was added to the gel and the reaction was incubated for

10 minutes at 50˚ with repeated vortexing to ensure the gel was completely dissolved. One gel volume of isopropanol (Fisher Scientifics, UK) was added and mixed with the dissolved gel mixture. To bind DNA, the mixture was then applied to a new QIAquick column and centrifuged for 1 minute at 13000 rpm. The flow-through was discarded. Then 0.5 ml of QG buffer was added to the column and centrifuged for 1 minute at 13000 rpm (10000 g). The flow-through was discarded. To wash the DNA, 0.75 ml of PE buffer was then added to the column and centrifuged for 1 minute at 13000rpm and the flow-through was discarded. The centrifugation was repeated for 1 minute at 13000 rpm to ensure complete removal of the buffer. To elute the DNA, the column was then put in a new 1.5ml micro centrifuge tube, and 30 µl of BE buffer was applied on column, centrifuged for 1 minute at 13000 rpm. The DNA yield was measured using a Nanodrop spectrophotometer.

### 2.6.8   DNA Sequencing:

For PCR products with single band, a clean-up step was applied prior to DNA sequencing, while TA cloning or gel extraction products were sequenced directly without this step. The sequence clean-up reaction included adding 2U of EXOSAP-IT enzyme (USB, USA) to treat 5 µl of the PCR products (to remove the residual primers and dNTPs from the PCR products to be sequenced). The samples were then incubated in thermal cycler (Applied Biosystems) using the following programme (1 cycle): 15 minutes at 37˚C and then 15 min at 80˚.

Sequencing of PCR products was carried out using one of the primers used for the initial PCR amplification (Table 2-2, Table 2-3, & **Error! Reference**

**source not found.**). The samples were prepared on ice using the following: 1 µl of sequencing primer (5µM), 2 µl of Big Dye v3.1 Terminator Ready reaction mix (Applied Biosystem)  and 1x sequencing buffer (Big Dye terminator, v3.1) (Applied Biosystem) in addition to the DNA template. Quantities of DNA template used ranged from 3µl of the enzyme-treated PCR product (~20ng), 5-20 ng of gel-extracted product, or 150-500ng of TA cloned DNA. The mixture was incubated in a thermal cycler (Applied Biosystems) using the following programme (25 cycles):  10 seconds at 90˚, 5 seconds at 50˚, 4 minutes at 60˚ and the samples then kept at 10˚ for hold. For post-reaction clean up, the sequence column purification tubes (Edge Biosystems Performa DTR Gel Filtration Cartridges) were used to remove dye terminators, dNTPs, and other low molecular weight materials from the sequencing reaction. First, the columns were centrifuged for 3 minutes at 4000 rpm (850 g) to remove excess buffer from tubes. The columns were then transferred to a new tube, and the sequencing products were then added on the column membrane, and centrifuged for 3 minutes at 4000 rpm (850 g). The eluates were transferred into new 0.5 ml tubes to be evaporated to dryness at 90˚. To confirm the correct sequence for each of the transcripts detected from PCR products, gel extractions, or bacterial clones, DNA sequencing was applied using forward and reverse primers for each transcript. All samples were separated by capillary electrophoresis for sequencing.

The sequences for primers M13 forward and reverse are shown in **Error! Reference source not found.**.

**Table 2-5: Nucleotide sequence for pSC-A-amp/kan vector primers.**

| Primer ID | Primer sequence |
|-----------|-----------------|
| **M13 F** | GGAAACAGCTATGACCATG |
| **M13 R** | GTAAAACGACGGCCAGT |
| **F: Forward primer, R: Reverse primer, M13 primers (100 ng/µl, Invitrogen)** | |

### 2.6.9    Quantitative Real-time PCR (QRT-PCR):

To estimate the expression of the two major transcripts of CHRNA7 by A549 cells, QRT-PCR was used. In order to avoid detection of the CHRFAM7A transcripts, the sense primer was located outside of the duplicated region, in exon 4 (Table 2-2). The antisense primers were designed to bridge the exon-exon boundaries between exons 8 and 9 (for full-length CHRNA7 transcript) and between exons 8 and 10 (for missing exon 9 CHRNA7 transcript) (Table 2-6) (Figure 2-12).

**Table 2-6: Reverse primers used for QRT-PCR amplification of CHRNA7 two major transcripts in A549 cells.**

| Primer ID | Primer sequence | Position |
|-----------|-----------------|----------|
| **A7E8/9_AS** | CGAAGTACTGGGCTATCAATGG | Exons 8-9 junction |
| **A7E8/10_AS** | GACTCTGGTCTATCAATGGTACC | Exons 8-10 junction |
| All primers are from Eurogentec, Belgium | | |
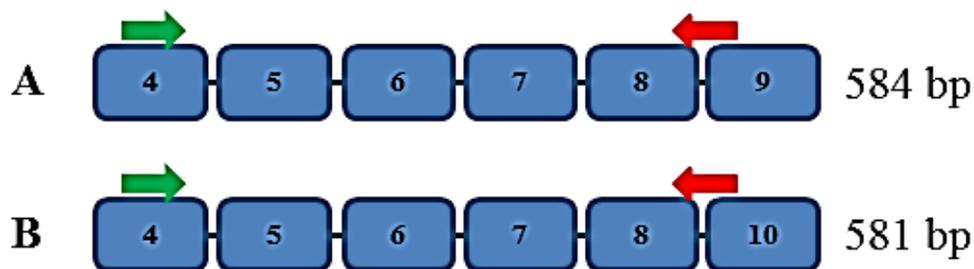
**Figure 2-12: CHRNA7 amplified parts for QRT-PCR.**
Diagram showing location of primers used to amplify two major transcripts of CHRNA7 in A549 using QRT-PCR. Blue boxes: exons, Green arrow: forward primer, Red arrow: reverse primer. The products are *A:* (exons 4-8/9), *B:* (exons 4-8/10). The expected product size (in bp) is shown on the right.

SYBR GREEN BRILLIANT III master mix (Agilent technologies) was used for the QRT-PCR. The QRT-PCR master mix was prepared into a final volume of 20 µl by adding 10 µl/reaction of SYBR green brilliant III master mix, 0.3 µl/reaction of the diluted reference dye (1/50), 0.3 µM primer concentration and 5 µl/reaction of the diluted A549 cDNA and the reaction volume was completed using RNase-free water. The cDNA: RNase-free water dilutions used were 1:4 (1/5) and 1:19 (1/20) in duplicates for each primer set. Then PCR products were subjected to real time PCR machine analysis in addition to agarose gel electrophoresis (as mentioned in section 2.6.5).

For primer specificity check of each primer set used, the cDNA templates used for QRT-PCR were used for RT-PCR using the same primer sets. These were then used for TA cloning to separate the resulted products (as mentioned in section 2.6.6). This was followed by using the prepared clones as templates representing full-length or missing exon 9 (according to the primer set used) for another RT-PCR by using both primer sets in two separate reactions. This test was aimed at showing whether the primer sets designed were specific or not (that is full-length primers were supposed to amplify full-length clones, and

the same for missing exon 9 part). Previously prepared clones (representing full-length exons 1-10 and missing exon 4) were used as control templates.

### 2.6.10  RT-PCR semi-quantitative assay:

This assay used same RT-PCR described in section 2.6.4. The primer sets used to amplify exons 1-10 of CHRNA7 (Table 2-2; Figure 2-9) and exons A-10 of CHRFAM7A (Table 2-3; Figure 2-10). PCR samples were prepared in duplicates for cycles 27-35, and cycle-specific samples were removed from the thermocycler after each specified cycle and kept on ice (e.g. 27 cycle-specific samples were removed after cycle 27, and so on). After completing second stage of RT-PCR, the samples were re-introduced into the thermocycler for the elongation step till the end of the program. Then PCR products were subjected to gel electrophoresis (as mentioned in section 2.6.5). The procedure was repeated at least three times to achieve reproducible results.

### 2.6.11  AlphaDigiDoc software and statistical analysis:

To evaluate the relative percentage of mRNA expression for the major transcripts of CHRNA7 and CHRFAM7A, the gel electrophoresis pictures were analysed using AlphaDigiDoc1201 software. The parameter used for comparisons was the band intensity for each transcript, which was evaluated across duplicated samples for different PCR cycles (27-35) to determine the exponential phase and the plateau phase (relative start and end points), and to determine the relative percentage of expression of each transcript by A549 cells (for each gene transcripts).

**2.6.12  Bioinformatics used for analysis of transcripts:**

First, for the extraction of the reference sequence for CHRNA7 and CHRFAM7A, database websites for corresponding human reference sequences were used from *NCBI*, *UCSC*, and *Ensembl*. Then for sequence alignment of the common parts of the two genes, *ClustalW* software program was used ((Larkin et al., 2007). For prediction of the possible restriction enzyme recognition sites within exon 9b and flanking introns, *Web Cutter* and *Neb Cutter* software programs were used. For primer design, *primer3* (v.0.4.0) software was used for designing the primers, *Primer-Blast/NCBI* software was used to check for the primer specificity. To look for the consensus sequences of ESE and ESS, the predicted ESE (PESE) and predicted ESS (PESS), *RESCUE-ESE* and *Human splicing finder* software programs were used (Fairbrother et al., 2002). Using *ORF finder* software programme, the mRNA of resultant splice variants was translated into corresponding protein code. Then for determining the KOZAK consensus sequence, *weakAUG* software program (Tikole and Sankararamakrishnan, 2008) was used. This was followed by using the *SOSUI signal* software program (Masahiro, 2005) for the prediction of a possible signal peptide. Then using *PSIPRED* software (Jones, 2007), the secondary protein structure was predicted for each transcript. A summary for the software programs using for bio-informatics analysis and their corresponding URLs are listed in (Table 2-7).

**Table 2-7: Summary for the bio-informatics software programs and tools used for analysis of CHRNA7 and CHRFAM7A transcripts.**

| Software program | URL |
|---|---|
| *ClustalW (sequence alignment)* | http://www.ebi.ac.uk/Tools/msa/clustalw2/ |
| *Web cutter (searching for restriction recognition sites)* | http://users.unimi.it/~camelot/tools/cut2.html |
| *Neb cutter (as above)* | http://tools.neb.com/NEBcutter2/ |
| *Primer 3 (v.0.4.0) (PCR primer design)* | http://frodo.wi.mit.edu/ |
| *Primer blast/NCBI (check primer potential target)* | http://www.ncbi.nlm.nih.gov/tools/primer-blast/ |
| *RESCUE-ESE (look for potential ESE sequences)* | http://genes.mit.edu/burgelab/rescue-ese/ |
| *Human Splicing Finder (look for splicing sequences)* | http://www.umd.be/HSF/ |
| *ORF finder (look for ORF with a start and stop codons)* | http://www.ncbi.nlm.nih.gov/gorf/gorf.html |
| *WeakAUG (detect possible weak start codons)* | http://bioinfo.iitk.ac.in/AUGPred/index.php |
| *SOSUI (detect signal peptides)* | http://bp.nuap.nagoya-u.ac.jp/sosui/ |
| *PSIPRED (test for possible protein secondary structure)* | http://bioinf.cs.ucl.ac.uk/psipred/ |

**2.6.13 Western blotting:**

**2.6.13.1      Cell lysate preparation:**

For cell lysate preparation, all reagents were kept on ice throughout the procedure. RIPA buffer (20 mM Tris pH7.5, 150 mM NaCl, 1% triton x-100, 0.5% sodium deoxycholate, 1 mM EDTA and 0.1% SDS) was prepared and mixed with 5x protease inhibitor mixture (Complete Mini protease inhibitor cocktail tablets, Roche). First, the cells were cleaned from the old medium and dead cells, and then washed with ice-cold PBS. Then 600 μl of the prepared buffer mixture was added into the cells for 5 minutes, and then cells were detached and transferred into 1.5ml micro centrifuge tubes. The mixture was then rotated in a cold room for 30 minutes to allow solubilisation of the proteins, before separating the lysate from the cell remnants using a precooled centrifuge at 4°C for 20 minutes at 12000 rpm (4000 g). After that, the supernatant was collected and transferred into a new tube and stored at -80°C. Then the lysate protein contents were quantified using BCA protein quantitative assay kit (Thermo Fisher Scientific, USA), according to manufacturer's guidelines. The protein samples were quantified after that using FluoSTAROPTIMA spectrometer at 550 nm wavelength (BMGs LABTECH, Germany).

**2.6.13.2      SDS-PAGE:**

Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) was used to separate the proteins according to their sizes. This constituted discontinuous 5-10% gel. The constituents of gel used are included in (Table 2-8).

**Table 2-8: SDS-PAGE reagents.**

| Substance | Stacking gel | Separating gel |
|---|---|---|
| Gel percentage | 5% | 10% |
| Polyacrylamide mix 40% | 625 µl | 5 ml |
| Tris base | 625 µl (pH 6.8), 1M | 5 ml (pH 8.8), 1.5 M |
| Ammonium Persulfate (APS) 10% | 50 µl | 200 µl |
| Sodium Dodicyl Sulfate (SDS) 10% | 50 µl | 200 µl |
| Tetramethylethyl diamine (TMED) | 5 µl | 8 µl |
| Distilled water | 3.65 ml | 9.6 ml |
| Total volume | 5 ml | 20 ml |

The samples were prepared by adding aliquots of A549 cell lysate mixed with 4x loading buffer (240 mM Tris-HCL/Tris base pH 6.8, 8% SDS, 0.04% bromophenol blue, 5% β-mercaptoethanol and 40% glycerol). The mixture was then incubated at 70˚C for 10 minutes, put on ice before loading on the gel. To monitor the protein migration during SDS-PAGE, a protein ladder was used to identify the protein size between 10-260 kDa (Fermentas life sciences, UK). The power supply was set to 150 volts, 50 mA, for 1:30 hr.

**2.6.13.3        Semi-dry Transfer:**

The transfer was carried out using PVDF (polyvinylidene fluoride) membrane. The transfer sandwich was prepared as follows (from down upwards): Three blotting papers, PVDF, Gel, and three blotting papers on top (with the gel near the cathode), all soaked in transfer buffer for 5 minutes prior to arrangement, except PVDF which was soaked in absolute methanol then in the transfer

buffer. The transfer buffer was composed of the following: 48 mM Tris HCL/Tris base, 39 mM glycerin and 20% Methanol.


### 2.6.13.4        Immunoblotting:

The membrane was blocked overnight with 5% skimmed milk. This was followed by incubation with primary rabbit polyclonal IgG antibodies for $\alpha 7$ protein (ab10096, Abcam, UK) for 1:30 hr., and secondary goat polyclonal to rabbit antibodies conjugated to horse radish peroxidise (HRP) (ab6721, Abcam, UK) for another 1:30 hr. Each of the blocking and antibody incubation steps was followed by washing with PBS-Tween buffer (0.1% Tween-20 and phosphate buffered saline) for three times, (10-15 minute duration/each). The membranes were probed with antibodies for $\beta$-actin as a positive control for protein in A549 (primary mouse monoclonal IgG (SIGMA-ALDRICH, UK), and secondary goat anti mouse IgG (SIGMA-ALDRICH, UK)). The blocking agent and antibodies were all diluted in the 5% milk-PBS-Tween wash buffer. The dilutions used for all antibodies are summarized in (Table 2-9).

**Table 2-9:  Sources, dilution and expected protein sizes for antibodies used.**

| Antibody | Source | Dilution | Expected band size |
|---|---|---|---|
| **Primary antibody for $\alpha 7$** | Abcam, UK | 1:400 | 56 kDa |
| **secondary antibody for $\alpha 7$** | Abcam, UK | 1:3000 | |
| **Primary antibody for $\beta$-actin** | SIGMA-ALDRICH, UK | 1:4000 | 46 kDa |
| **secondary antibody for $\beta$-actin** | SIGMA-ALDRICH, UK | 1:10000 | |

**2.6.13.5        Enhanced chemiluminescence (ECL):**

The expected antibody-protein complexes were detected using ECL reagents spray (Calbiochem). The membranes were sprayed twice, and incubated on the bench at room temperature for 1 minute. This was followed by using Kodak films in a dark room to visualise the protein bands. The exposure duration time used was for 15, 30, 45, 60, 90, 120, 300, and 600 seconds.

**2.6.13.6        Protein deglycosylation:**

Using deglycosylation kit (New England Biolabs), PNGase F enzyme (peptide: N-Glycosidase F) was used for deglycosylation of A549 lysates prior to Western blot analysis. For each reaction, 50 µg of A549 lysate protein was used.  This was mixed with 1x glycoprotein denaturing buffer, and distilled water was used to complete the 20 µl reaction volumes. To denature the glycoprotein, the mixture was heated at 100˚ for 10 minutes. Then, the following reagents were added: G7 reaction buffer (1x), NP-40 (1%) and 1000 units of PNGase F enzyme. The final reaction volume (of 40 µl) was completed using distilled water. RNase B, a positive control, was used for the deglycosylation reaction.

## 2.7   Results:

### 2.7.1   Characterisation of CHRNA7 transcripts in A549 cells:

Using primers specific for different parts of CHRNA7 (Table 2-2), RT-PCR was used to amplify the target parts as illustrated in Figure 2-9. The PCR products detected from A549 cells cDNA using exons 2-6 primer sets are shown in (Figure 2-13).
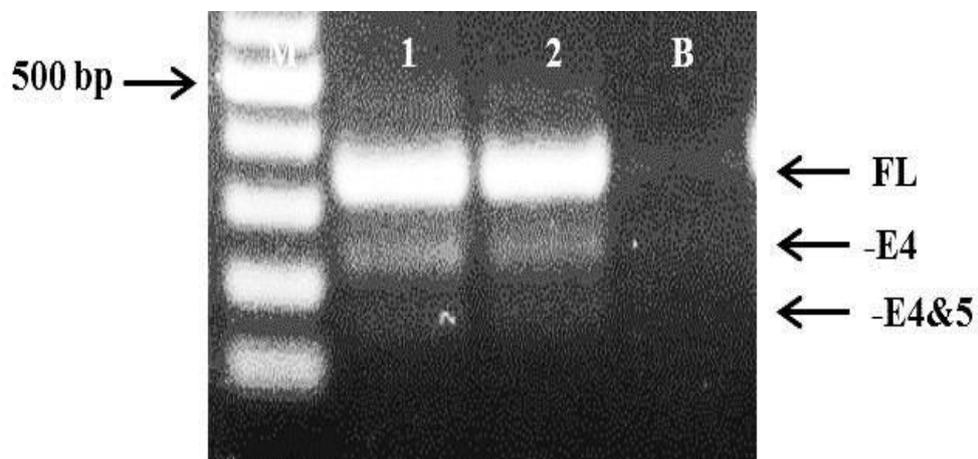


**Figure 2-13: Agarose gel electrophoresis for CHRNA7 transcript E2-6.**
*Lanes 1-2:* duplicate samples of exon 2-6 product from cDNA of A549 (expected size is 336 bp=full-length (FL). Two smaller bands of lower intensity than the full length were detected, the middle one can be observed (matching to predicted sizes of transcripts missing exon 4 (226 bp size=-E4) and the lowermost one that was too faint on the gel (matching the predicted size of missing exons 4&5 (146 bp size=-E4&5)), respectively. These bands were visible on the gel but were too faint to be captured on the image. M: 100 bp DNA ladder, B: blank.

The use of other primer sets (spanning different parts of CHRNA7) showed the expression of expected band sizes. However, additional novel transcripts were also detected (Figure 2-14).
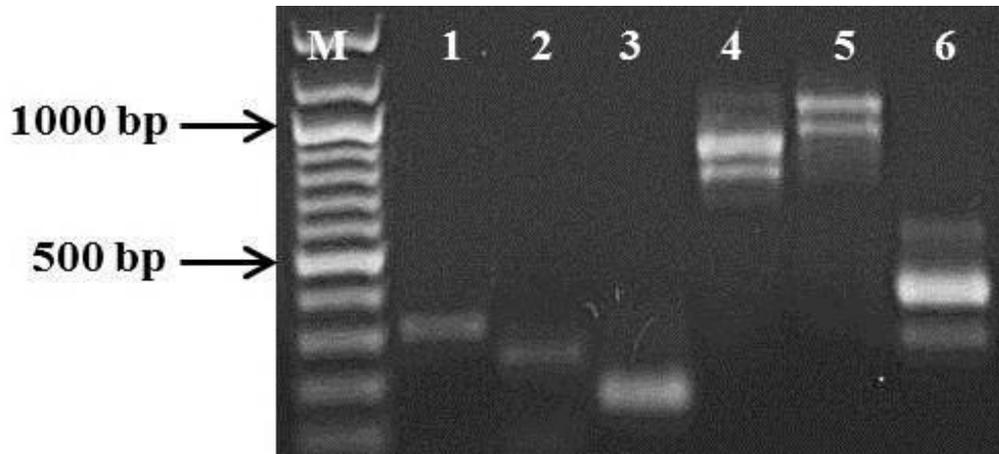


**Figure 2-14: Agarose gel electrophoresis for all tested CHRNA7 transcripts.**
Different primer sets were used to amplify CHRNA7 exons (E). Lanes 1-6 show A549 amplified cDNA using CHRNA7 specific primers. *1:*E2-4a (expected size is 312 bp), *2:*E4a-6 (expected size=242 bp), *3:* E4-6 (expected size=160 bp), *4:* E4-10(expected size= 857 bp), *5:* E2-10 (expected size= 1033bp), *6:* E7-10 (expected size= 399 bp). Larger size products (lanes 4-6) showed additional products with deletion (lower than the expected sizes for the primer pairs used), while only lane 6 (the only product amplified from both genes) showed an extra product with insertion (larger than the expected size for the primer pair used). M: 100 bp DNA ladder.

For short amplicons (1-3), only a single product was detected. For larger amplicons (4-6), two major products were detected. However, PCR product 6 (amplifying exons 7-10) showed a brighter main band (middle band), a second transcript with deletion similar to those noticed with PCR products 4 & 5 (lower band) and a third transcript with an insertion (top band).

Following the RT-PCR results, direct DNA sequencing for products containing a single band (E2-4a, 4a-6, and 4-6) and the top band products of (E2-6) was gel extracted and showed that these were full-length products, as shown in (Table 2-10).

**Table 2-10: Sequencing results for PCR products.**

| PCR product | Exons included | Product size (bp) |
|---|---|---|
| Exons 2-6 | 2,3,4,5,6 | 336 |
| Exons 2-4a | 2,3,4,4a | 312 |
| Exons 4a-6 | 4a,5,6 | 242 |
| Exons 4-6 | 4,5,6 | 160 |

The full length of CHRNA7 was amplified using specific primers between exons 1-10. Similar results were obtained from RT-PCR with two major transcripts, the full-length, and a product with around 100bp deletion (Figure 2-15).
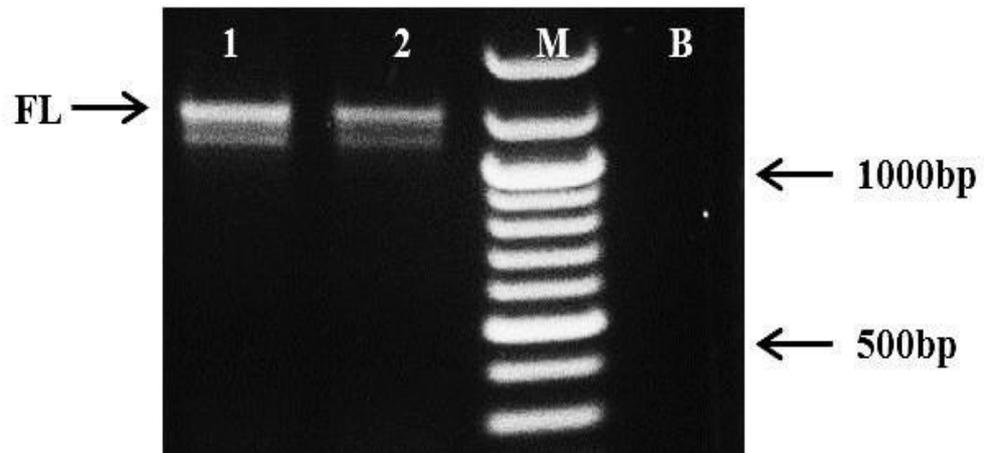


**Figure 2-15: Agarose gel electrophoresis for CHRNA7 transcript E1-10.**
*Lanes 1-2:* duplicate samples of exon 1-10 product from cDNA of A549 (expected size is 1173 bp=FL). The FL PCR product (arrowed) was detected in addition to another product of smaller size (by about 100 bp) and of almost similar intensity. M: 100 bp DNA ladder, B: blank.

For products with multiple bands, TA cloning or gel extraction or both were applied. For products with relatively separable bands after gel electrophoresis, gel extraction was used; otherwise, the TA cloning technique was applied for products with band closer to each other.

The use of the gel extraction strategy made it possible for several bands to be separated using a high percentage gel electrophoresis, extracted and then sequenced. Surprisingly, all large size products showed the expression of a novel transcripts missing exon 9. This transcript was not reported in any previous study of CHRNA7 transcripts. The results of gel extraction are summarised in Table 2-11.

**Table 2-11: Summary for sequencing of gel extraction products.**

| PCR product | Product size (bp) | Transcripts detected |
|---|---|---|
| E2-6 | 366 | E2-6 FL |
| | 226 | missing exon 4 |
| | 146 | missing exons 4&5 |
| E1-10 | 1173 | E1-10 FL |
| | 1063 | missing exon 9* |
| E2-10 | 1033 | E2-10 FL |
| | 923 | missing exon 9* |
| E4-10 | 857 | E4-10 FL |
| | 747 | missing exon 9* |
| E7-10 | ~ 600 | Inserted product |
| | 399 | E7-10 FL |
| | 289 | missing exon 9* |

E: exon, FL: full length included. *: indicates novel transcripts detected.

The sequencing results were non-informative for the larger PCR product (with insertion) from amplifying exons 7-10. This gel-extracted product was further explored using the TA cloning method.

Whenever the use of gel extraction was impractical, TA cloning was applied. First, the PCR products that were used for TA cloning included products amplifying exons 2-6, 1-10, and 7-10 of CHRNA7. For all TA cloning, ligation and transformation were applied to insert the PCR product into pSC-A-amp/kan cloning vector. This was followed by PCR screening to distinguish clones with same insert and avoid repeating plasmid-insert extraction of the same transcript. Following plasmid-insert DNA extraction, the samples were subjected to screening using EcoRI restriction enzyme to distinguish transcripts prior to sequencing.

**2.7.1.1 TA cloning for exons 2-6 PCR product of CHRNA7:**

This PCR product included CHRNA7 transcripts enclosed between exons 2 &

6. All the clones were sequenced and a summary of the results is detailed in

(Table 2-15). Then PCR screening (using exons 2-6 primer set) was used prior

to miniprep plasmid preparation showed products of full-length exons 2-6 and

missing exon 4 based on expected sizes (Figure 2-16).
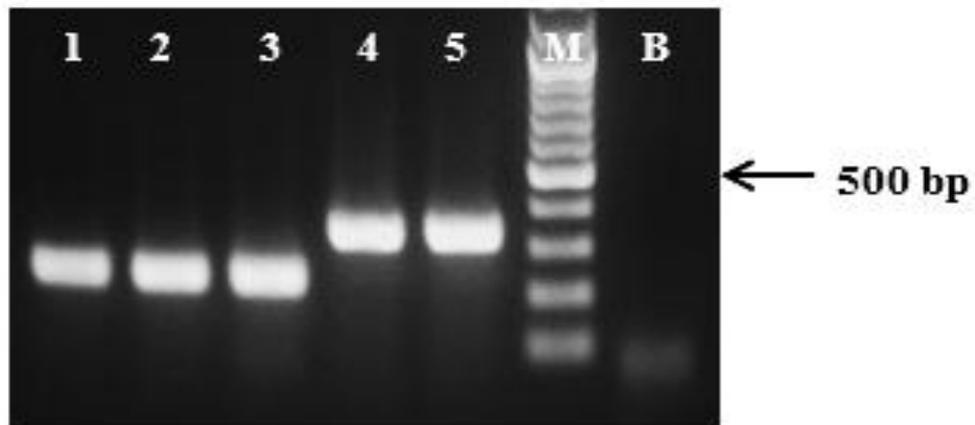


**Figure 2-16: Agarose gel for PCR screening of exons 2-6 products.**
PCR screening of TA clones from exons 2-6 PCR product of A549 (*samples 1-5*). The

clones 4 & 5 match the full-length transcript of exons 2-6. Clones 1-3 match transcript

missing exon 4. M: 100 bp DNA ladder, B: blank.

EcoRI restriction sites on pSC-A-amp/kan vector will cut the vector from the insert PCR product. In addition, EcoRI will cut the insert within exon 5 (at base 45 of exon 5) cutting the insert product of 336 bp into 264bp and 72bp parts in addition to the band corresponding to the plasmid backbone (Figure 2-17) (Table 2-12).
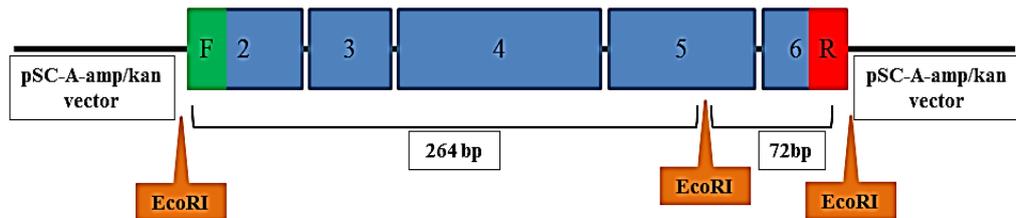


**Figure 2-17: Restriction digestion of plasmid-insert (exons 2-6) DNA using EcoRI.**
The figure shows EcoRI recognition sites on pSC-A-amp/kan vector and exons 2-6 insert (original size=336bp) and possible digested insert parts sizes=264bp and 72bp, respectively. F: forward PCR primer, R: reverse PCR primer, 2-6: CHRNA7 exons amplified.

**Table 2-12: Predicted digestion product sizes for alternative transcript inserts from cloning of exons 2-6 PCR products.**

| Exons included | Product(s) size(s) |
|---|---|
| **Missing exon 3** | 219, 72 |
| **Missing exon 4** | 154, 72 |
| **Missing exon 5** | 256 (uncut) |
| **Missing exons 3&4** | 110, 72 |
| **Missing exons 4&5** | 146 (uncut) |
| **Missing exons 3,4&5** | 101 (uncut) |
| **With exon 4a** | 392, 72 |
| **Without exon 4a** | 264, 72 |

According to the restriction digestion results, samples were sequenced (Figure 2-18).
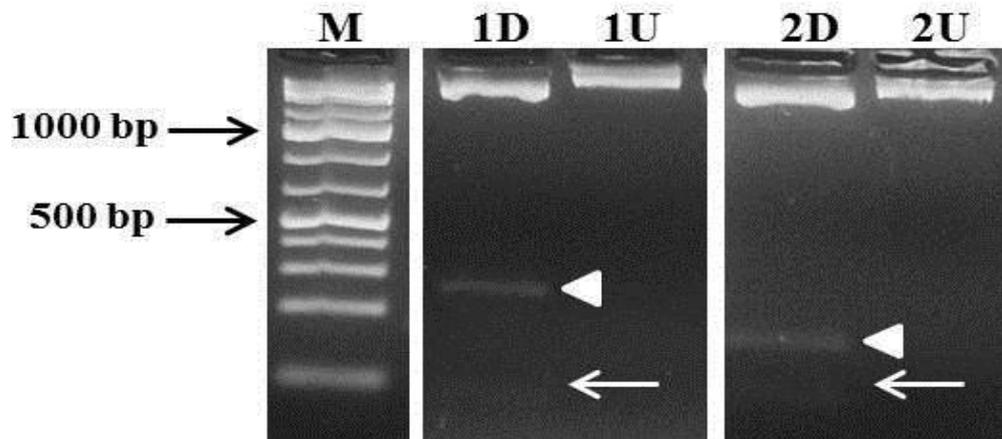


**Figure 2-18: Agarose gel for restriction digestion using EcoRI for exons 2-6 clones.**
*Lanes 1D/1U:* Full-length exons 2-6 product, lanes *2D/2U:* missing exon 4 product (refer to Table 2-12 for product sizes). D: digested DNA using EcoRI, U: undigested DNA. Upper bands (white arrowhead) and lower bands (white arrow) representing the two digested parts, with the lower bands were visualised on the gel but were too faint to be captured on the image. M: 100bp DNA ladder.

**2.7.1.2 TA cloning for exons 1-10 PCR product of CHRNA7:**

This PCR product included CHRNA7 transcripts enclosed between exons 1 &

10. About 25 % of the clones were sequenced and a summary of the results is

detailed in Table 2-15.Then PCR screening was used prior to miniprep plasmid

preparation with most of the clones showing products of full-length exons 1-10

based on expected sizes (Figure 2-19).
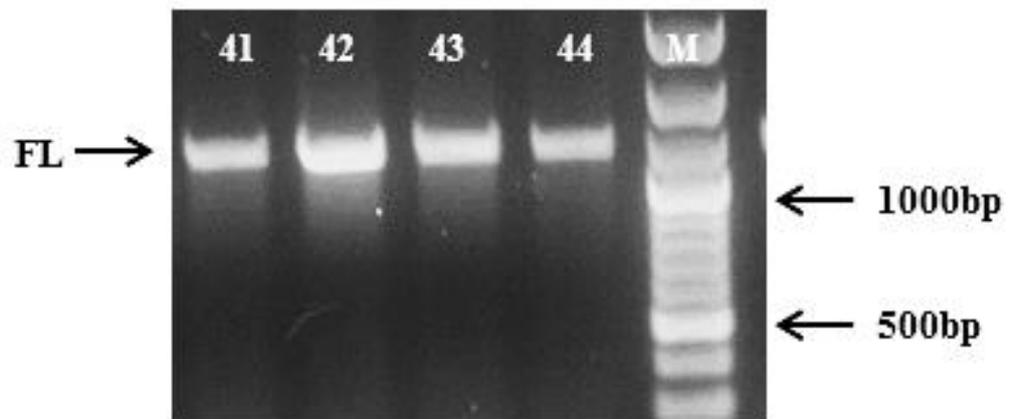


**Figure 2-19: Agarose gel for PCR screening for exons 1-10 products.**
PCR screening of TA clones from exons 1-10 PCR product of A549 (*samples 41-44*).

All the clones shown in this figure match the full-length transcript of exons 1-10

(size=1173 bp). M: 100 bp DNA ladder.

EcoRI restriction sites on pSC-A-amp/kan vector will cut the vector from the insert PCR product. In addition, EcoRI will cut the insert within exon 5 sequence (at base 45 of exon 5) cutting the insert of 1173 bp into 403bp and 770bp fragments in addition to the band corresponding to the plasmid backbone (Figure 2-20) (Table 2-13).
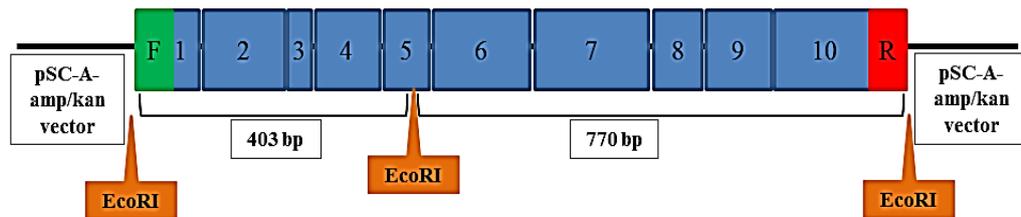


**Figure 2-20: Restriction digestion of plasmid-insert (exons 1-10) DNA using EcoRI.**
The figure shows EcoRI recognition sites on pSC-A-amp/kan vector and exons 1-10 insert (original size=1173bp) and possible digested insert parts sizes=403bp and 770bp, respectively . F: forward PCR primer, R: reverse PCR primer, 1-10: CHRNA7 exons amplified.

**Table 2-13: Predicted digestion product sizes for alternative transcript inserts from cloning of exons 1-10 PCR products.**

| Exons included | Product(s) size(s) |
|---|---|
| **Missing exon 3** | 358, 770 |
| **Missing exon 4** | 293, 770 |
| **Missing exon 5** | 1093 (uncut) |
| **Missing exons 3&4** | 248, 770 |
| **Missing exons 4&5** | 983 (uncut) |
| **Missing exons 3,4&5** | 938 (uncut) |
| **With exon 4a** | 530, 770 |
| **Without exon 4a** | 403, 770 |
| **Missing exon 9** | 403, 660 |

According to the restriction digestion results, samples were sequenced (Figure

2-21).
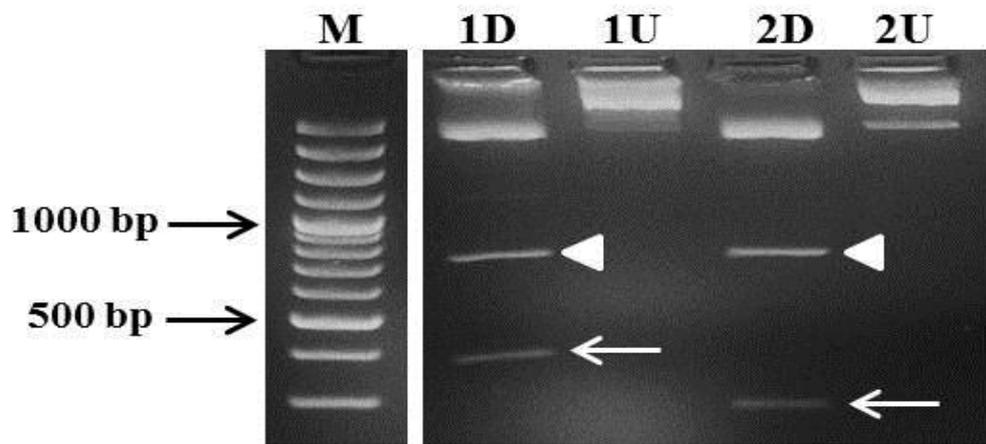


**Figure 2-21: Agarose gel for restriction digestion using EcoRI for exons 1-10 clones.**
*Lanes 1D/1U:* Full-length exons 110 product, lanes *2D/2U:* missing exon 4 product

(refer to Table 2-13 for product sizes). D: digested DNA using EcoRI, U: undigested.

Upper bands (white arrowhead) and lower bands (white arrow) representing two the

two digested parts. M: 100bp DNA ladder.

### 2.7.1.3 TA cloning for exons 7-10 PCR product of CHRNA7:

This PCR product included CHRNA7 transcripts enclosed between exons 7 &
10. About 25 % of the clones were sequenced and a summary of the results is
detailed in Table 2-15. This is the only primer set that was designed to amplify
the exons found in both genes CHRNA7 and CHRFAM7A. Then PCR
screening was used prior to miniprep plasmid preparation with most of the
clones showed products of full-length exons 7-10 based on expected sizes
(Figure 2-22).
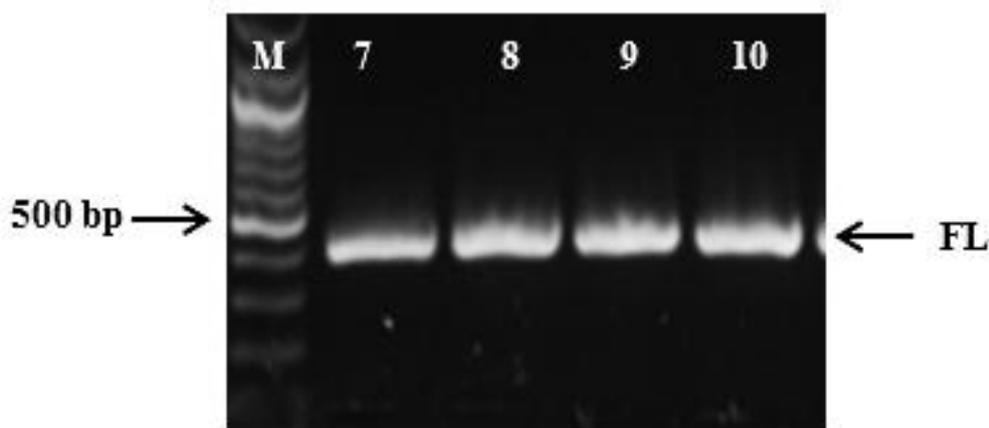


**Figure 2-22: Agarose gel for PCR screening for exons 7-10 product.**
PCR screening of TA clones from exons 7-10 PCR product of A549 (*samples 7-10*).
All the clones match the full-length transcript of exons 7-10 (size =399 bp). M: 100 bp
DNA ladder.

EcoRI restriction sites on pSC-A-amp/kan vector will cut the vector from the insert PCR product. However, EcoRI will not cut the insert leaving one band for the insert full length = 399bp in addition to the band corresponding to the plasmid backbone (Figure 2-23).
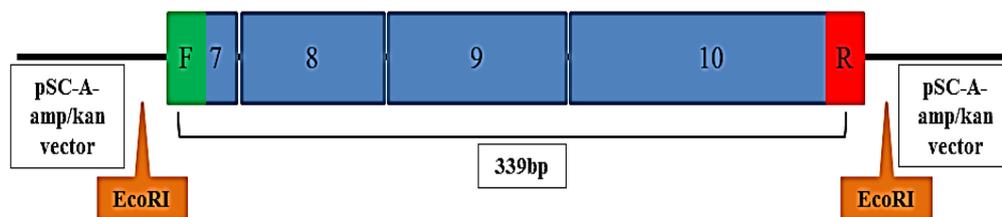


**Figure 2-23: Restriction digestion of plasmid-insert (exons 7-10) DNA using EcoRI.**
The figure shows EcoRI recognition sites on pSC-A-amp/kan vector and exons 7-10 insert (original size=399bp) uncut. F: forward PCR primer, R: reverse PCR primer, 7-10: CHRNA7 exons amplified.

The transcript that was expected to be noticed is missing exon 9 (proved earlier by gel extraction) in addition to the transcript with addition with larger size than exons 7-10 (Figure 2-14). However, only full-length transcript (exons 7-10) could be detected. The earlier results of gel extraction and sequencing for the isolated three bands detected for exons 7-10 PCR product showed that the middle band is full length (exons 7-10) and the lower band is missing exon 9, but the upper band was detected as having only full-length transcripts. It was assumed that the transcript with insertion (Figure 2-14, lane 6, upper band) have less copies than those of the full-length transcripts. As the two PCR products' (insertion and full-length transcripts) bands were so close to each other on the gel, the gel extraction might have included a mixture of both products during the extraction process for each product. This might explain the finding of only full-length copies sequence results for the insertion transcript. These results highlighted to the need for a separate TA cloning for the top gel

extracted band (insertion transcript) to isolate this transcript from the full-length transcripts (as discussed below).

**2.7.1.4 TA cloning for top band of exons 7-10 PCR product of CHRNA7:**

For characterising the upper band detected for the exons 7-10 PCR product, TA cloning of the gel extracted upper band DNA followed. Then PCR screening was used prior to miniprep plasmid preparation with most of the clones showing products of transcript with insertion (around 600 bp size) (Figure 2-24).
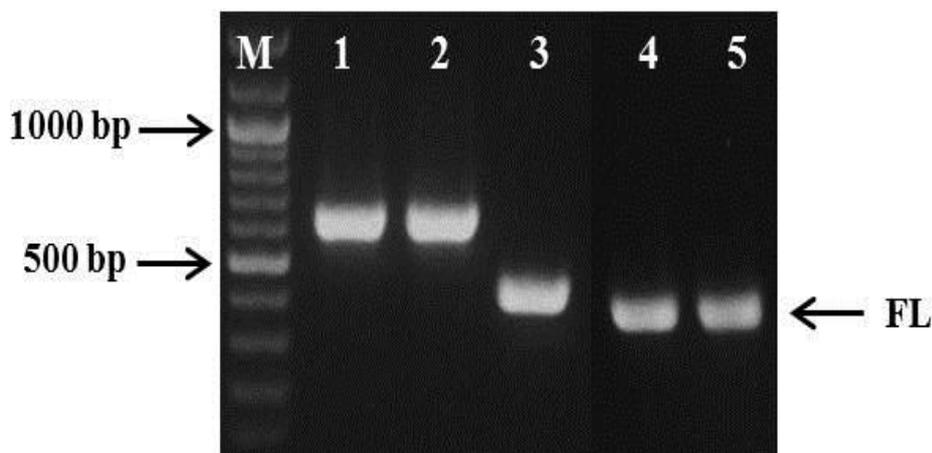


**Figure 2-24: Agarose gel for PCR screening for exon 7-10 top band products.**
After using TA cloning for the upper band products amplified using exons 7-10 primer set in A549 cells, PCR screening using the same primer set was carried out (lanes **1-5**). Based on the band sizes from the gel picture, lanes **1** & **2** showed transcript with insertion (around 600 bp size), lane **3** showed transcript slightly larger than the full length, lanes **4** & **5** showed clones match the full-length transcript of exons 7-10 (399 bp size). M: 100 bp DNA ladder. Note: the figure is made from different gel images.

These results showed the presence of full-length transcripts within the gel extracted mixture that explains the sequencing results compatible with full-length transcripts.

These samples were sequenced and the results are shown in (Table 2-14).

**Table 2-14: Summary for sequencing results of TA clones from exons 7-10 upper gel extraction.**

| Exons included | Product(s) size(s) | Sequencing result |
|---|---|---|
| **Full length + insertion** | 399 bp + additional 197 bp | Exons 7-10 + 197 bp from CHRFAM7A* intron 9 sequence (7,8,9,I9,10) |
| **Missing exon 9 + insertion** | 289 bp (missing exon9) + additional 197 bp | Exons 7,8,10 + 197 bp from CHRFAM7A* intron 9 sequence (7,8,I9,10) |
| **Full length (exons 7-10)** | 399 bp | Exons 7-10 |
| *: The sequencing results were aligned with intron 9 sequences of both genes (CHRNA7 and CHRFAM7A). The results showed that the insertion was part of intron 9 matching that from CHRFAM7A (further details will be discussed for this transcript later in this chapter and in chapter 3). | | |

From this point onward, this inserted sequence of intron 9 was termed exon 9b (E9b).

The total number of TA clones with inserts and the results of transcripts detected is summarised in Table 2-15. Surprisingly, the novel transcripts missing exon 9 could not be detected on TA cloning.

**Table 2-15: TA cloning summary for CHRNA7 transcripts.**

| PCR product | Clones tested | Transcripts detected | Number |
|---|---|---|---|
| **Exons 2-6** | 11 | full length (exons 2-6) | 6 |
| | | missing exon 4 | 4 |
| | | missing exons 3 & 4 | 1 |
| **Exons 1-10** | 9 | full length (exons1-10) | 5 |
| | | missing exon 4 | 3 |
| | | missing 1$^{st}$ 24 bp of exon 4 | 1 |
| **Exons 7-10** | 15 | full length (exons 7-10) | 8 |
| | | with exon 9b | 6 |
| | | with exon 9b & missing exon 9 | 1 |
| Note that no transcripts missing exon 9 were detected from all amplified sequences. This might highlight the effect of missing exon 9 on these transcripts when transformed into E.coli, by either causing toxic effect on the cell, inhibiting cell growth, reproduction, transcription, or other effects. | | | |

In summary, different methods were employed for the detection of CHRNA7 transcripts in A549 cells. These included the use of specific PCR primers, gel extraction, or TA cloning method. Some of the transcripts should be detected using one or more methods depending on that specific method (Table 2-16).

**Table 2-16: A summary of CHRNA7 transcripts detected in A549 cells.**

| Gene | Transcript | Detection method |
|------|-----------|------------------|
| | Full length (exons 1-10) | RT-PCR, gel extraction, TA cloning |
| | Missing exon 9 | Gel extraction |
| | Missing exon 4 | Gel extraction, TA cloning |
| **CHRNA7** | Missing exons 3&4 | TA cloning |
| | Missing exons 4&5 | Gel extraction |
| | Exon 4a (in short transcripts) | RT-PCR |

It is worth noting that full-length transcripts could be detected using all methods due to their abundant expression. Missing exon 9 transcripts were clearly detected by using PCR primers (shown as a second sequence peak that mixed up with the full-length one) or by using gel extraction method. However, none was detected using the TA cloning method. For missing exon 4 transcripts, they were detected both by gel extraction and TA cloning methods, but much less than missing exon 9 transcripts when comparing the gel extraction method.

### 2.7.2 Characterisation of the main CHRFAM7A transcripts in A549 cells:

Using primers specific for different parts of CHRFAM7A (Table 2-3), RT-PCR was used to amplify the target parts as illustrated in (Figure 2-10). The PCR products detected from A549 cells cDNA are shown in (Figure 2-25).
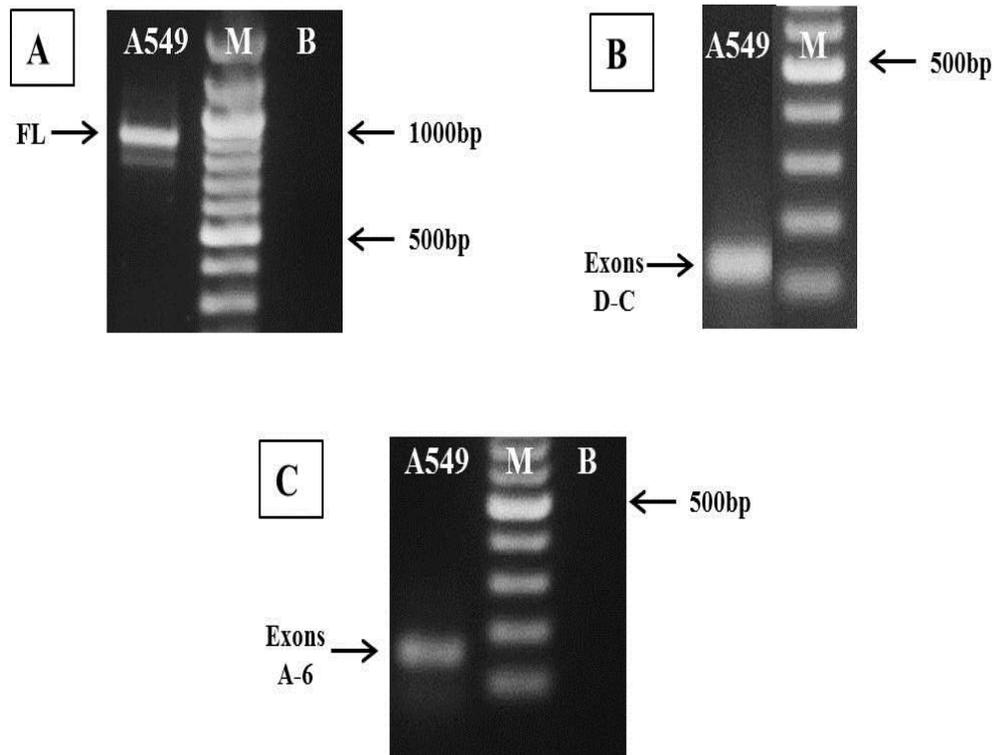


**Figure 2-25: Agarose gel electrophoresis for CHRFAM7A transcript.**
*A:* exons A-10 (expected size is 848 bp=FL) with one band smaller than FL noted, *B:* exons D-C (expected size is 124 bp), *C:* exons A-6 (expected size is 151 bp). M: 100 bp DNA ladder, B: blank.

Interestingly, the same novel transcript detected for CHRNA7 (missing exon 9) was detected in CHRFAM7A transcript.

In addition, a transcript missing bases 67-68 (TG) of exon 6 that was reported before (Gault et al., 1998) was detected from A549 cells (Figure 2-26).



Sequence 1: CAAAC**TG**AAGTTT (higher peaks)
Sequence 2: CAAACAAGTTTGG (lower peaks)
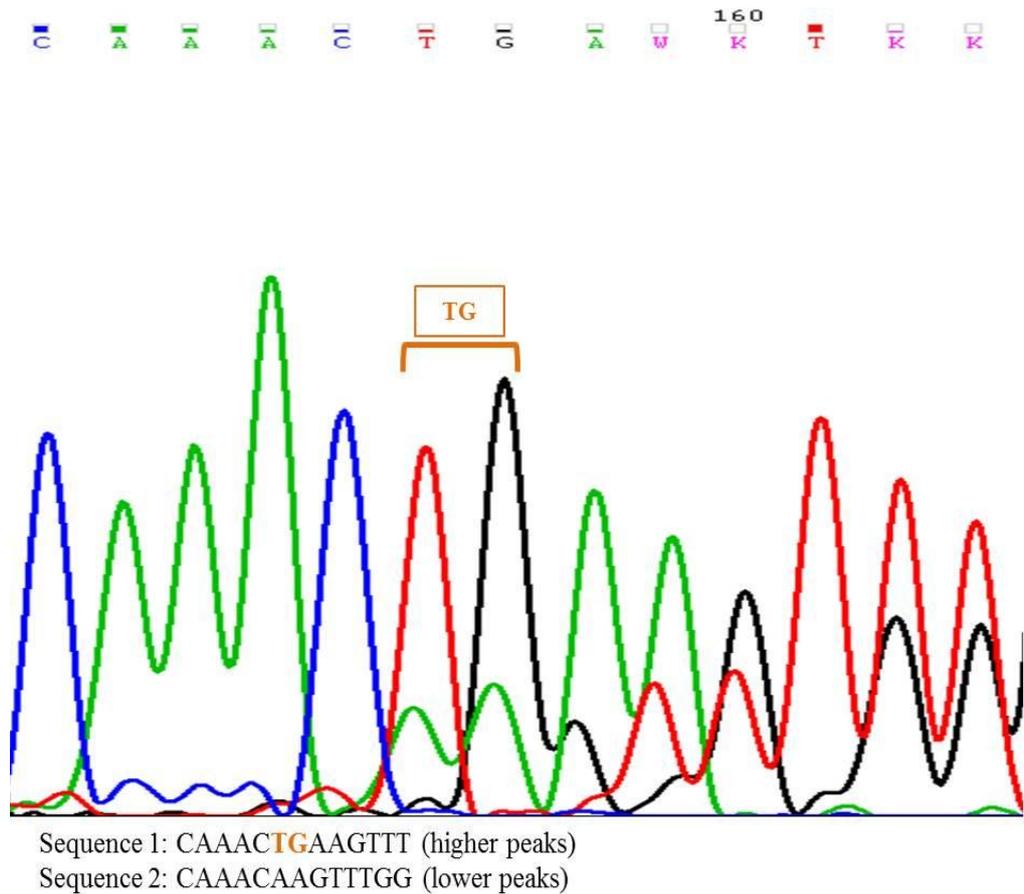
**Figure 2-26:CHRFAM7A transcript missing TG bases of exon 6.**
TG bases at sequence 67-68 of exon 6 are indicated. From that point, two sequences could be detected: a major transcript (with TG bases)-*Sequence 1*; and a minor transcript (without TG bases)-*Sequence 2*. Base colour reference: A=green, T=red, G=black, C=blue.

However, regarding the common exons between CHRNA7 and CHRFAM7A genes, exons 7-10 part was amplified earlier (results shown in Figure 2-14). These results showed the expression of a transcript with an insertion (larger than the size of exons 7-10 (Figure 2-27).
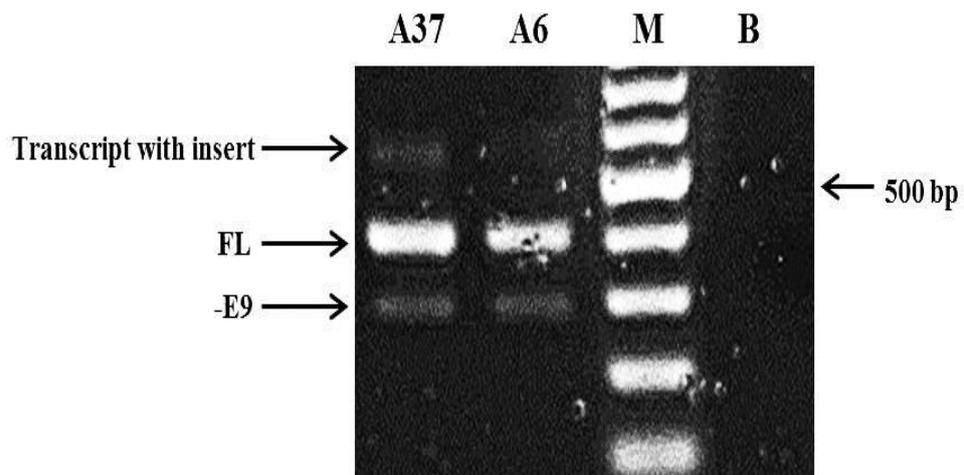


**Figure 2-27: Agarose gel electrophoresis for CHRNA7/CHRFAM7A transcript E7-10 in A549 cells.**
Lanes A37 & A6: cDNA samples from A549 (expected size is 399 bp=FL). Two additional bands of larger (transcript with insert ~ 600 bp size) and smaller (matching to predicted sizes of transcripts with missing exon 9, -E9= 289 bp size) product sizes could be observed. M: 100 bp DNA ladder, B: blank.

### 2.7.2.1 Initial characterization of insert transcripts in A549 cells:

The larger product (with insertion) was targeted first using gel extraction and sequencing. However, the sequencing results came back with a mixed signal of full length and an additional sequence (of lower peaks height). We aimed for targeting this transcript using TA cloning of the gel extracted largest size band (Figure 2-27).

**2.7.2.1.1   TA cloning of the largest product from RT-PCR for exons 7-10:**

This TA cloning is different from that mentioned in section 2.7.1.3 because it was preceded by gel extraction of the exon 7-10 top band prior to cloning. PCR screening was used prior to plasmid-insert extraction according to which the clones were selected for extraction (Figure 2-28).
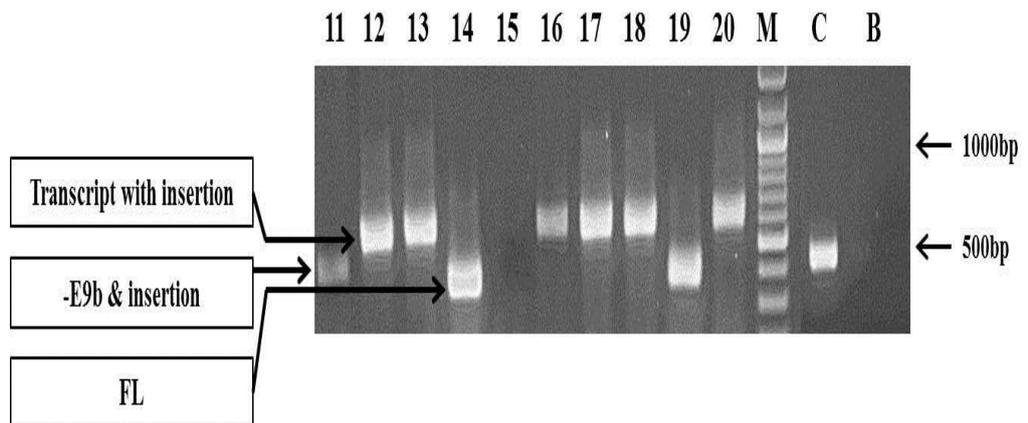


**Figure 2-28: Agarose gel for PCR screening for exon 9b inserted products.**
PCR screening of TA clones from exons 9b inserts PCR product of A549 (*samples 11-20*). The clones12-13, 16-18, & 20 match exons 7-10 transcript with insertion. The clones 14 & 19 match the full-length transcript (FL: exons 7-10). The clone 11 matches transcripts missing exon 9 and with insertion. C: exons 7-10 plasmid insert as control for FL. M: 100 bp DNA ladder, B: blank

The TA cloning results showed the presence of an inserted sequence (termed in this study as exon 9b) between exons 9 and 10 sequences (Figure 2-29).
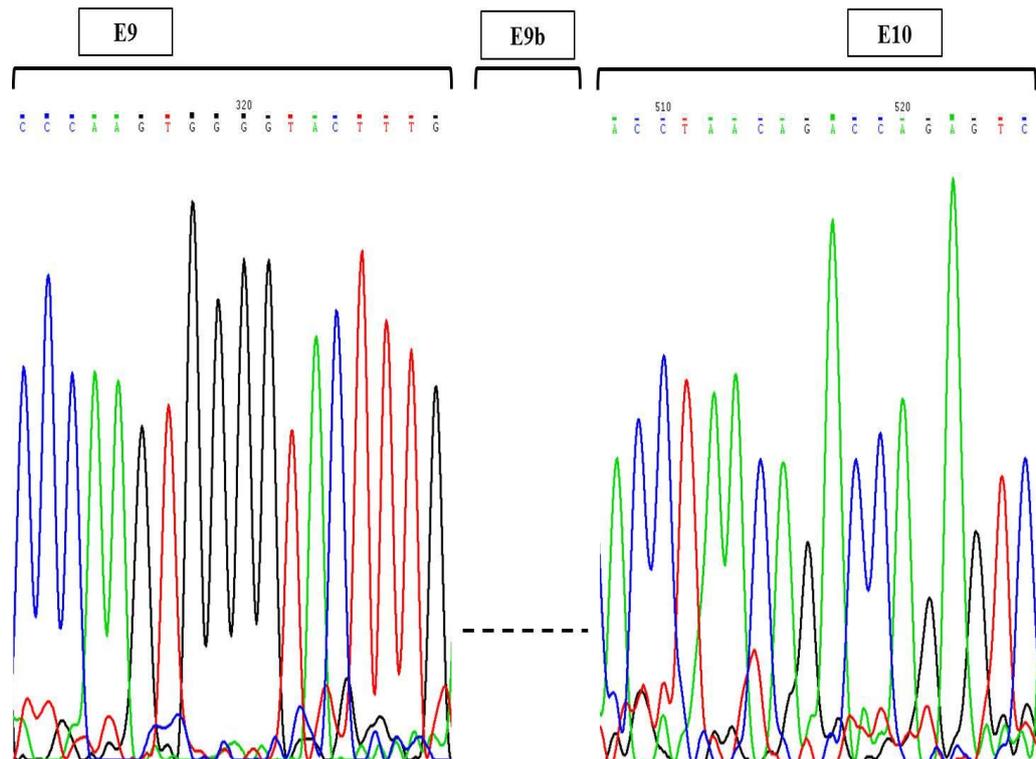
**Figure 2-29: Exon 9b inserted sequence position.**
Exon 9b sequence inserted between exons 9 and 10 sequences. Base colour reference:
A=green, T=red, G=black, C=blue.

Out of 9 clones, 7 were detected to match exon 9b insertion (one of them showed combined missing exon 9 with exon 9b insertion) and only two showed exons 7-10 matching the expected product from reference sequence of both genes (Figure 2-30).
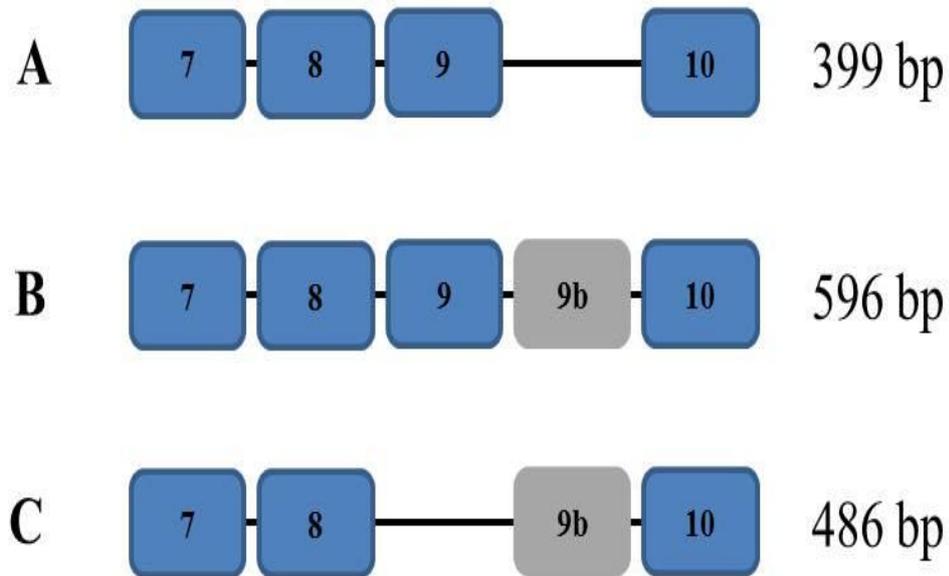
**Figure 2-30: exon 9b transcripts detected in A549.**
Blue boxes: constitutive exons, grey box: inserted exon 9b. The transcripts are *A:* (exons7-10), *B:* (exons 7-10, +E9b), *C:* (exons 7-10, +E9b, -E9). The expected product size (in bp) is shown on the right.

### 2.7.2.1.2  Sequencing the resulting products:

The inserted sequence showed matched to the part of intron 9 sequence from CHRNA7 and CHRFAM7A. However, only two bases were different within the exon 9b sequence from the intron 9 part of CHRNA7 but matching that of CHRFAM7A when aligned with the reference sequences of both genes (CHRNA7 accession number NM_00074 and CHRFAM7A accession number NM_139320.1, NCBI) (Figure 2-31).
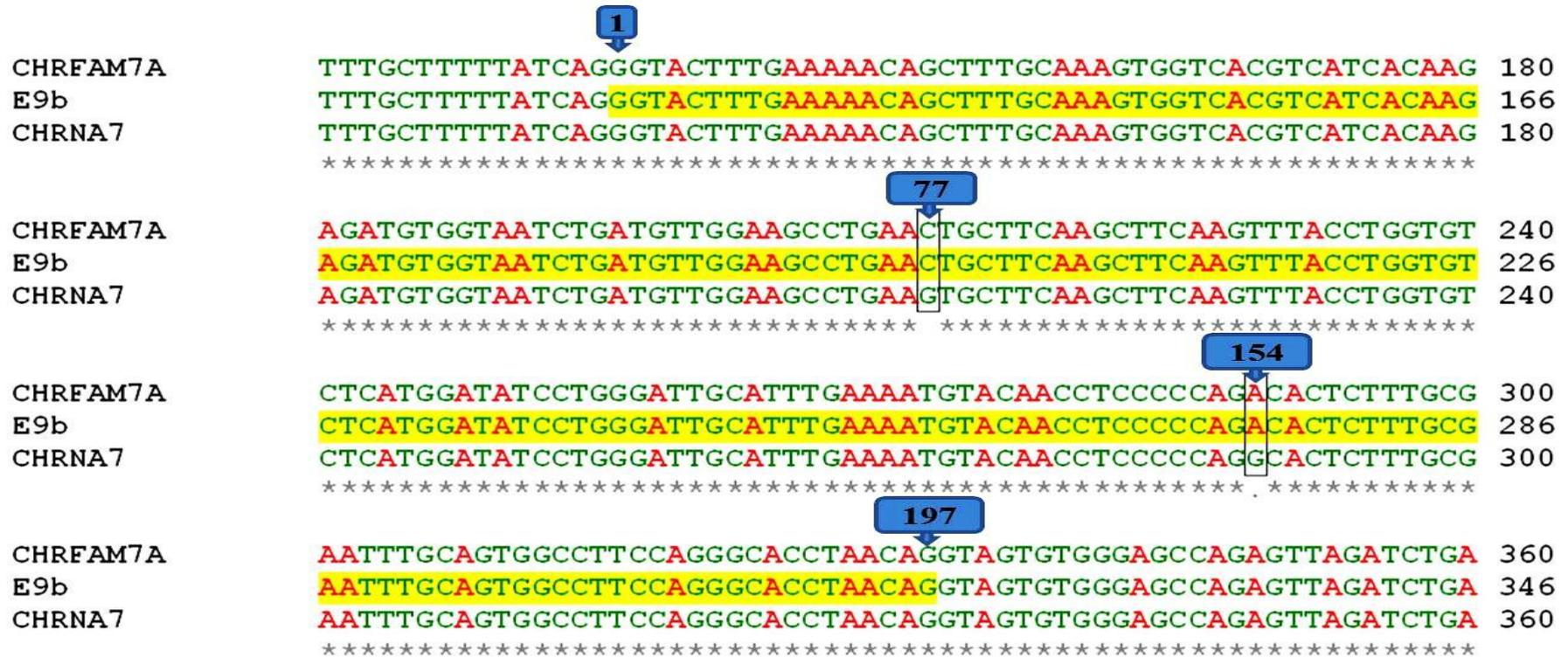
**Figure 2-31: Sequence alignment of exon 9b detected.**
The detected exon 9b sequence was aligned with intron 9 of CHRNA7 and CHRFAM7A genes. Two nucleotide bases were different between the two genes at positions 77 & 154 of exon 9b (enclosed within black boxes), and these two bases in the exon 9b transcripts (detected in our study) matched that of CHRFAM7A. The numbers above indicate the sequence from start of exon 9b (where 1 is start of exon 9b, 77 & 154 represent the two different positions between both genes, and 197 is the end of exon 9b).

At first, this exon 9b was expected to match that found in mice neuronal tissues within a conserved region between human and mice (Saragoza et al., 2003). In that study, CHRNA7 transcripts with inserted part of intron 9 (termed in the study as exon 9b) were detected. This region corresponds to the nucleotide sequences 2028-2574 of the human intron 9, while exon 9b detected in our study is located within nucleotide sequences 4025-4222 of human intron 9.

**2.7.2.2 Characterization of exon 9b in BE (2)-c cells transcripts:**

After confirming the expression of exon 9b in CHRFAM7A transcripts in A549 cells we tried to answer the following question: does this represent a tissue specific splicing process? To answer this question, the exons 7-10 were amplified from BE (2)-c cDNA using the same primer sets and PCR conditions to those used with A549. The results came back supporting those found in A549 cells (Figure 2-32).
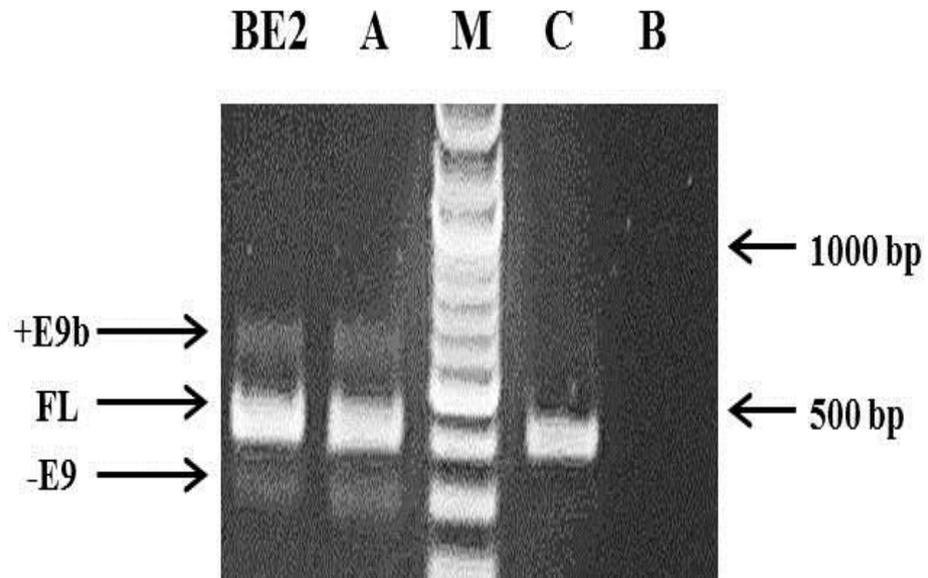
**Figure 2-32: Agarose gel electrophoresis for CHRNA7/CHRFAM7A transcript exons 7-10 in BE (2)-c cells.**
Lanes 1 & 2: cDNA samples from BE (2)-c (expected size is 399 bp=FL). A: cDNA from A549 tested as control for exons 7-10 transcripts. Two additional bands of larger (+E9b= 596 bp size) and smaller (-E9= 289 bp size) product sizes can be observed (matching to predicted sizes of transcripts with inserted exon 9b and missing exon 9, respectively). C: exons 7-10 plasmid-insert as control for FL. M: 100 bp DNA ladder, B: blank.

**2.7.2.2.1  TA cloning of the largest product from RT-PCR for exons 7-10:**

Similar to results shown early with A549, the results of PCR screening and restriction digestion (using EcoRI enzyme) in BE (2)-c cells showed products with the predicted sizes expected for exon 9b transcripts (full length (FL), with inserted exon 9b (+E9b) and missing exon 9 (-E9)) (Figure 2-33).
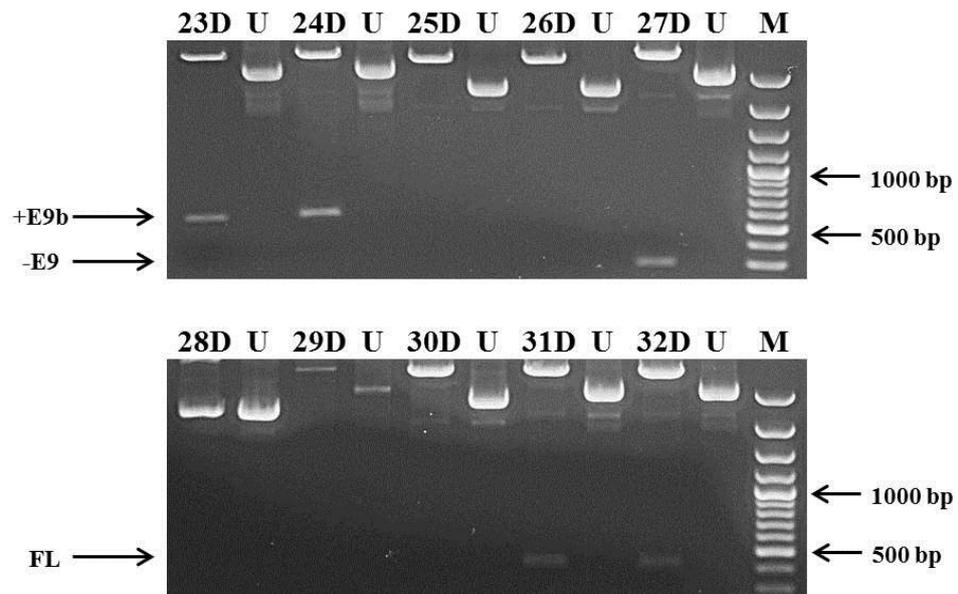
**Figure 2-33: Agarose gel for restriction digestion using EcoRI for exons 7-10 clones from BE (2)-c cells.**
Lanes 23D/U& 24D/U: inserted exon 9b products (+E9b), lane 27D/U: missing exon 9 transcript (-E9), lanes 31D/U & 32D/U: full-length exons 7-10 (FL). D: digested DNA using EcoRI, U: undigested. These bands were visible on the gel but sometimes were too faint to be captured on the image. M: 100bp DNA ladder.

Although showing exon 9b inserted transcripts, we could not detect transcripts with inserted exon 9b and deleted exon 9 in BE (2)-c cells (Like those found in A549 cells).

### 2.7.2.2.2  Sequencing the resulting products:

The sequencing results of the detected transcripts were similar to those detected in A549 (exon 9b sequence matched the corresponding part in CHRFAM7A) (Figure 2-31).

**2.7.2.3 Exon 9b sequence analysis:**

After detection of exon 9b sequences within CHRFAM7A transcripts in A549 and BE (2)-c cells, several bioinformatics tools were applied to the sequence to check for the splicing driving sequences within exon 9b and the flanking introns.

**2.7.2.3.1   Clustal W:**

This software was used for comparing specific sequences for nucleotide base differences. The sequencing results for CHRNA7, CHRFAM7A, and exon 9b transcripts detected were compared and shown to differ at positions 77 and 154 of exon 9b sequence. These two bases could differentiate the origin of transcripts with exon 9b (CHRNA7 or CHRFAM7A) (Table 2-17)

**Table 2-17: Exon 9b sequences detected at positions 77 and 154.**

| Sequence | Position 77 of exon 9b | Position 154 of exon 9b |
|---|---|---|
| CHRNA7 | G | G |
| CHRFAM7A | C | A |
| Exon 9b transcripts in A549 | C | A |
| Exon 9b transcripts in BE (2)-c | C | A |

The sequence of exon 9b detected in A549 transcripts matched that of corresponding intron 9 part of CHRFAM7A. Only two nucleotide bases were different between the two genes at positions 77 & 154 of exon 9b. This lead to concluding that exon 9b transcripts in A549 cells produced by CHRFAM7A.

**2.7.2.3.2   Exon scan:**

This software was used to predict the exonic sequences within a given sequence. The software requires the introduction of whole or part of a genomic

sequence to predict the presence of exons within a given sequence that should be flanked by 20 and 60 nucleotide bases sequences upstream and downstream of the exon tested, respectively. To test for exon 9b, sequences from CHRNA7 (NM_000746) and CHRFAM7A (NM_139320.1) were analysed. The first set of data included intron 8, exon 9 and intron 9 while the second set of data included only intron 9. The results for both data sets showed that exon 9b was not detected as an exon. This means that exon 9b is not a constitutive exon.

### 2.7.2.3.3  Splice site prediction:

For predicting the splice sites, *BDGP* software was used to test for the strength of the splice signal (Figure 2-34).



**Donor site predictions for 128.243.253.116.8034.0 :**

| Start | End | Score | Exon | Intron |
|-------|-----|-------|------|--------|
| 261 | 275 | 0.78 | ctaacag$gt$agtgtg | |

**Acceptor site predictions for 128.243.253.116.8034.0 :**

| Start | End | Score | Intron | Exon |
|-------|-----|-------|--------|------|
| 50 | 90 | 0.64 | ttggtgtttgcttttttatc$ag$ggtactttgaaaaacagctt | |

**Figure 2-34: Splice site prediction test.**
The strength of the splice sites flanking exon 9b sequence was checked using BDGP software program. Start/ end: refers to the start/end of splice site detected within the sequence tested. The score represents the strength of splice site. Score cut-off =0.4, the exon/intron boundary is shown in big letters.

Given that the threshold value used by the software was 0.4 for both splice sites, it seems that the splice sites flanking exon 9b are much higher than that of the threshold values. When using *Human splice finder* software, similar results were detected. The results of these two software programs for predicting potential splice sites are based on testing sequences flanked by AG and GT intronic sequences. This is a different analysis method from that used in Exon scan software program where it utilises additional ESEs, ESSs and give an overall score for a given sequence to predict a defined exon.

In summary, the search for the main transcripts of CHRFAM7A was carried out during this study. The main aim was to characterise the main transcripts expressed by A549 cells, and only RT-PCR and gel extraction methods were employed for this purpose. The results showed that A549 cells express two major transcripts for this gene: full-length and missing exon 9 (similar to that expressed by CHRNA7) (Table 2-18).

**Table 2-18: A summary for CHRFAM7A transcripts detected in A549 cells.**

| Gene | Transcript | Detection method |
|---|---|---|
| **CHRFAM7A** | exons A-10 | RT-PCR |
| | Missing exon 9* | gel extraction |
| | Exons D-C | RT-PCR |
| | Missing GT 67-68 of exon 6 | Gel extraction |
| | Exon 9b* | Gel extraction & TA cloning |
| **\*:** indicates novel transcripts detected. | | |

### 2.7.3   CHRNA7 and CHRFAM7A transcripts detected in other cell lines:

The main transcripts of CHRNA7 and CHRFAM7A were tested in other cell lines (mentioned in section 2.6.1). These transcripts included the full length, as well as the two novel transcripts detected in this study (missing exon 9 and exon 9b).

First, the main transcripts of these genes were tested and compared at the same time for relative expression in A549 and BEAS2B cells. The expression of CHRFAM7A transcripts was higher than the expression of CHRNA7 transcripts in both cell lines, although much less of the latter transcripts could be detected in BEAS2B cells (Figure 2-35).
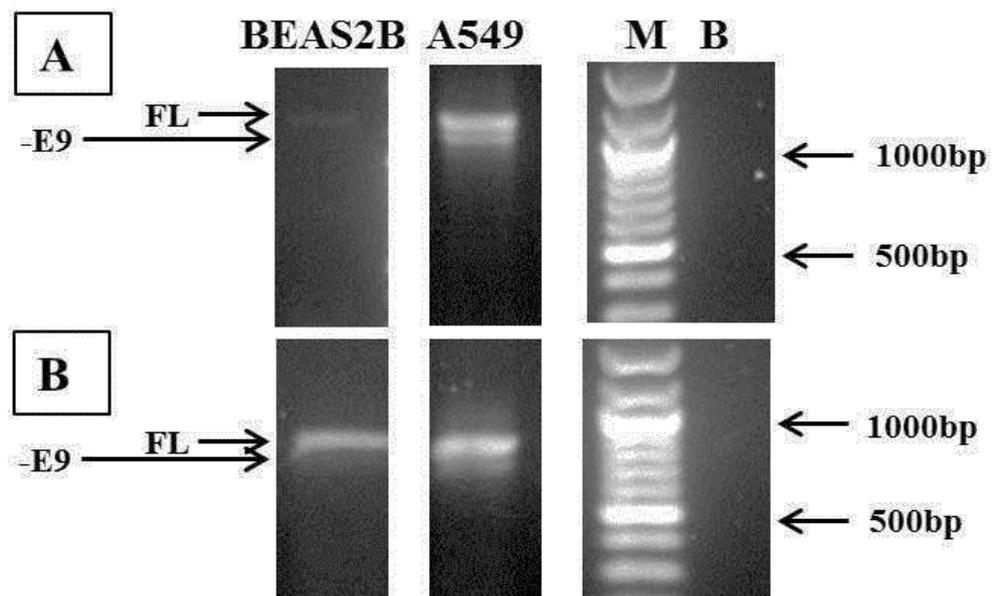


**Figure 2-35: Comparison of the relative expression of CHRNA7 and CHRFAM7A by A549 and BEAS2B cells.**
*A:* CHRNA7 main transcripts. *B:* CHRFAM7A main transcripts. Cells tested A549 and BEAS2B. FL: Full-length transcript (referred here for CHRNA7 exons 1-10; or for CHRFAM7A exons A-10), -E9=missing exon 9. Product sizes were FL= 1173 bp & -E9= 1063bp for CHRNA7, and FL= 848bp & -E9=738bp for CHRFAM7A. These bands were visible on the gel but were too faint to be captured on the image.  M: 100 bp DNA ladder, B: blank.

This comparison was also repeated for peripheral blood mononuclear cells (PBMC), which again showed a relatively higher expression of CHRFAM7A, than CHRNA7 transcripts (Figure 2-36).
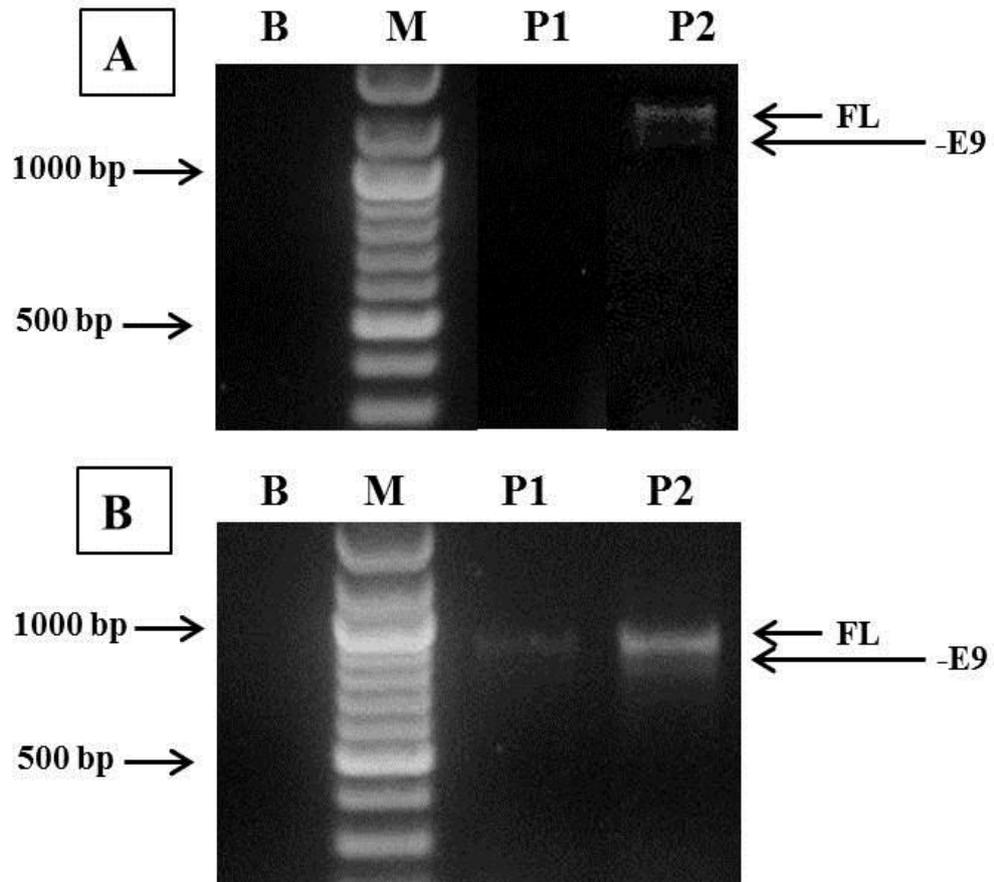


**Figure 2-36: Comparison of the relative expression of CHRNA7 and CHRFAM7A by PBMC cells.**
*A:* CHRNA7 main transcripts. *B:* CHRFAM7A main transcripts. P1-2: PBMC cDNA preps, FL: full-length transcript (referred here for CHRNA7 exons 1-10; or for CHRFAM7A exons A-10),   -E9=missing exon 9. Product sizes were FL= 1173 bp & -E9= 1063bp for CHRNA7, and FL= 848bp & -E9=738bp for CHRFAM7A.

P1 PBMC CHRNA7 transcripts were visible on the gel, but were too faint to be captured by the image. M: 100 bp DNA ladder, B: blank.

Several cDNA preps of PBMC were compared with A549 expression of the

two genes (Figure 2-37).
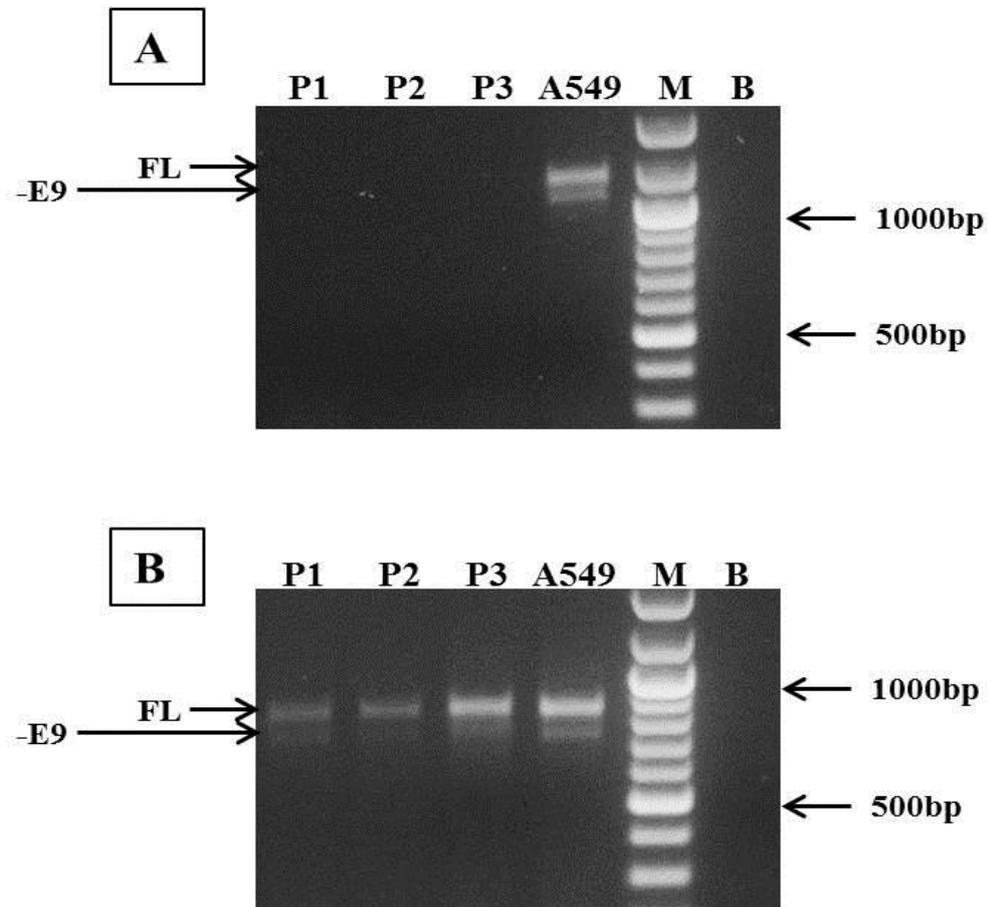


**Figure 2-37: Comparison of the relative expression of CHRNA7 and CHRFAM7A by PBMC and A549 cells.**
*A:* CHRNA7 main transcripts. *B:* CHRFAM7A main transcripts. P1-3: PBMC cDNA

preps, A549: A549 cDNA prep, FL: full-length transcript (referred here for CHRNA7

exons 1-10; or for CHRFAM7A exons A-10),    -E9=missing exon 9. The CHRNA7

transcripts were visible on the gel, but were too faint to be captured by the image. M:

100 bp DNA ladder, B: blank.

The same comparison was repeated using BE (2)-c cells cDNA. However, the results in these cells were the opposite of that expressed by the other cell lines used showing more expression of CHRNA7 than the CHRFAM7A (Figure 2-38 and Figure 2-39).
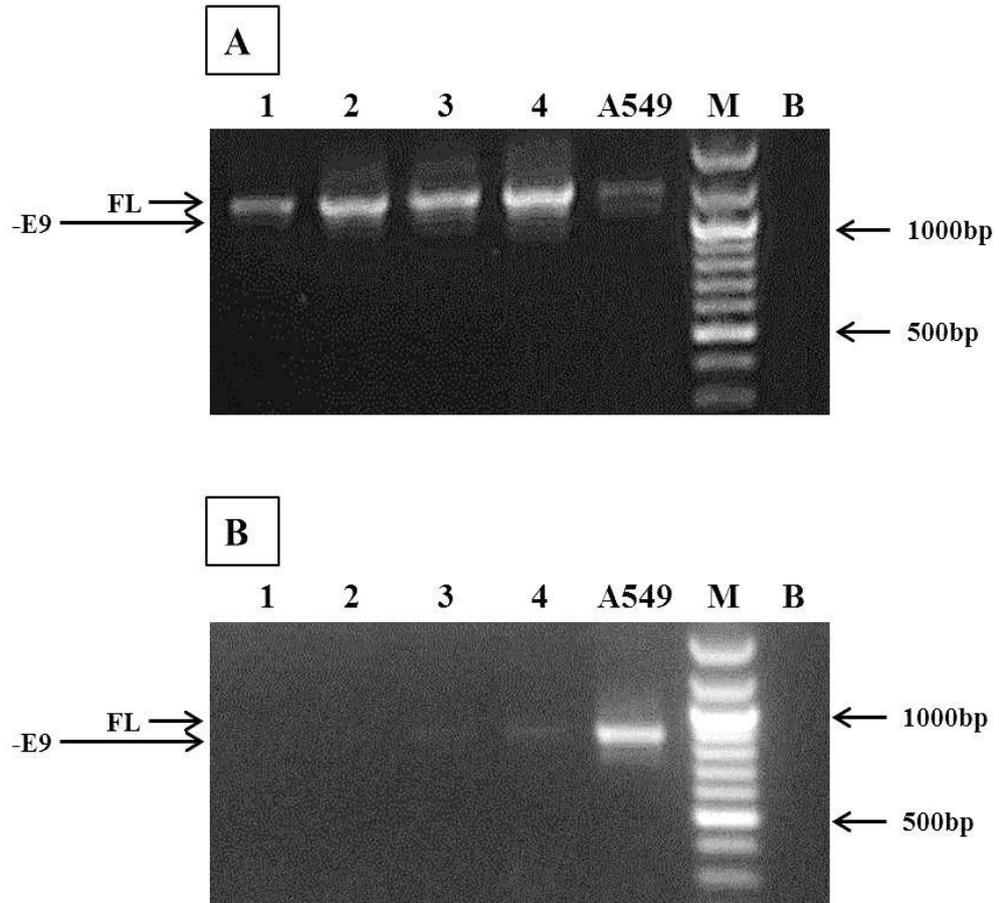


**Figure 2-38: Comparison of the relative expression of CHRNA7 and CHRFAM7A by BE (2)-c and A549 cells.**
*A:* CHRNA7 main transcripts. *B:* CHRFAM7A main transcripts. 1-4: BE (2)-c cDNA preps, A549: A549 cDNA prep, FL: full-length transcript (CHRNA7=exons 1-10; CHRFAM7A=exons A-10),  -E9=missing exon 9. The CHRFAM7A transcripts were visible on the gel, but were too faint to be captured by the image. The opposite relation between the two genes transcripts is evident on this gel picture, showing more CHRFAM7A in A549 cells, while more CHRNA7 in be (2)-c cells.  M: 100 bp DNA ladder, B: blank.
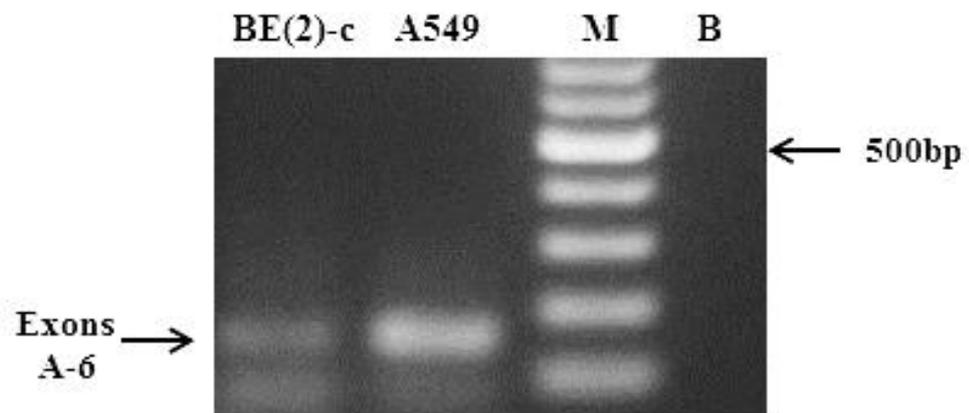
**Figure 2-39: Comparison of the relative expression of short fragment of CHRFAM7A between BE (2)-c and A549 cells.**
BE (2)-c and A549 cDNA preps were tested to confirm the expression of exons A-6

part of CHRFAM7 of 140bp size (arrowed). M: 100 bp DNA ladder, B: blank.

For three cell lines (A549, BEAS2B and PBMC), the main transcripts of CHRNA7 and CHRFAM7A genes were characterized by RT-PCR and gel extraction (Table 2-19). Interestingly, all of the three cells expressed the full lengths and missing exon 9 transcripts of both genes. Furthermore, the exon 9b transcripts could not be detected in PBMC cells, while -TG transcripts could not be detected in BEAS2B cells.

**Table 2-19: A summary for the main transcripts of both genes in three main cells tested.**

| Gene | Transcript | Cells | | |
|---|---|---|---|---|
| | - | A549 | BEAS2B | PBMC |
| **CHRNA7** | FL | ✓ | ✓ | ✓ |
| | -E9* | ✓ | ✓ | ✓ |
| | E9b* | ✓ | ✓ | X |
| **CHRFAM7A** | FL | ✓ | ✓ | ✓ |
| | -E9* | ✓ | ✓ | ✓ |
| | -TG | ✓ | X | ✓ |

Note: FL=full length (for CHRNA7=exons 1-10; for CHRFAM7A=exons A-10),

-E9=missing exon 9, E9b=with exon 9b, -TG=missing bases 67-68 of exon6.

✓ : detected transcripts, X: non-detected transcripts

**\*:** indicates novel transcripts detected

### 2.7.4    Quantitative Real-time PCR results:

For estimating the relative percent expression of full length and missing exon 9 CHRNA7 transcripts by A549 cells, QRT-PCR was used. For this purpose, the primers were designed in a way to avoid amplifying the common sequence between CHRNA7 and CHRFAM7A (exons 5-10). Thus, the forward primer was used in common for the detection of both transcripts of CHRNA7 (full-length and missing exon 9) (forward primer on exon4-Table 2-2) while a separate reverse primer for each transcript was designed on exon-exon junctions (Table 2-6, Figure 2-12). Thus, the two amplified parts of CHRNA7 would be one representing the full-length transcripts (amplifying exons 4-exon junction 8/9 part) and another representing missing exon 9 transcripts (amplifying exons 4-exon junction 8/10).

### 2.7.4.1 QRT-PCR optimization:

First, the primer sets were optimized using RT-PCR. For this purpose, SYBR green brilliant III was tested first on a Verity block PCR machine using a temperature gradient (Figure 2-40).
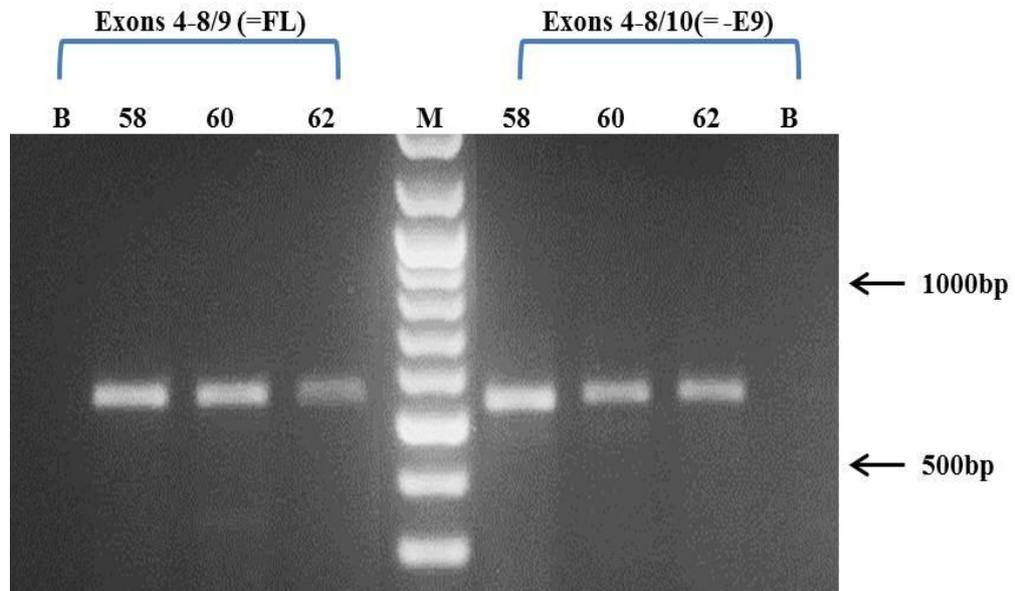


**Figure 2-40: RT-PCR using QRT-PCR primers and mastermix.**
A549 cDNA preperations were tested using QRT-PCR primers for detection of full length (FL, expected product size=584bp) and missing exon 9 (-E9, expected product size=581bp) transcripts. It is worth noting that the final temperature used for the QRT-PCR was 60º. The corresponding products from both primer sets were sequenced and the results confirmed the predicted sequences. M: 100 bp DNA ladder, B: blank.

The next step included testing these primers on A549 cDNA using QRT-PCR (Figure 2-41).
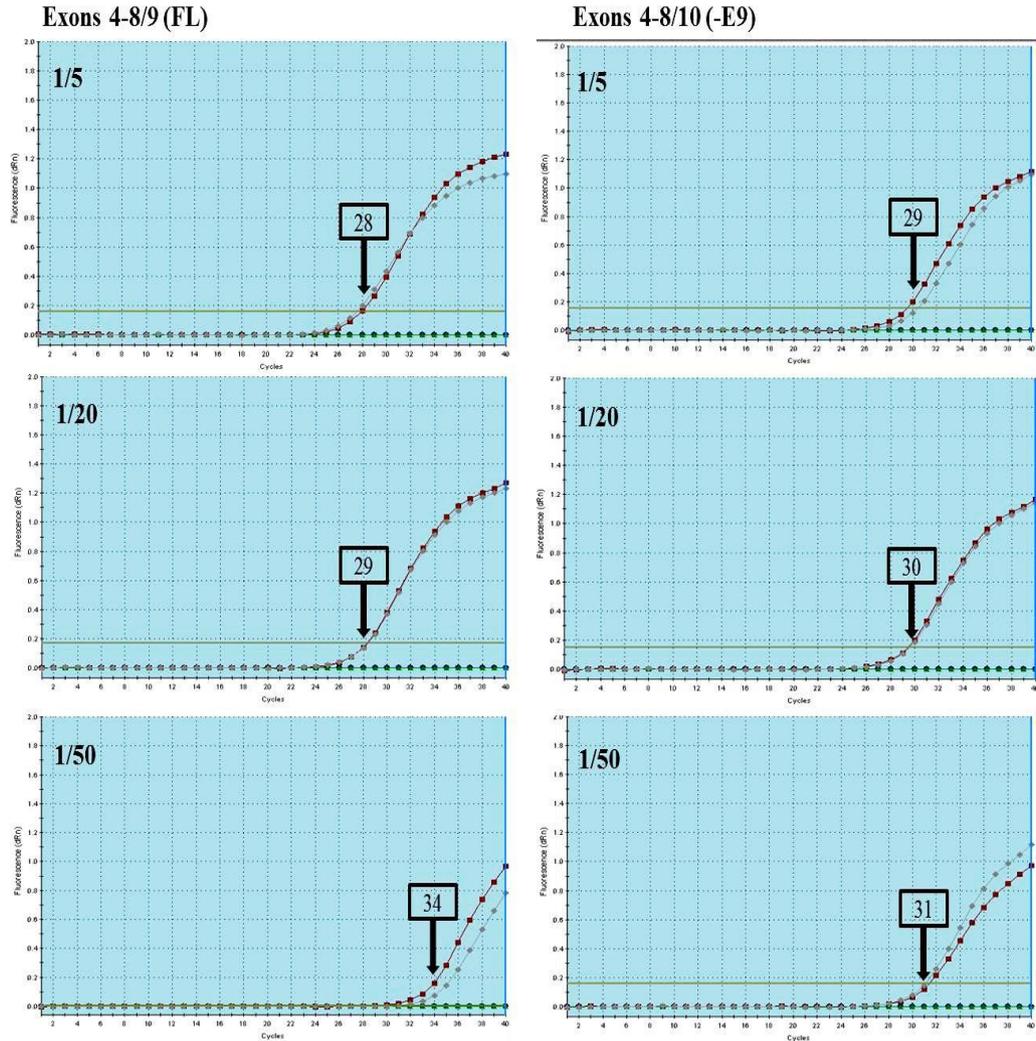


**Figure 2-41: Amplification plots for QRT-PCR.**
*Left:* Exons 4-8/9 (full length=FL) and *Right:* exons 4-8/10 (missing exon 9= -E9) transcripts of CHRNA7 were detected using QRT-PCR. The amplification plots show the delay in amplification Ct values associated with decreasing the amount of template starting the reaction (with lower Ct values with 1/5 cDNA dilution and higher Ct values with 1/50 cDNA dilution). The dilutions of the cDNA used were 1/5, 1/20 and 1/50 in duplicates (indicated on top left corner of each plot) and the Ct values for are indicated (a box and arrow). These results showed that with increasing the dilution of the tested cDNA samples, Ct values were higher, indicating that less template were available at the start of the reaction.

In addition, the dissociation curves for the above shown QRT-PCR results indicated that multiple products could be amplified. This was later confirmed when the products were subjected to agarose gel electrophoresis following each PCR reaction, and an extra band below the main band of each amplified product was noticed just below the main product bands (for all the dilutions and all the duplicates) (Figure 2-42).
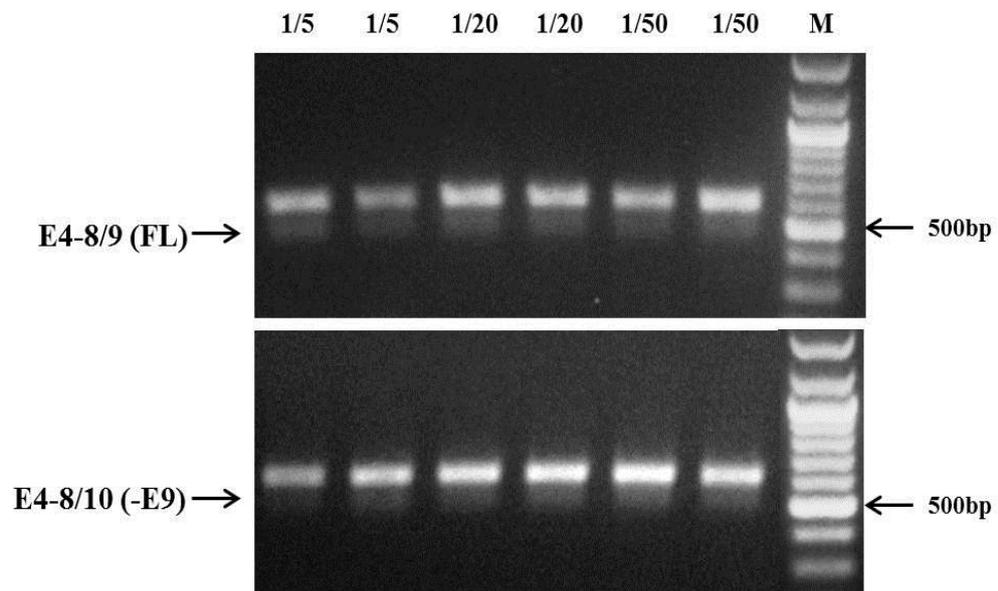


**Figure 2-42: Agarose gel electrophoresis for QRT-PCR products.**
cDNA dilutions (in duplicates) are indicated above each lane. FL: indicates exons 4-8/9 products, -E9: indicates missing exon 9 products. Lower band is noticed below the main band of all products. M: 100 bp DNA ladder, B: blank.

RT-PCR was repeated to check for the double banding pattern; however, the same results were obtained (Figure 2-43).
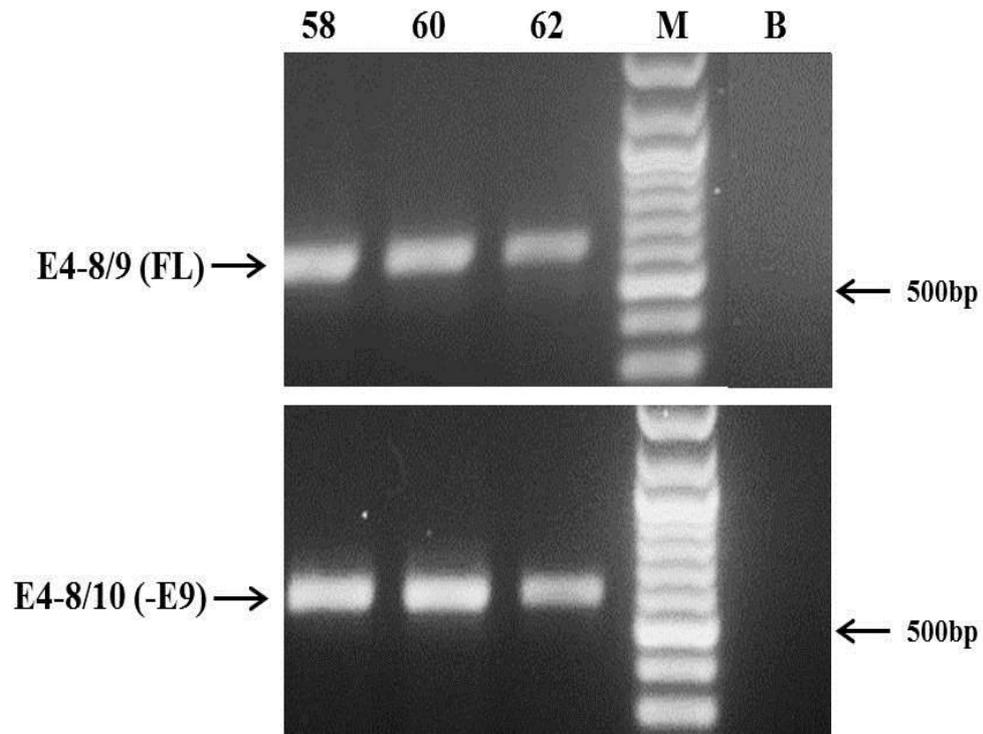


**Figure 2-43: Agarose gel electrophoresis for QRT-PCR products(using temperature gradiant).**
Temperature gradient RT-PCR was repeated for QRT-PCR primers and master mix using A549 cDNA. Temperatures used indicated above the lanes, FL: indicates exons 4-8/9 products, -E9: indicates missing exon 9 products. Lower band is noticed below the main band of all products. These bands were visible on the gel but were too faint to be captured on the image. M: 100 bp DNA ladder, B: blank.

Due to the presence of a second band of similar size when using QRT-PCR for amplifying both products (exons 4-8/9 (FL) and exons 4-8/10(-E9)), the sequence similarities between the reverse primers were checked for matching bases. First, both primers have a common sequence (the part of the primer on exon 8), this narrowed the comparison area to the 5` end of the primers (the parts on exon 9 or 10). However, four more bases were identical between the two primers (Figure 2-44).
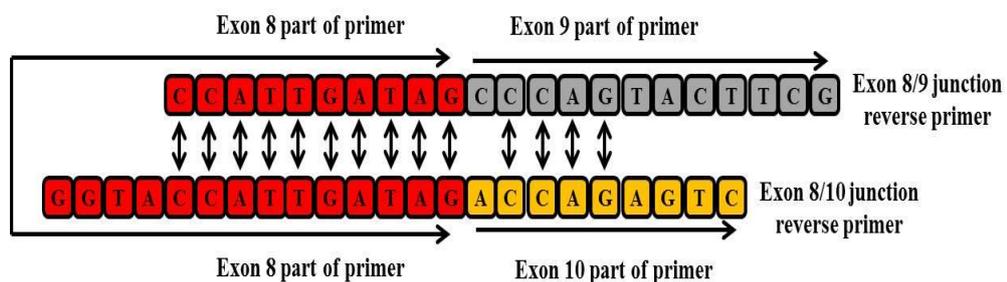


**Figure 2-44: QRT-PCR primers for CHRNA7.**
The primer nucleotides included from exon 8 (in red), exon 9 (in grey) or exon 10 (in yellow). The matching nucleotides between the two reverse primers are indicated by arrows (↕).

Due to the presence of 14 nucleotide bases matching between the two reverse primers, it was decided to use TA cloning for generating exons 4-8/9 (FL) and exons 4-8/10 (-E9) templates test for the primer specificity.

**2.7.4.2 TA cloning for exons 4-8/9 & 4-8/10 RT-PCR product of CHRNA7:**

Initially, the RT-PCR products amplified using primer sets for exons 4-8/9 and exons 4-8/10 were used as the target for TA cloning. The aim was to generate full length and missing exon 9 templates. The next step was to use these templates to test both primer sets for specificity of each set. Further confirmation of the primer specificity was done by using previously prepared clones containing exons 1-10 with or missing exon 4, respectively. These templates should confirm if the exon 4 specificity is to CHRNA7 and not to CHRFAM7A transcripts. As with other TA cloning used before, PCR screening was used after TA cloning to limit the extraction of the plasmid-insert DNA. This was done by using the primers used initially for RT-PCR amplification. This means that exons 4-8/9 primers were used for RT-PCR and then TA cloning were used for PCR screening and according to the band size difference clones were selected for plasmid-extraction and sequencing (the same was done for exons 4-8/10 primer sets) (Figure 2-45).
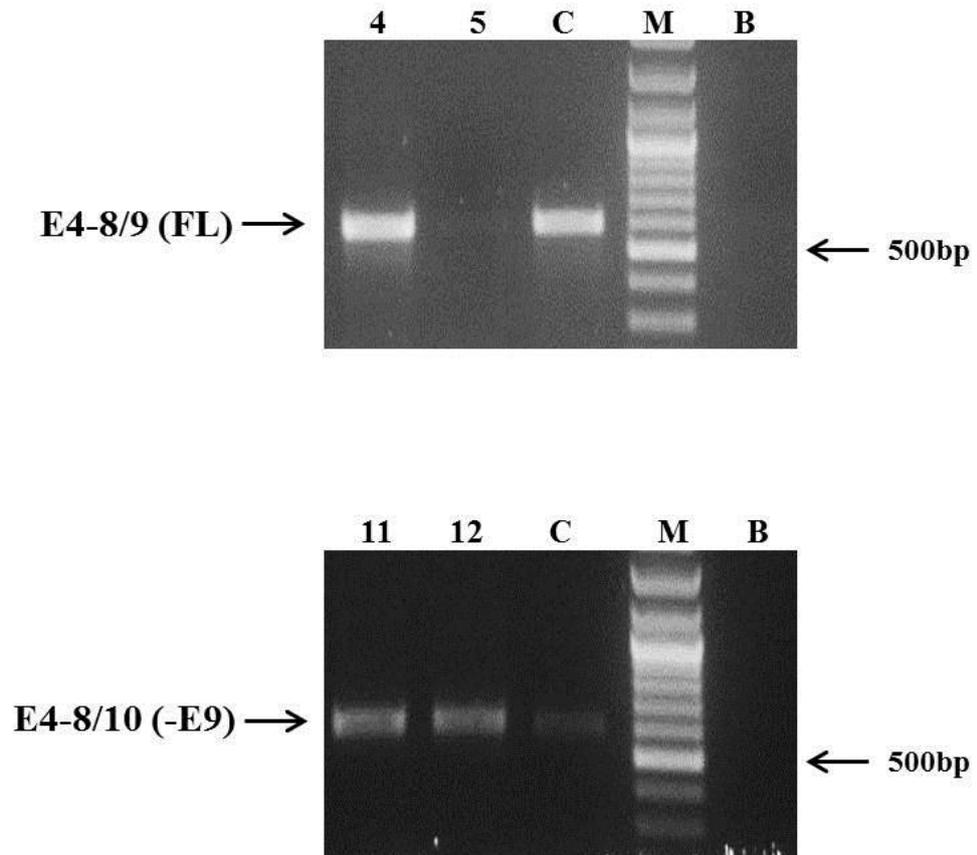
**Figure 2-45: Agarose gel for PCR screening.**
PCR screening of TA clones from exons 4-8/9 (*Top*) (with Clone 4 match the full-length transcript of exons 4-8/9 of 584 bp size) & 4-8/10 (*Bottom*) (with Clones 11 & 12 matching transcripts missing exon 9 of 581 bp size) PCR products of A549. M: 100 bp DNA ladder, C: controls (from gel extraction), B: blank.

A further screening step was followed (prior to sending the plasmid inserts for
sequencing) by using EcoRI for digestion of the plasmid from the insert (with
no recognition site for EcoRI in exons 4-8/9 or 4-8/10) (Figure 2-46).
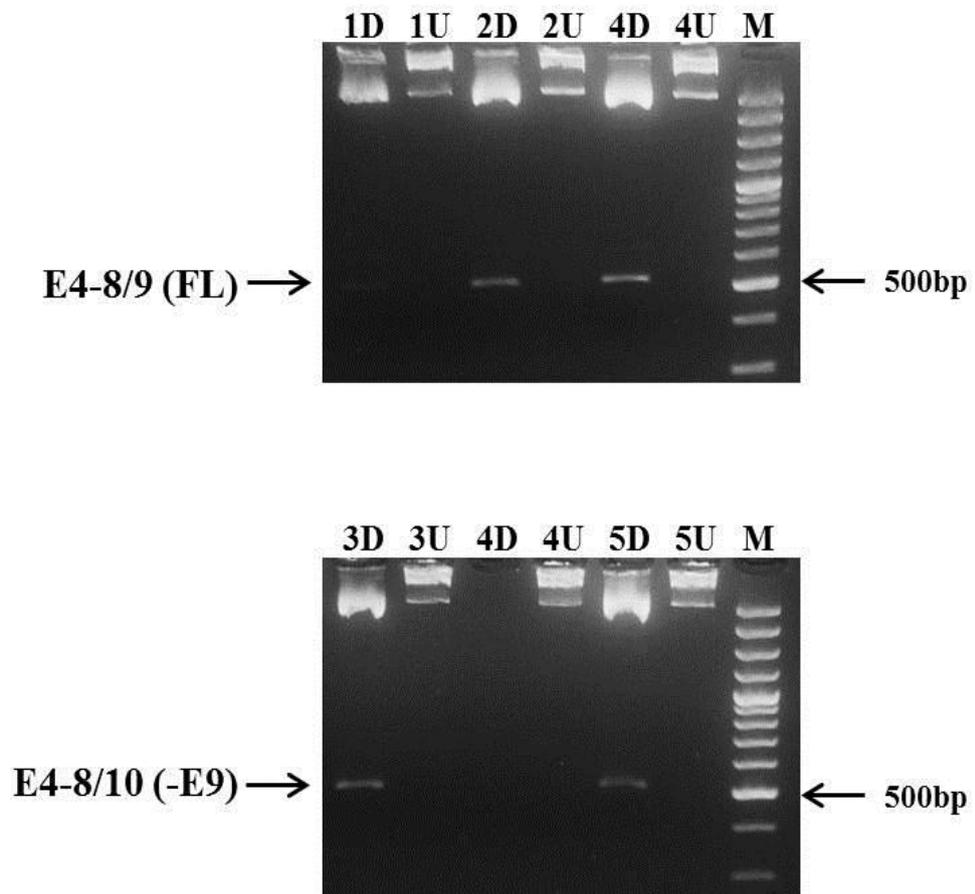


**Figure 2-46: Agarose gel for restriction digestion using EcoRI for exons 4-8/9 &
4-8/10 clones.**
*Top:* exons 4-8/9 clones digestion (with lanes 1, 2 & 4D matching uncut exons 4-8/9

product size=584 bp), *Bottom:* exons 4-8/10 clones digestion (with lanes 3 & 5D

matching uncut exons 4-8/10 product size=581 bp). D: digested DNA using EcoRI, U:

undigested, M: 100bp DNA ladder.

After extracting the plasmid-insert DNA, the samples were sent for sequencing
and matched the expected products: with sequence of exons 4-8/9 matched full
length CHRNA7 transcripts and the inserts of exons 4-8/10 matched missing
exon 9 CHRNA7 transcripts (Figure 2-47).
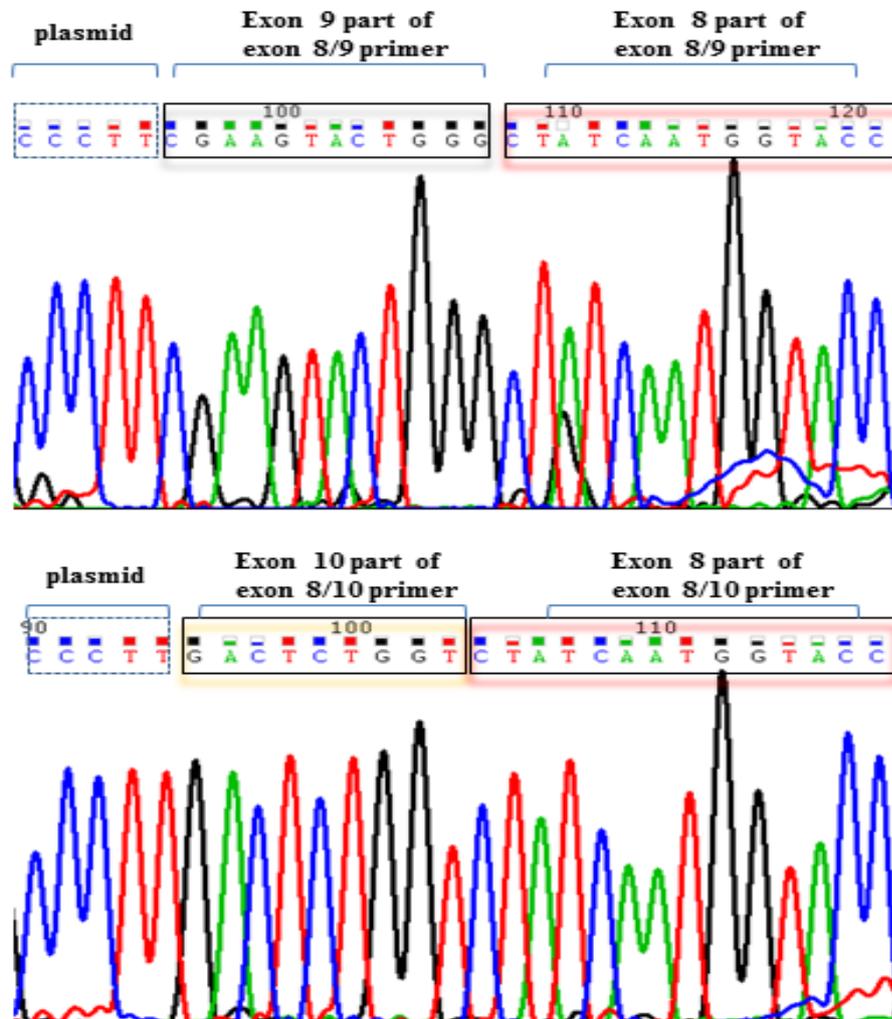


**Figure 2-47: DNA sequencing of extracted plasmid-insert for QRT-PCR primer products.**
*Upper panel:* exons 4-8/9 (full-length CHRNA7 transcript) sequencing results, *Lower
panel:* exons 4-8/10 (missing exon 9 CHRNA7 transcript) sequencing results. For
each sequencing result, three areas are shown: plasmid, exon 9 or 10 and then exon 8
of the reverse primer, respectively. Sequencing primer used was exon 4S primer
(Table 2-2). Base colour reference: A=green, T=red, G=black, C=blue.

### 2.7.4.3 Testing QRT-PCR primer specificity:

These DNA extracts were used as templates for further testing. By using QRT-PCR primer sets, PCR amplification was applied and both templates sets (full length and missing exon 9) were tested by both primer sets (full length and missing exon 9) to test for the primer specificity (Figure 2-48).
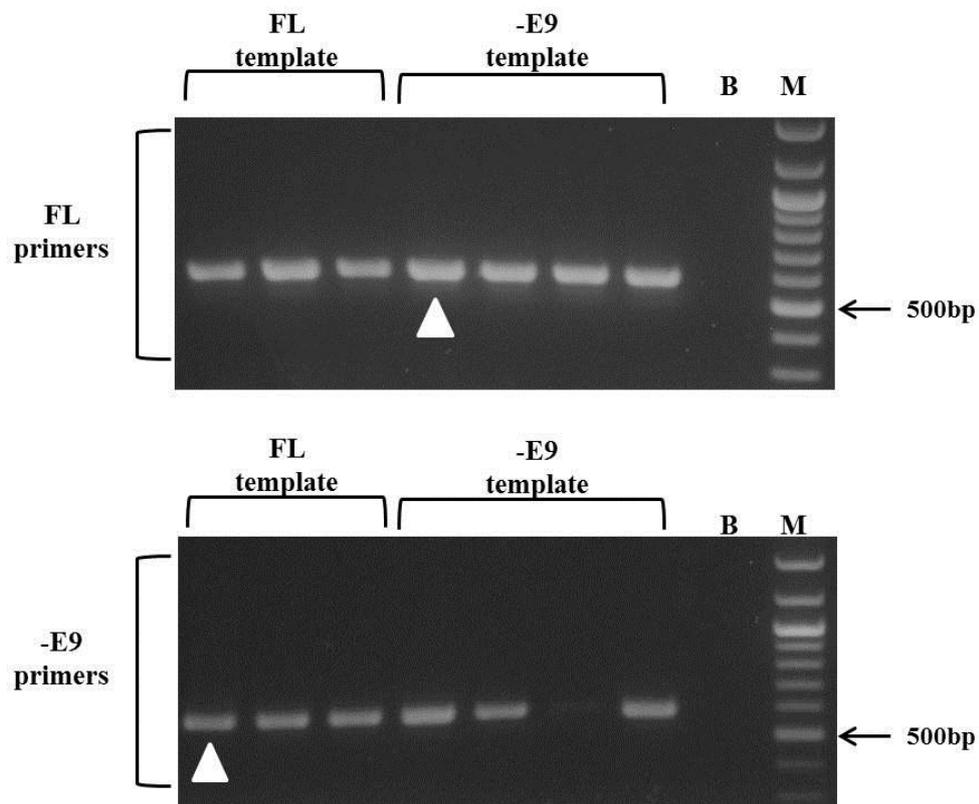


**Figure 2-48: Agarose gel for PCR of testing QRT-PCR primer specificity.**
PCR using QRT-PCR primers on A549 DNA clones representing full length (FL) and missing exon 9 (-E9) templates. *Upper panel:* FL primers used on both templates, *Lower panel:* -E9 primers used on both templates. Both primer sets amplified both template sets, and faint lower bands could be detected for most of the PCR products. White arrowheads indicate the products selected for the illustrated DNA sequence in the next figure (Figure 2-49). These are representing a missing exon 9 template amplified by using exons 4-8/9 (FL) primer set (on the upper panel) and a full-length template amplified by using the exons 4-8/10 (-E9) primer set (on the lower panel). M: 100 bp DNA ladder, B: blank.

This result indicates that the primers are non-specific and can bind to both templates leading to the non-specific nature of the reaction. The PCR products of the above reaction (Figure 2-48) were sent for sequencing, and the results showed sequence of all the products matching the expected sequence except the last part, represented by the sequence of the 5` end of the reverse primers (Figure 2-49).
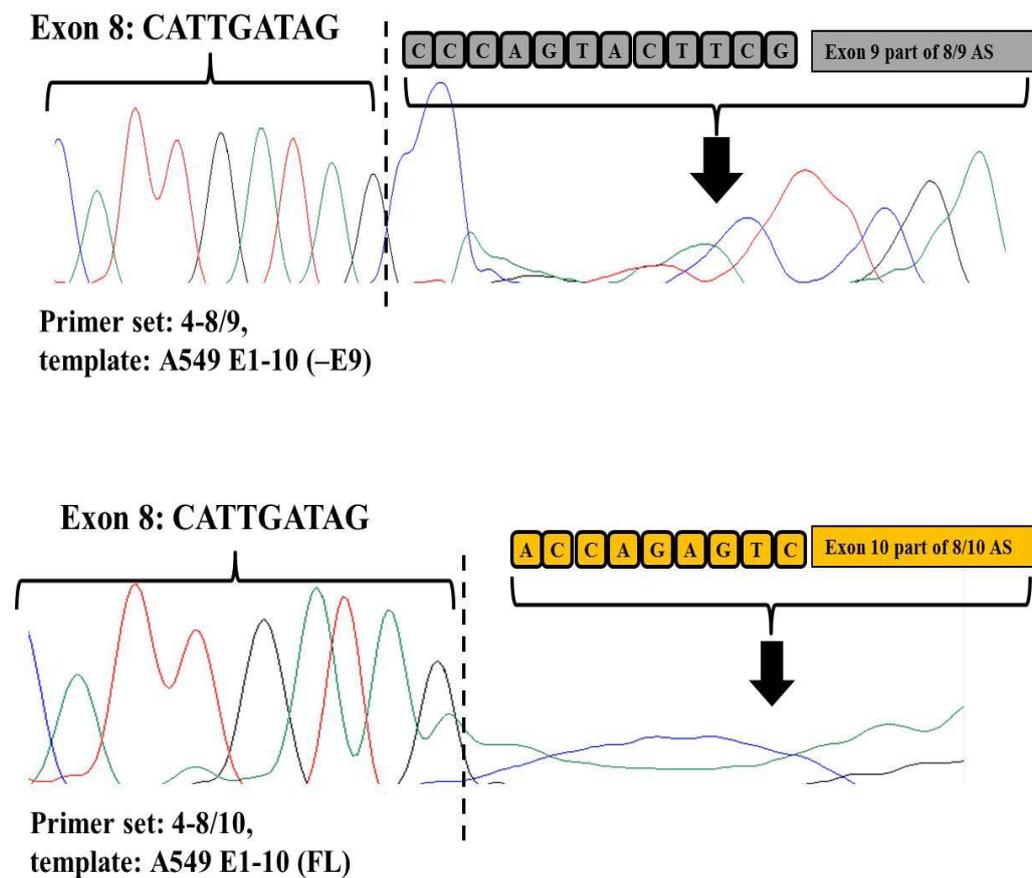




**Figure 2-49: DNA sequencing results for PCR products of QRT-PCR primer specificity testing.**
This figure shows the sequence of two of the PCR products shown on Figure 2-48. *Upper panel:* The missing exon 9 template amplified by full-length primers (indicating that the 5` end of this primer which represents exon 9 sequence was not annealing to anything (in the PCR)). *Lower panel:* The full-length template amplified by missing exon 9 primers (indicating that the 5` end of this primer which represents exon 10 sequence was not annealing to anything (in the PCR)).

To prove that the sense primer for QRT-PCR (exon 4S-Table 2-2) was specific to CHRNA7 transcripts, another PCR was performed on clones prepared from A549 cDNA by using exons 1-10 primer set (Table 2-2) and were proved by sequencing to be matching full length and missing exon 4, respectively. These template DNA samples were subjected to amplification by PCR using QRT-PCR primer sets for full length and missing exon 9 transcripts. The results showed that exon 4S primer is detecting CHRNA7 with exon 4, and could not detect those missing exon 4, indicating that this primer is specific to CHRNA7 transcripts with exon 4 (Figure 2-50).
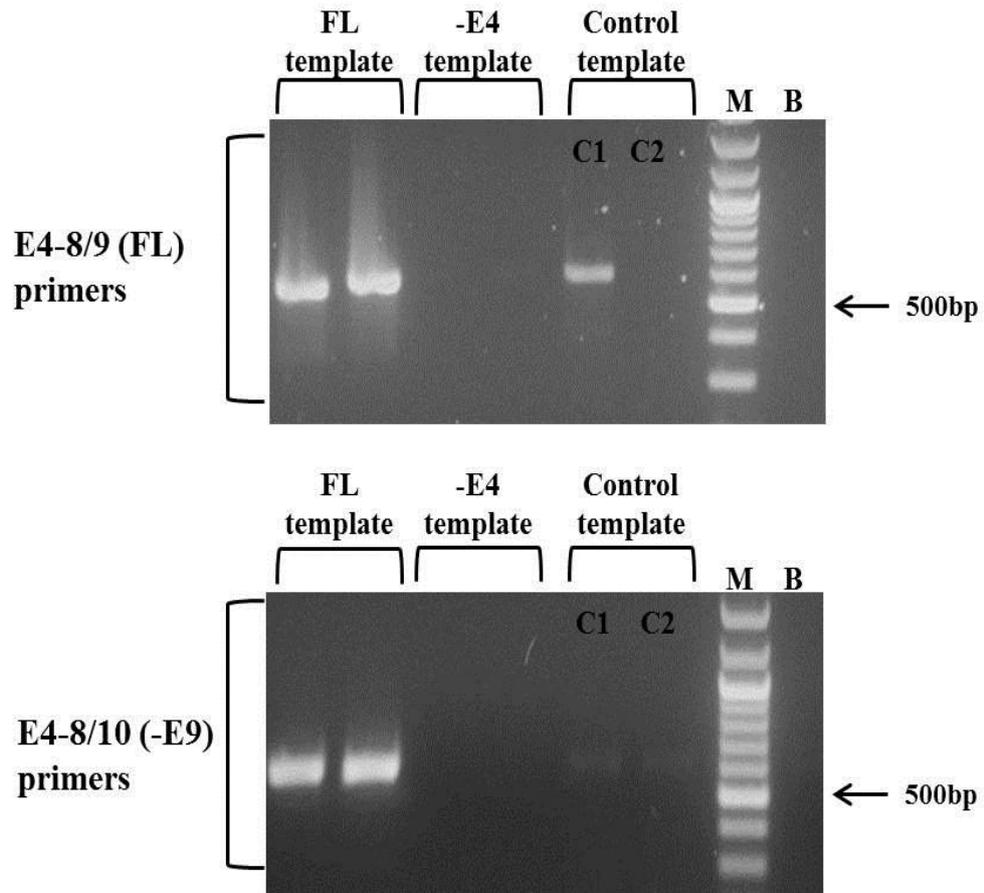
**Figure 2-50: Agarose gel for PCR products testing of exon 4S primer specificity.**
PCR using QRT-PCR primers on A549 DNA clones representing full length (FL) and
missing exon 4 (-E9) templates. *Upper panel:* FL primers used on both templates,
*Lower panel:* -E9 primers used on both templates. Only CHRNA7 templates with
exon 4 were amplified using both primer sets. The use of –E9 QRT-PCR primer set
(exons 4-8/10 primers) in addition to the FL ones aimed at testing both primer sets for
annealing when exon 4 was within (FL) or missing from (-E4) the template.  C1: A549
cDNA, C2: BEAS2B cDNA. M: 100 bp DNA ladder, B: blank.

In addition, the sequence of exon 4S primer was aligned with exons 1-10 of CHRNA7 and it matched sequence of exon 4. When exon 4S sequence was aligned with exons D-A of CHRFAM7A, it did not match to any sequence (Figure 2-51).

**Figure 2-51: Sequence alignment of exon 4S primer and CHRFAM7A exons D-A.**
1: with exon D, 2: with exon C, 3: with exon B, 4: with exon A. No complete
matching of exon 4S primer was detected on any of these exon sequences (asterisks
below the sequences indicates the matching between the nucleic acids aligned) .

**2.7.5   PCR semi-quantitative assays:**

Due to the difficulties encountered while attempting QRT-PCR amplification of CHRNA7 main transcripts (full-length and missing exon 9) an alternative semi-quantitative RT-PCR assay was used. This method depends on the use of primer sets that were used to amplify the main transcripts of both genes (exons 1-10 primers for CHRNA7 and exons A-10 for CHRFAM7A).  This method helped to estimate the relative percent expression of the main transcripts of these genes by A549 cells. This step was followed by quantifying the intensity of the representative bands present on gel electrophoresis using a densitometry method.

The target templates were amplified using RT-PCR (as described in section 2.6.4) and the results were then analysed using AlphaDigiDoc 1201 software (as described in section 2.6.11). For this purpose, A549 cDNA was used. Before the quantitation assay, the cDNA was tested using RT-PCR to check for CHRNA7 and CHRFAM7A expression. Once confirmed, this cDNA was used for all subsequent runs for quantitation of both genes transcripts'.

**2.7.5.1 PCR semi-quantitative assays for CHRNA7 transcripts:**

RT-PCR was used for amplifying CHRNA7 in A549 cells using different cycle number for amplifications (cycles 21-35). This helped narrow the cycle numbers used. For optimization purposes, the first run was performed on single samples. The results showed that the amplification starts at cycle 27 and reaches a plateau phase after cycle 33 (Figure 2-52).
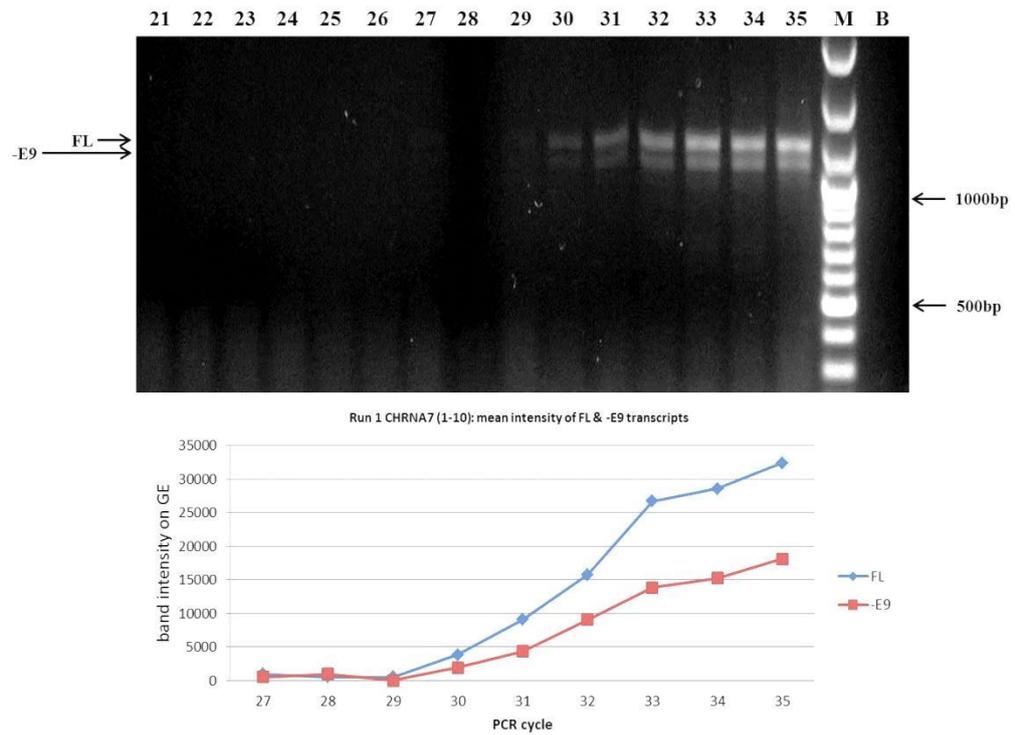
**Figure 2-52: Semi-quantitative assay optimisation for CHRNA7 transcripts.**
*Upper panel:* Agarose gel electrophoresis of RT-PCR applied on A549 cDNA. PCR sample tubes were removed from PCR block after cycles 21-35 (cycle number shown on top). *Lower panel:* AlphaDigiDoc analysis for PCR products that were visually detectable on the gel (for cycles 27-35) (using band intensity). M: 100 bp DNA ladder, B: blank.

From this point onward, the cycles selected were 27-35 (27 cycles which showed the first visible CHRNA7 transcripts). The results for CHRNA7 transcripts semi-quantitative assay is shown in (Figure 2-53).
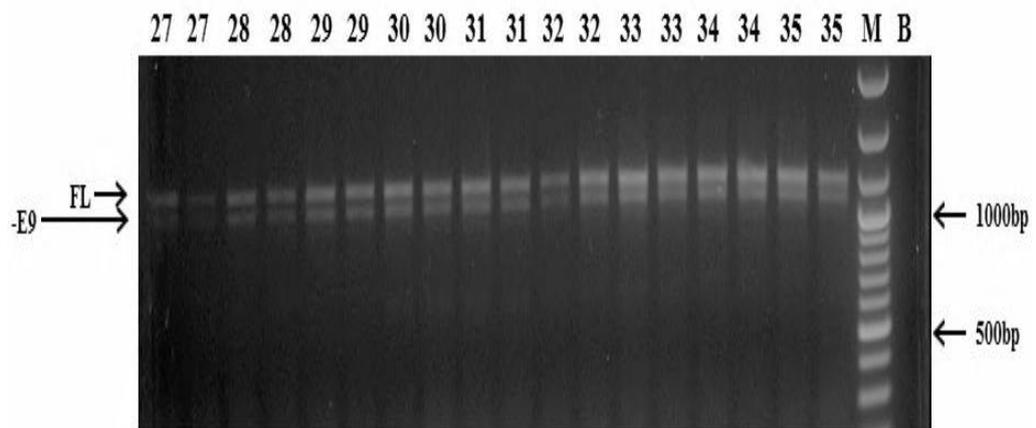


**Figure 2-53: Agarose gel electrophoresis for Semi-quantitative assay of CHRNA7.**
RT-PCR applied on A549 cDNA to test for CHRNA7 transcripts. M: 100 bp DNA ladder, B: blank.

After completing 3 runs for CHRNA7 semi-quantitation, AlphaDigiDoc analysis was applied to measure the band intensities for each duplicate of each cycle. The mean intensity of the bands at each cycle and the means from three runs for each cycle were calculated. These intensities were subsequently used to calculate the corrected intensity adjusted for the band size of the two major bands (full length and missing exon 9 transcripts). Finally, the percent expression of these two transcripts in each cycle was calculated (Figure 2-54).
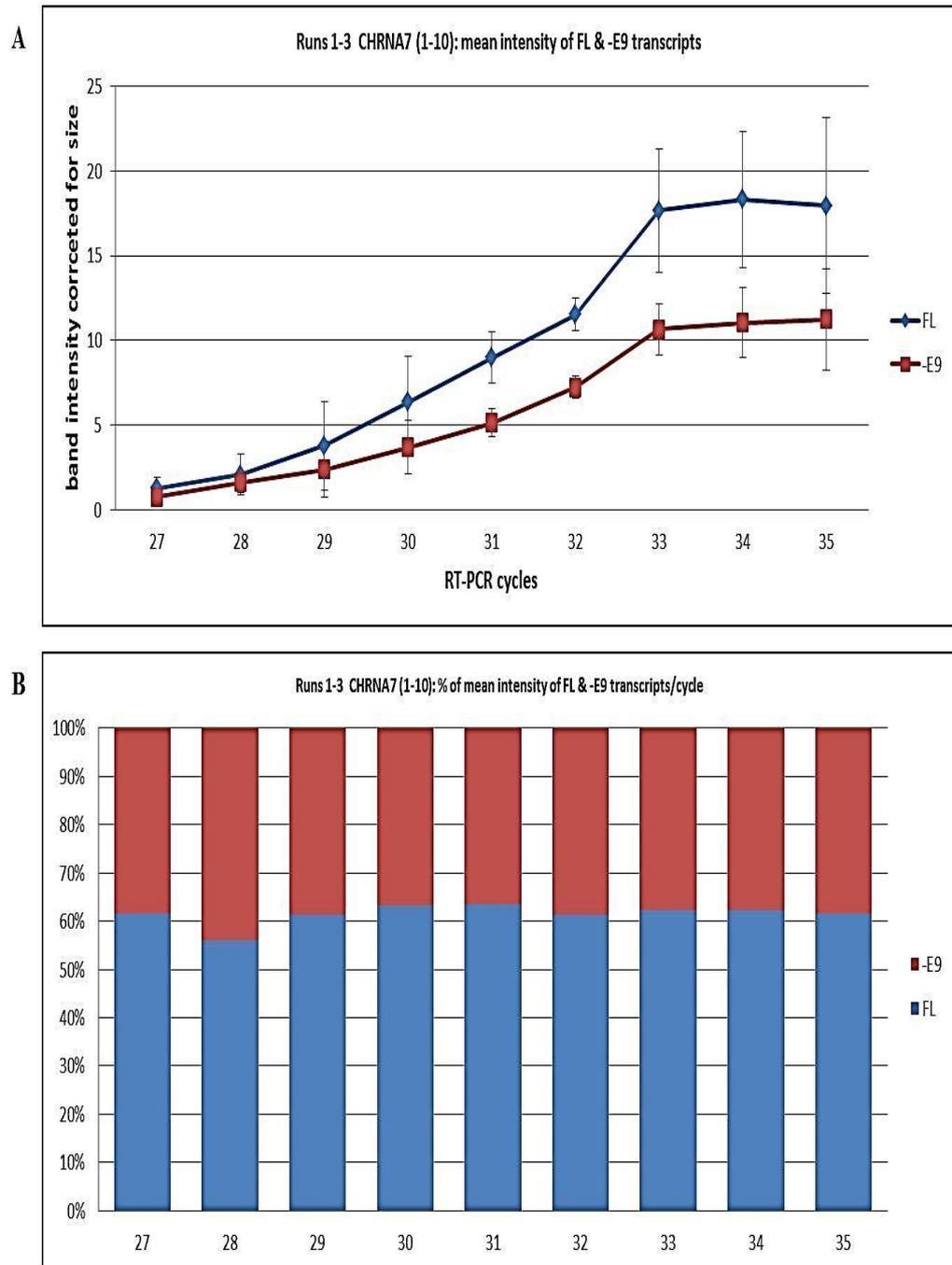
**Figure 2-54: AlphaDigiDoc analysis of CHRNA7 semi-quantitative assays.**
*A:* The intensity curve is showing the correlated increase in intensities of the bands with increasing PCR cycles (mean $\pm$ SEM). This intensity was adjusted for the size of the PCR product on the Y-axis while the PCR cycles on the X-axis. *B:* The percent expression of both transcripts at each cycle presented as bar charts. The percentage is shown on the Y-axis while the PCR cycles are shown on the X-axis. This figure shows the exponential phase starts at cycle 27-33. The relative ratio of full-length: missing exon 9 is ~2:1.

### 2.7.5.2 PCR semi-quantitative assays for CHRFAM7A transcripts:

The same steps were followed for CHRFAM7A transcripts as for CHRNA7.

The results for CHRFAM7A transcripts semi-quantitative assay is shown in

(Figure 2-55).
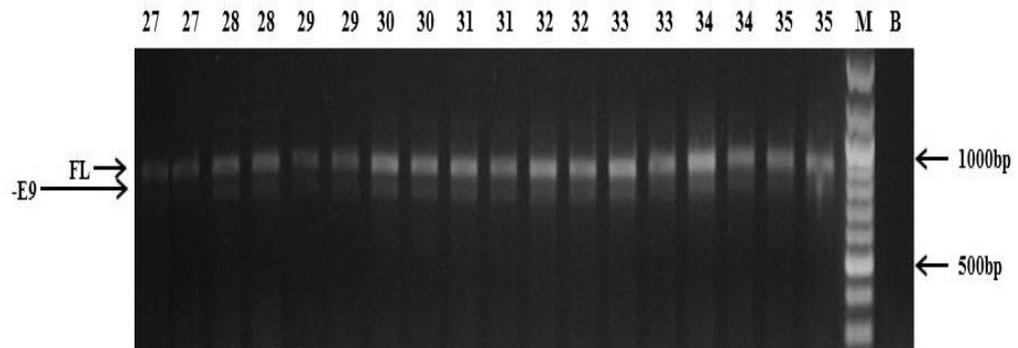


**Figure 2-55: Agarose gel electrophoresis for Semi-quantitative assay of CHRFAM7A.**
RT-PCR applied on A549 cDNA to test for CHRFAM7A transcripts expression. M:

100 bp DNA ladder, B: blank.

After completing 3 runs for CHRFAM7A semi-quantitation, AlphaDigiDoc analysis was applied to measure the band intensities for each duplicate of each cycle. The mean intensity of the bands at each cycle and the means from three runs for each cycle were calculated. These intensities were subsequently used to calculate the corrected intensity adjusted for the size of the two major bands (full length and missing exon 9 transcripts). Finally, the percent expression of these two transcripts in each cycle was calculated (Figure 2-56).
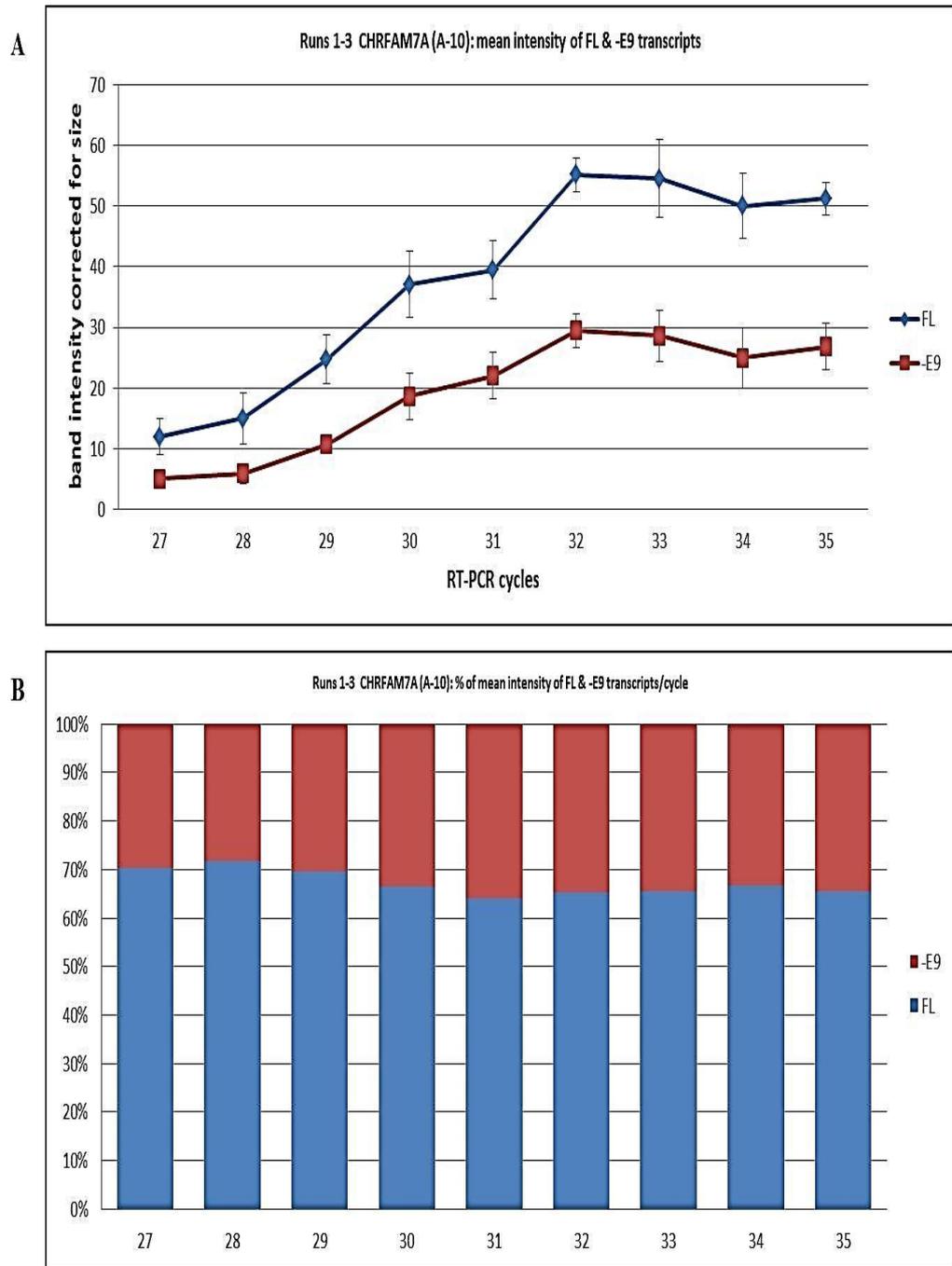
**Figure 2-56: Alpha DigiDoc analysis of CHRFAM7A semi-quantitative assays.**
*A:* The intensity curve is showing the correlated increase in intensities of the bands with increasing PCR cycles (mean + SEM). This intensity was corrected for the size of the PCR product on the Y-axis while the PCR cycles on the X-axis. *B:* The percent expression of both transcripts at each cycle presented as bar charts. The percentage is shown on the Y-axis while the PCR cycles are shown on the X-axis. . This figure shows the exponential phase starts at cycle 27-33. The relative ratio of full-length: missing exon 9 is ~2:1.

### 2.7.6 Bio-informatics analysis of alternative transcripts and protein structure:

**2.7.6.1 Using splicing/alternative splicing tools:**

For detection of ESE sequence motifs within each exon of both genes, and to calculate the ratio of predicted ESE (PESE) to predicted ESS (PESS) the *Human Splice Finder* software program was used (Table 2-20 and Table 2-21). The default thresholds of the above mentioned software programs were used as the reference values. These included the consideration of a constitutive exon for each exon with a PESE/PESS ratio value more than 5.5 compared to alternatively spliced exons (with a ratio of 3.6) and pseudoexons which have a ratio of around 0.63 (Zhang, 2004). Taking these threshold values into consideration, it seems that exons 1, 2, 6, 7, and 10 have a constitutive or alternatively spliced exon characteristic while the same was not true for exons 3, 4, 4a, 8 and 9 when CHRNA7 exons were tested (Table 2-20). These results were somewhat consistent with the previous results detected for alternative splicing affecting CHRNA7 as exons 3, 4, and 4a were found to be frequently missing from RNA transcripts solely or in combinations (Gault et al., 1998, Severance and Yolken, 2008). While no transcripts missing exon 8 were reported before in studies testing human tissues, our study show the existence of transcripts missing exon 9 and backing the PESE/PESS ratio results predictions.

**Table 2-20: ESE analysis for CHRNA7 exons.**

| Exon | ESE | PESE/PESS ratio |
|------|-----|-----------------|
| 1 | 1 | 14 |
| 2 | 15 | 12 |
| 3 | 9 | 0.52 |
| 4 | 16 | 0.38 |
| 4a | 7 | 2.25 |
| 5 | 2 | No silencer motif detected |
| 6 | 12 | 3.86 |
| 7 | 12 | 3.63 |
| 8 | 2 | 0.79 |
| 9 | 6 | 1.08 |
| 10 | 31 | 6.89 |

The PESE/PESS ratio for exons 1,2,6,7 and 10 showed values of constitutive or alternatively spliced exons (>3.6), while those of exons 3,4,5,8 and 9 showed smaller values.

When testing CHRFAM7A exons using PESE/PESS ratios, the results showed that exons D, 6, 7, and 10 have a constitutive or alternatively spliced exons characteristics, while the same was not true for exons C, B, A, 8, 9, and 9b (Table 2-21). Only transcripts missing exon B were detected before (Gault et al., 1998) in addition to missing exon 9 transcripts detected in our study that back up the PESE/PESS ratio results.

**Table 2-21: ESE analysis for CHRFAM7A exons.**

| Exon | ESE | PESE/PESS ratio |
|------|-----|-----------------|
| D | 10 | 3.5 |
| C | 17 | 1.83 |
| B | 8 | 0.24 |
| A | 2 | 0.06 |
| 5 | 2 | No silencer motif detected |
| 6 | 12 | 3.86 |
| 7 | 12 | 3.63 |
| 8 | 2 | 0.79 |
| 9 | 6 | 1.08 |
| 9b | 15 | 2.63 |
| 10 | 31 | 6.89 |

The PESE/PESS ratio for exons D, 6, 7 and 10 showed values of constitutive or alternatively spliced exons (>3.6), while those of exons C, B, A, 5, 8, 9 and 9b showed smaller values.

**2.7.6.2 Predicting the ORF and secondary protein structure:**

For the analysis of the results obtained and for further prediction of the impact on protein structure and expression of the α7 receptor, a number of bioinformatics software programs were used.

**2.7.6.2.1  ORF finder results:**

For the splice variants detected for both genes in A549 cells, the transcripts were subjected to *ORF finder* software for initial detection of the mRNA open reading frame (ORF), possible start and stop codons and by comparing the preserved and the lost parts of mRNA and their corresponding translated parts of the protein (Table 2-22 & Table 2-23).

**Table 2-22: Summary for bioinformatics software analysis for CHRNA7 main transcripts detected in A549 cells.**

| Transcript | Predicted protein amino acids | | | ORF | | | | | Possible effects on the protein | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Total | Conserved | New | ORF | Start codon | Exon | Stop codon | Exon | Preserved | lost |
| **FL** | 502 | 502 | - | +1 | 1 | E1 | TAA | E10 | All | - |
| **Missing E9** | 380 | 306 | 13 | +1 | 1 | E1 | TGA | E10 | SP, N-end, all GLY, all BD, TMD1-3 | All P, most of IC loop, TMD4, C-end |
| FL=full length, E=exon, IC=intracellular, GLY=glycosylation site, N-end= aminoterminal, C-end= carboxy terminal, TMD= trans-membrane domain, BD= binding domain, P= phosphorylation sites, Start codon: the cDNA sequence of starting Adenine nucleotide of ATG of the predicted ORF, SP: signal peptide. Start codon: refers to c.1 of the reference sequence of CHRNA7 accession number NM_000746. | | | | | | | | | | |

**Table 2-23: Summary for bioinformatics software analysis for CHRFAM7A main transcripts detected in A549 cells.**

| Transcript | Predicted protein amino acids | | | ORF | | | | | Possible effects on the protein | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Total | Conserved | New | ORF | Start codon | Exon | Stop codon | Exon | Preserved | lost |
| **FL** | 412 | 412 | - | +1 | 319 | EB | TAA | E10 | All (No SP) | - |
| **Missing E9** | 290 | 230 | 13 | +1 | 319 | EB | TGA | E10 | N-end, all GLY, all BD, TMD1-3,  C-end | All P, IC loop, TMD4, |
| **Inserted E9b** | 264 | 246 | 4 | +1 | 319 | EB | TGA | E9b | N-end, all GLY, all BD, TMD1-3,  C-end | All P, part of IC loop, TMD4 |
| **Missing E9 & inserted E9b** | 221 | 212 | 1 | +1 | 319 | EB | TAA | E9b | N-end, all GLY, all BD, TMD1-3 | All P, IC loop, TMD4, C-end |
| FL= full length, E= Exon, EB= exon B, IC= intracellular, GLY= glycosylation site, N-end= aminoterminal, C-end= carboxy terminal, TMD= trans-membrane domain, BD= binding domain, P= phosphorylation sites, Start codon: the cDNA sequence of starting Adenine nucleotide of ATG of the predicted ORF. Start codon: refers to c.1 of the reference sequence of CHRFAM7A accession number NM_139320.1. | | | | | | | | | | |

**2.7.6.2.2   Predicting the start codon:**

Several software programs were employed to predict the start codon of different transcripts detected earlier (like *the ORF finder* and *the Human Splicing Finder*). In addition, start codons within weak contexts were predicted using the *Weak AUG* software program. This program used the KOZAK consensus sequence to predict the translation initiation site. The presence of Adenosine nucleotide base at position -3 or guanosine nucleotide base at position +1 of an AUG (or ATG in corresponding DNA sequence) can help predict the Kozak consensus sequence for a start codon. The transcripts were first subjected to ORF finder and the Human Splicing Finder analysis to help predict the start codon, and then the ORF predicted the RNA to be translated. Further analysis using WeakAUG was undertaken to test for possible weak consensus start codons. Based on the results shown above (Table 2-22 & Table 2-23), only CHRNA7 transcripts missing exons 4, 3 & 4 or 4 & 5 in addition to all CHRFAM7A transcripts were shown to have a weak start codon for translation initiation (using *weakAUG* software program) (Figure 2-57).
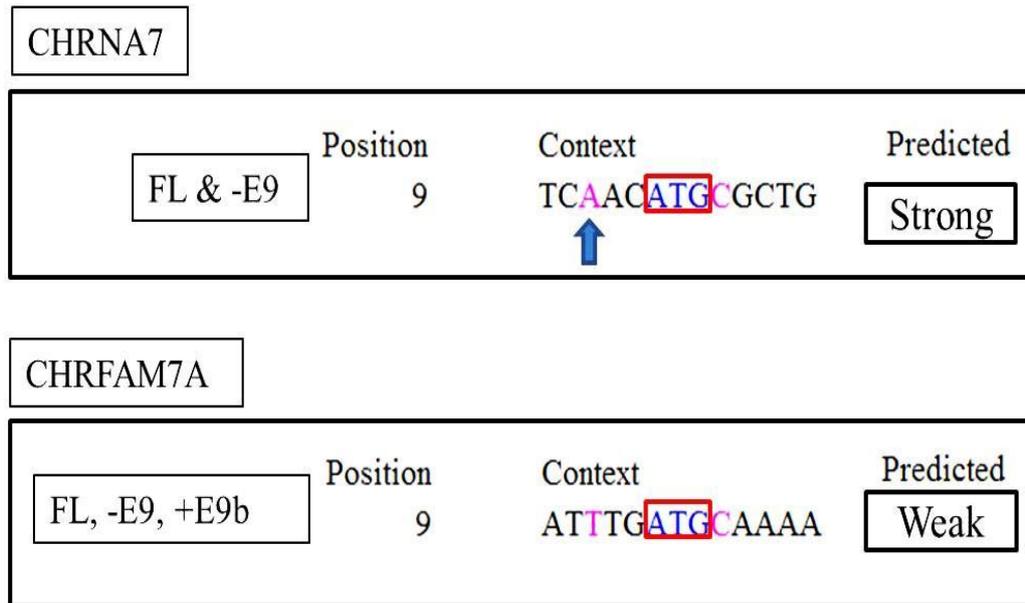
**Figure 2-57: KOZAK consensus sequence prediction.**
*Upper panel:* CHRNA7 transcripts detected in A549 cells. *Lower panel:* CHRFAM7A transcripts detected in A549 cells. Red box: ATG start codon, blue arrow: A at -3 or G at +4 positions required for recognising A at position +1 as the start of start codon. Predicted: indicate a **Weak** or **Strong** ATG predicted in a given sequence. This test showed that CHRFAM7A transcripts had a weak AUG, unlike CHRNA7 transcripts.

**2.7.6.2.3   SOSUI signal software program:**

The transcripts that were detected in A549 for both genes were tested for the presence of a signal peptide within the translated mRNA sequence using *SOSUI* software program. Only CHRNA7 transcripts were detected shown to have signal peptide while CHRFAM7A transcripts did not.

**2.7.6.2.4   PSIPRED software program:**

For prediction of the secondary structure of the proteins translated from the transcripts detected for both genes in A549 cells, the *PSIPRED* software program was used.  After using ORF finder, transcripts with start and stop codons were picked up for subsequent bioinformatics testing. The corresponding amino acid sequences of the selected ORF sequences were subjected to secondary structure prediction using *PSIPRED* software program. The results for CHRNA7 transcripts showed that full length and inserted exon 4a have the same structure with aminoterminal and carboxyterminal ends both were extracellular, 4 transmembrane domains and intracytoplasmic loop (between TMD 3 & 4). The results for missing exon 9 transcripts showed that the translated protein would lose transmembrane domain 3 and part of the intracytoplasmic loop modifying the protein structure with the carboxyterminal end converted intracytoplasmic in position (Figure 2-58).
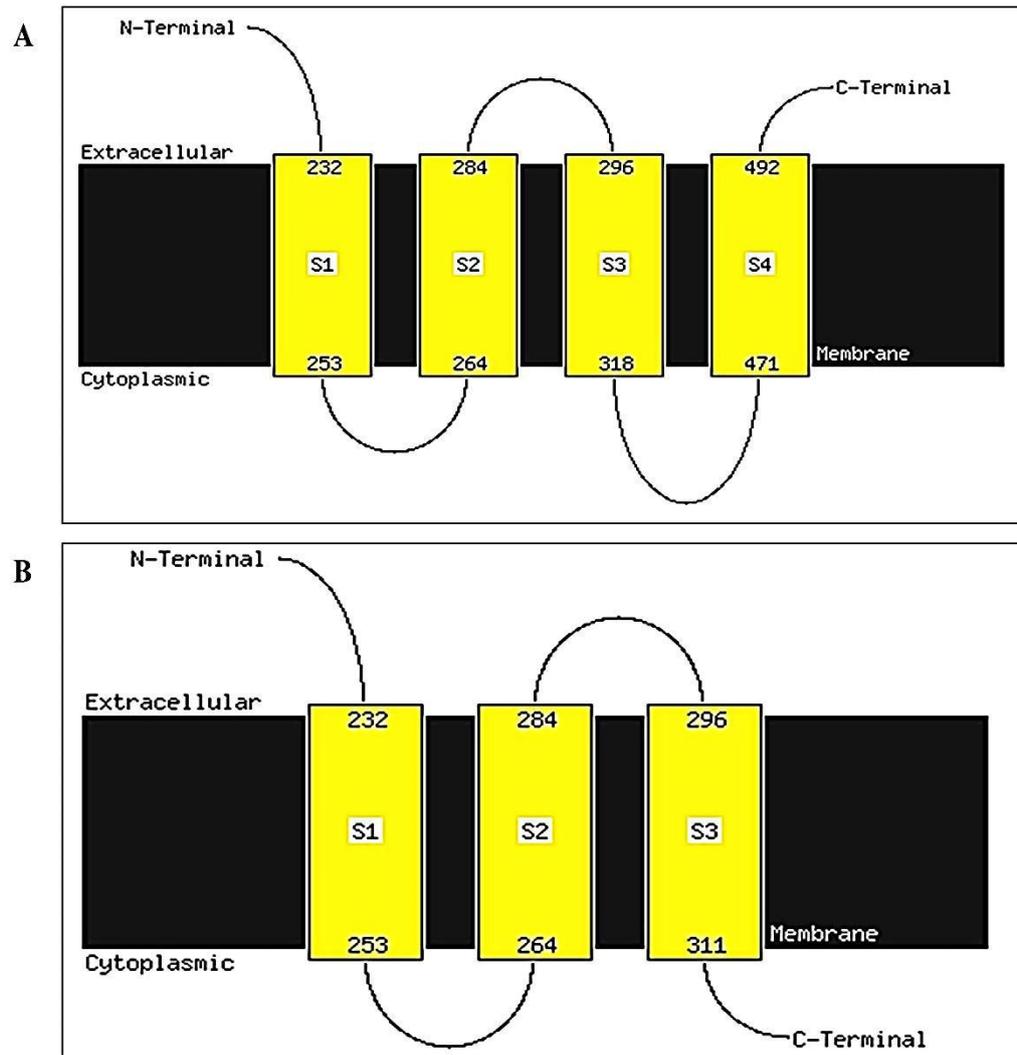
**Figure 2-58: Secondary protein structure prediction for CHRNA7 transcripts.**
*A:* protein structure for the full-length CHRNA7 transcripts. *B:* protein structure for

the missing exon 9 CHRNA7 transcripts.

The results for protein structure of CHRFAM7A transcripts showed that full-length transcripts have a protein with similar structure to that of CHRNA7 full length transcript. The same was true for results of missing exon 9 transcripts, inserted exon 9b and combined inserted exon 9b and missing exon 9 transcripts were all showed similar protein structure for that of CHRNA7 missing exon 9 (Figure 2-59).
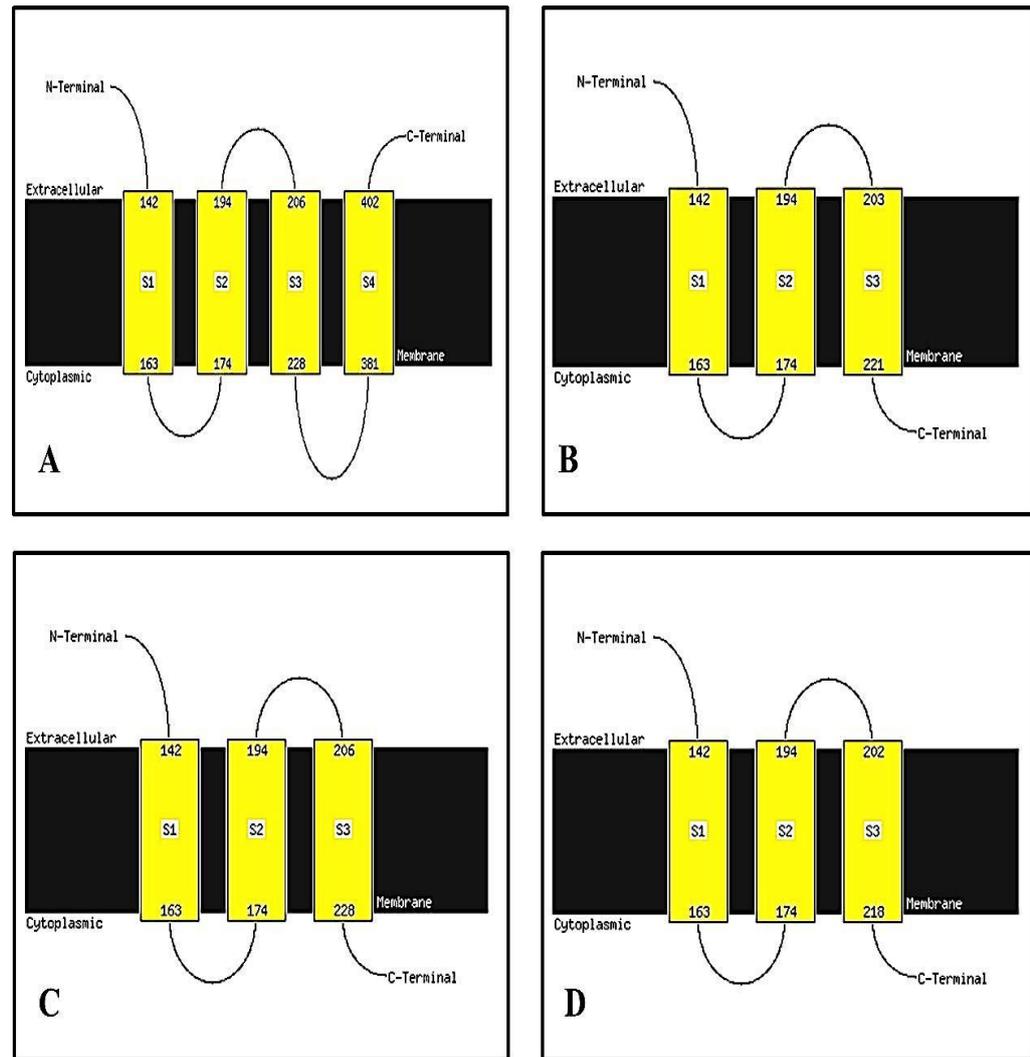
**Figure 2-59: Secondary protein structure prediction for CHRFAM7A transcripts.**
*A:* protein structure for the full-length and. *B:* protein structure for the missing exon 9 transcript protein. *C:* protein structure for the inserted exon 9b transcript, *D:* protein structure for the missing exon 9 and inserted exon 9b transcript. Note: -GT (67-68) of exon 6 predicted protein is similar to full length transcript protein, except for a shorter N-end. This means that missing exon 9, including exon 9b or both could lead to the same results of missing TMD 3 and reversing C-terminus intracellularly.

**2.7.7   Western blotting results:**

Western blot analysis was performed on A549 cell lysates to confirm the expression of α7 protein. After the transfer process, the gel was stained with Coomassie brilliant blue stain to check for the efficiency of transfer of proteins to the membrane. The preliminary results for detecting α7 proteins (using the antibody against the N-terminus part of CHRNA7–from Abcam-) in A549 cells showed many bands detected by Western blotting. However the brightest of these were two bands of around 42 and 34 kDa sizes, respectively. The next run showed clearer, but still abundant bands (run 2). However after changing the blocking into 1 hour rather than overnight, and incubating the PVDF membrane with primary antibody specific for α7 protein overnight, two major bands were noticed above 55 kDa size. Furthermore, another antibody that can detect the intracytoplasmic loop part of the protein (expressed by both α7 and dup α7 proteins- from Santa Cruz-) was used to probe the A549 lysate.  These latter antibodies could detect three bands of 55+ kDa sizes (the predicted size of the protein is 56 kDa) (Figure 2-60).
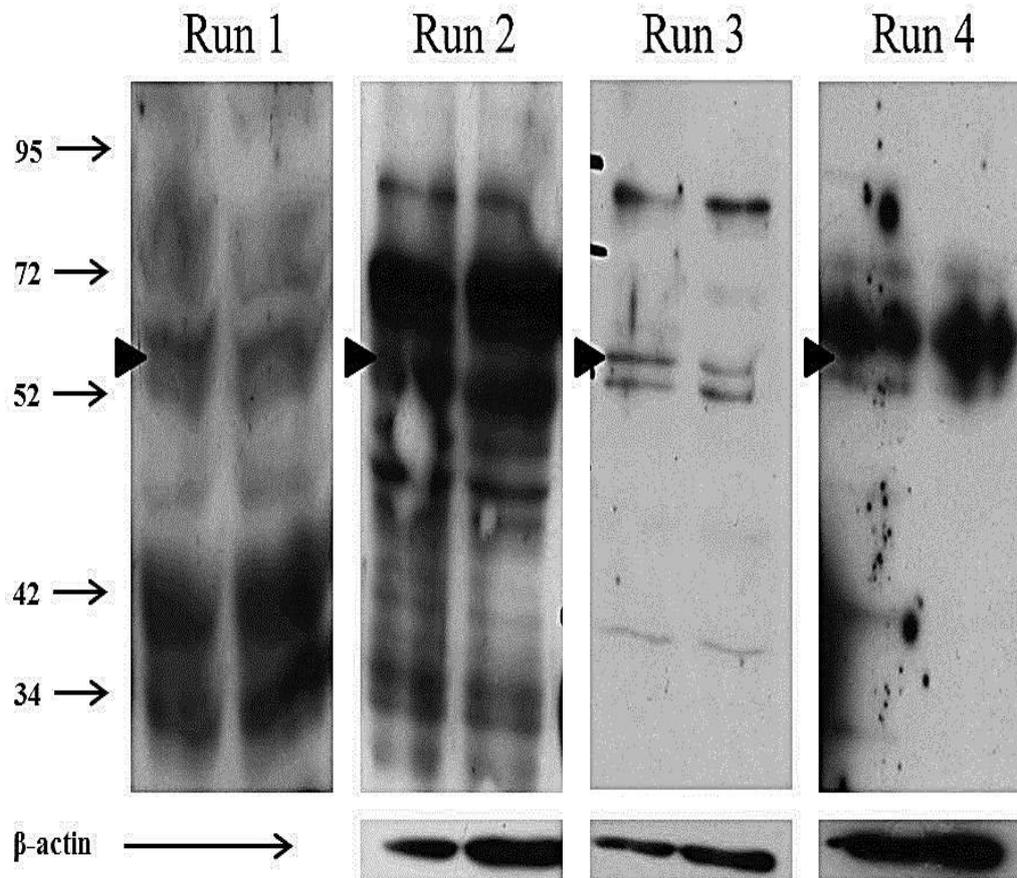
**Figure 2-60: Western blot analysis for A549 cell lysate using α7 antibodies.**
A549 cell lysates were used for testing the antibodies. Size indicator (in kDa) is shown on left of Western blot with arrows, β-actin: indicate control protein (of 46 kDa size) (shown at the lower part of this figure), Black triangle: indicate expected α7 protein band (~57 kDa size). *Run 1*: Many bands were detected, but the brightest two were around 42 and 34 kDa sizes, respectively. *Run 2:* similar to run 1, but with two main large bands around 55+ and 75+ kDa sizes, respectively. *Run 3:* two major bands at 55+ kDa size, *Run 4:* three bands at 55+ kDa size. Note that the antibodies used in runs 1-3 (from Abcam) detect the aminoterminal part of α7 protein only, while the antibodies used in run 4 are (from Santa Cruz ) could detect the intracytoplasmic loop in the common region of both proteins (α7 and dupα7).

The primary antibody against α7 protein was designed to target the amino acid range 22-71 of the human α7 protein. The human α7 protein harbors three N-Glycosylation sites, one of which is at position 46 (within the range targeted by the primary antibody used) which directed us towards the use of a deglycosylation enzyme PNGase F (New England Biolabs). After the deglycosylation step, the Western blot analysis for A549 cells revealed two bands of sizes between 34 and 42 kDa that were different from those reported by Plummer et al (Plummer et al., 2005) but similar to those reported by Drisdel and Green (Drisdel and Green, 2000) (Figure 2-61).
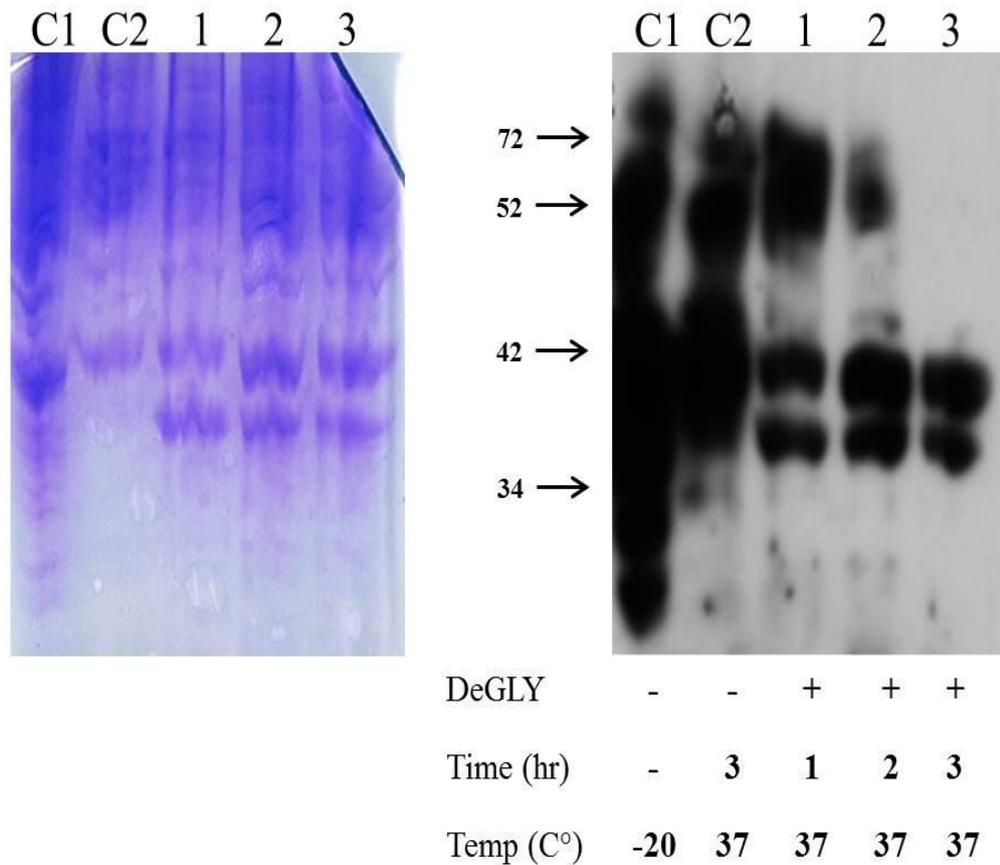
**Figure 2-61: Western blot analysis for α7 nAChR protein.**
A549 lysate was used for Western blot analysis (samples c1, c2, 1-3). *left:* Coomassie blue staining of the gel, *Right:* Western blot result: PNGase F deglycosylation kit was used on A549 lysates (DeGLY) using different incubation times (hr=hours) at 37°. Two A549 controls without deglycosylation were used: C1 (not heated) and C2 (heated at 37° for 3 hr). Size indicator (in kDa) is shown on left of Western blot with arrows.

This was followed by the use of mouse muscle lysate as a control for α7 receptor protein expression according to Abcam antibody website. In addition, PBMC and BE (2)-c cell lysates were checked for their expression of α7 protein (Figure 2-62) (mouse muscle and BE (2)-c cell lysates were kindly gifted by fellow researchers. The results showed that PBMC expressed several forms of the protein of different sizes, while mouse muscle and BE (2)-c cells expressed four different forms of the protein, three of which were almost identical (around 55+ kDa and around 95+ kDa).
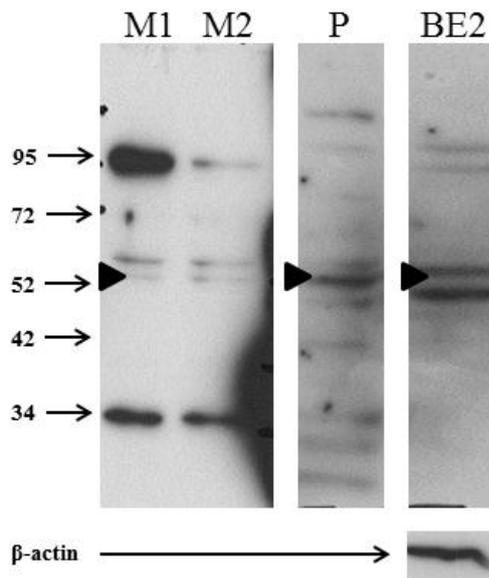


**Figure 2-62: Western blot analysis for protein lysate of other cells using α7 antibodies.**
*M1 & M2:* mouse muscle lysates, *P:* PBMC lysate, BE2: *BE (2)-c* cells lysate, Size indicator (in kDa) is shown on left of Western blot with arrows, *β-actin:* indicate control protein (of 46 kDa size), *Black triangle:* indicate expected α7 protein band (~57 kDa size).

## 2.8   Discussion:

This study involved the characterization of two genes, CHRNA7 and CHRFAM7A that were linked with the anti-inflammatory pathway. Our results show that both genes express a novel transcript that is described for the first time and is missing exon 9. As both genes express exons 5-10 within their transcripts of almost exact sequence, it is interesting to detect transcripts missing exon 9 from both genes, highlighting the fact that tissue splicing factors might influence such a mechanism. The cells tested include A549 and BEAS2B cells, which represent airway epithelial cells.

### 2.8.1   CHRFAM7A feedback control on CHRNA7:

CHRNA7 gene encodes for $\alpha$7 nAChR subunit to form an ion gated channel that helps in controlling inflammation. This receptor seems to be the key regulator for the 'anti-inflammatory pathway' to counteract the effects of inflammation once it started and help keeping it localized (Borovikova et al., 2000, Wang et al., 2003).

CHRFAM7A is a partial duplication of CHRNA7, including exons 5-10 of the latter. The first 4 exons of CHRNA7 are missing from CHRFAM7A and are replaced by exons A-C of ULK4 gene in addition to exon D (unknown origin) (Gault et al., 1998, Riley, 2002, Araud et al., 2011). The first 4 exons of CHRNA7 are responsible for encoding the agonist binding part of the protein, representing its main functional part, and by missing these exons, CHRFAM7A protein (dup$\alpha$7) was proposed to lack the ligand binding function of $\alpha$7 protein. Till recently, the dup$\alpha$7 protein role and function was unknown. Many studies since then have confirmed that the dup$\alpha$7 protein maintains a feedback control

on the α7 protein under basal conditions. Once inflammation starts, the production of inflammatory mediators is supposed to down regulate the CHRFAM7A and thus enhancing CHRNA7 upregulation (Benfante et al., 2011, de Lucas-Cerrillo AM, 2011) (Figure 1-11).

This indicated that in the normal condition, the expression of CHRFAM7A is higher than that of CHRNA7 (at transcriptional and translational levels).

This control mechanism may be affected by the alternative transcripts of both genes. It was shown before that some transcripts of CHRNA7 (missing exon 8 or including exon 9b) rendered the α7 receptor less or non-functional (García-Guzmán et al., 1995, Saragoza et al., 2003).

## 2.8.2    The alternative splicing of CHRNA7 and CHRFAM7A in airway epithelial cells: Novel transcripts detected

The choice of cells tested was based on their role in the anti-inflammatory pathway in the human. Airway epithelial cells (A549 and BEAS2B) and PBMC cells were targeted for testing CHRNA7 and CHRFAM7A expression.

Here we show that there are a number of alternative transcripts detected for CHRNA7 in A549 cells, ranging from insertion of exon 4a to deletions of exons 3, 4, & 5 alone or sometimes in combinations (Table 2-16). Among these transcripts is one missing exon 9 that is reported for the first time. This transcript was detected in all CHRNA7 transcripts targeting the amplification of exon 9 and its flanking coding part (Figure 2-14 & Figure 2-15). Surprisingly, missing exon 9 transcripts were not confined to CHRNA7 and this was also detected with corresponding CHRFAM7A transcripts (Figure 2-25). The expression of some of the transcripts detected in our study was reported before in human neuronal tissue such as missing exon 4, missing

exons 4 & 5 and missing exons 3 & 4 (Gault et al., 1998). Similarly, inserted exon 4a was reported previously in human neuronal tissue (Severance 2008). Interestingly, not all CHRNA7 transcripts could be detected by all methods used. For example, missing exon 9 transcripts could be detected by using RT-PCR or by using gel extraction methods, while missing exons 4 & 5 could only be detected by gel extraction method and missing exons 3 & 4 could only be detected by TA cloning. The lowermost band of amplified exons 2-6 (suggested to be missing exons 3 & 4 on Figure 2-13) was very faint and might get lost during the gel extraction. Similarly, low number of colonies were purified and sequenced from amplified exons 2-6, giving the chance to only some transcripts to be detected, like missing exons 3 & 4 and not missing exons 4 & 5. However, the most surprising results is that no single colony could detect missing exon 9 transcript, an effect that could be explained by causing an inhibitory effect on DNA replication in the host cell, but a definite explanation cannot be provided by our results.

So to check for tissue specific splicing effect, these transcripts (of both genes) were tested in other noncancerous cells, such as BEAS2B, PBMC or other cancerous cells such as BE (2)-c cells (Figure 2-35 to Figure 2-38). The expression of CHFRAM7A (transcript consisting of exons A-10) was higher than that of CHRNA7 (transcript consisting of exons 1-10) in A549 cells, BEAS2B (immortalized human bronchial epithelial cells) and PBMC (human primary peripheral blood mononuclear cells) which is similar to reports by other researchers in other cell lines (Villiger et al., 2002, Severance et al., 2009, de Lucas-Cerrillo AM, 2011). For PBMC, there were many reports that showed that there is no expression of CHRNA7 or of any functional α7

receptors in human leukemia cell lines or PBMC (Villiger et al., 2002, Severance et al., 2009, Benfante et al., 2011). However, a recent study showed that PBMC actually express CHRNA7 and functional α7 receptors although CHRFAM7A and its dupα7 protein were expressed in much higher concentrations (de Lucas-Cerrillo AM, 2011). The only cells that showed a reverse expression ratio (CHRNA7 > CHRFAM7A) were BE (2)-c cells. This was consistent with results from other studies (Araud et al., 2011).

It was shown that exon skipping could be used as a mechanism in enhancing protein function. This effect was shown in Apobec 3 gene, where a protein isoform translated from transcripts missing exon 5 helped to protect against viral infections better than the full-length protein (Li et al., 2012). This might explain the expression of missing exon 9 transcripts of CHRNA7 and CHRFAM7A in the cells we examined. This might shed the light on the possible control mechanism of CHRFAM7A over CHRNA7 via transcripts missing exon 9.

This transcript was not reported in human or animal tissues which might explain a role for species specific splicing (Gromoll et al., 2007).

In addition, our results show that CHRFAM7A transcripts contained another novel transcript including an inserted exon within intron 9 sequence (exon 9b) and a transcripts missing two nucleotide bases TG at position 67-68 of exon 6 (Table 2-19). This seems to be a reported deletion that has been reported in dbSNP and *Ensembl* databases (accession number is rs201490160, showing a MAF=0.37). (Gault et al., 1998). This means that the orientation of CHRFAM7A in the cells tested (A549, BEAS2B and PBMC) was the same to

that of CHRNA7 (tail-to-head orientation) (as mentioned in section 1.3.4.1) (Araud et al., 2011).

Testing for exon 9b showed that it was flanked by a splice site donor and acceptor sites (Figure 2-34). This part of intron 9 is included in sequences from both genes, CHRNA7 and CHRFAM7A. Thus, sequence alignment was used to differentiate the origin of exon 9b transcripts. The results showed that the detected exon 9b transcripts sequences matched those within intron 9 of CHRFAM7A (Figure 2-31). The main difference between the two genes within exon 9b sequence was the two bases at positions 77 and 154 of the exon. These bases were both G in CHRNA7 sequence and C & A (respectively) in CHRFAM7A sequence (termed from this point as G-G and C-A alleles of exon 9b, respectively). Several cell lines were selected for testing the expression of exon 9b, such as A549, BEAS2B, BE (2)-c and PBMC cells. The primer set that amplifies the common part between the two genes (exons 7-10) and RT-PCR were employed for this purpose. The detection of an additional PCR band that was larger than the expected size of exons 7-10 (399 bp) was the initial thread of evidence of inserted sequence expression (Figure 2-14). This transcript was detected in A549, BEAS2B, BE (2)-c cells, but not in PBMC (which could be due to tissue specific factors). All the resulted PCR products were subjected to sequencing, and were confirmed to harbour exon 9b sequence. For the upper band that was detected, the sequence was mixed in some instances that encouraged the use of TA cloning method to isolate the transcripts for subsequent sequencing. This was applied for A549 and BE (2)-c cells and the results were almost identical. Most transcripts from both cells

expressed exons 7-10 with and without exon 9b. The only difference was the inclusion of exon 9b and missing exon 9 in one transcript only in A549 cells. *In-silico* analysis was used for assessment of the nature of the detected sequence. First, splice site prediction was used that confirmed that exon 9b is flanked by a donor and an acceptor splice sites (Figure 2-34). For testing the possibility of exon inclusion, *Exon scan* software was used. However, the results showed that exon 9b was not detected as a constitutive exon.

### 2.8.3 Can the use of PESE/PESS ratio help to predict exon inclusion/exclusion?

The previous report on CHRNA7 transcripts showed some missing exon 3, 4, 5 in single or combined deletions (Gault et al., 1998, Severance and Yolken, 2008). The inclusion or exclusion of an exon within a transcript is dependent on many factors, including the splice sites, the branch point and the presence of enhancer or silencer elements within an exon. Among the factors that help defining an exon as `authentic` or `constitutive` is the presence of more ESEs and less ESSs within its sequence. This will help differentiate pseudo exons which usually have an estimated ESE/ESS ratio of around 0.63 while a constitutive exon will have about 8.6 fold this ratio (ratio ~5.5 ) (Zhang, 2004). By using bioinformatics to predict this ratio for CHRNA7 exons, exons 3 & 4 seem to have a ratio close to that of pseudo exons that might cause their deletion in many of CHRNA7 transcripts. This is not true for exon 5 as the same software could not detect silencer elements within exon 5 and that of exon 8 is still less than 3 (Table 2-20). However, that of exon 9 is lower than 3 (=1.08) and exon 9 is missing from almost half of CHRNA7 transcripts. The

same is true for CHRFAM7A, with the ratio of exons C, B, A, 8, 9 & 9b all less than 3. However, only transcripts missing exon B were reported before (Gault et al., 1998), and only transcripts with missing exon 9, including exon 9b or both were detected in our study (Table 2-18). This selective exon deletion or insertion might be dependent on the ESE and the SR proteins that bind to them, directing the splicing machinery towards or away from these exons, respectively. This may be in part due to the abundance of a specific SR that binds to such ESE. Another factor that affects such alternative splicing processes is a long intron length that precedes the exon to be excluded or included. This might be true for exons 3, 4 and 5 of CHRNA7 where they are preceded by relatively large introns (ranging from 10-70 kb in size) which might explain why the splicing machinery could not detect such exons and rather skip them from the final transcript. The intron size was proved to correlate with splicing, with larger intronic sizes favoring exon skipping. It was shown that reducing intronic sequence sizes flanking an alternatively spliced exon resulted in exon inclusion (Cote et al., 1997, Hertel, 2008). It was proposed that the recognition of an exon requires the interplay of many factors, such as splice sites and enhancer elements. However, for some exons, the minimal changes that might affect splicing may play a significant role in changing the status of that exon in the opposite way (from inclusion to exclusion). This may be due to, for instance,  a change of a SR protein concentration in a specific tissue (Hertel, 2008).

### 2.8.4    Aiming to quantify the full-length and missing exon 9 transcripts:

The next step we tried was to quantify the CHRNA7 transcripts percent expression by A549 (of full length and missing exon 9) using QRT-PCR. For this reason, we designed the primers for detecting these two transcripts within CHRNA7 sequence. This seemed to be best achieved by using the forward primer on exon 4 and the reverse primers bridging exon junctions (8-9 for full length and 8-10 for missing exon 9 transcripts) (Table 2-2 & Table 2-6). These primers sets were targeting the minimum possible amplicons of around 580 bp size for both products. The first results showed that there was a steady delay in amplification with increasing the cDNA dilution used for the reaction (Figure 2-41). However, when subjecting the reaction products to gel electrophoresis, the results showed that there were double products detected for both primer sets (Figure 2-42). This highlighted the fact that the primers might be non-specific and a further difficulty in checking such primers was added by the novelty of missing exon 9 transcripts (a novel transcript that is not yet available on the NCBI database) by using the primer blast software . When the sequence of both reverse primers was checked for sequence similarities between both sets, apart from the exon 8 part that was common between both primers, there were an additional 4 nucleotide bases common between the 5` ends of exons 9 & 10 (nucleotide bases 2-5 were exactly matching between the two exons, which were part of the reverse primers bridging exon junctions (Figure 2-44)). For this reason, the primer specificity was checked using TA cloning of the PCR products of both primer sets. Once detected, they were first checked by sequencing. The results confirmed that only single products were amplified from each primer set, and the sequencing results showed that only one product

was amplified by each set. However, this effect could be possibly due to the presence of the second product (that is amplified missing exon 9 within full-length RT-PCR products, and the reverse is true) at lower levels than the main amplified product. The next step was using the RT-PCR products for TA cloning to isolate the target inserts (full length and missing exon 9). This was completed successfully and confirmed by sequencing (Figure 2-47). The next step was to use these two prepared templates (full length and missing exon 9) for RT-PCR using QRT-PCR primers sets (4-8/9 & 4-8/10). The results showed that both primer sets were non-specific as both sets could amplify both templates (full length and missing exon 9) (Figure 2-48). Sequencing was carried out for all the RT-PCR products and it was expected to see full length template amplified by full length primers (and the same for missing exon 9). However, the surprising part was that full length template was amplified by missing exon 9 primers and a missing exon 9 template was also amplified using full length primers (Figure 2-49). This indicated that the primers were less specific than they were thought to be, and the QRT-PCR approach for detecting the relative percent expression of the two main transcripts was no longer applicable (Figure 2-63).
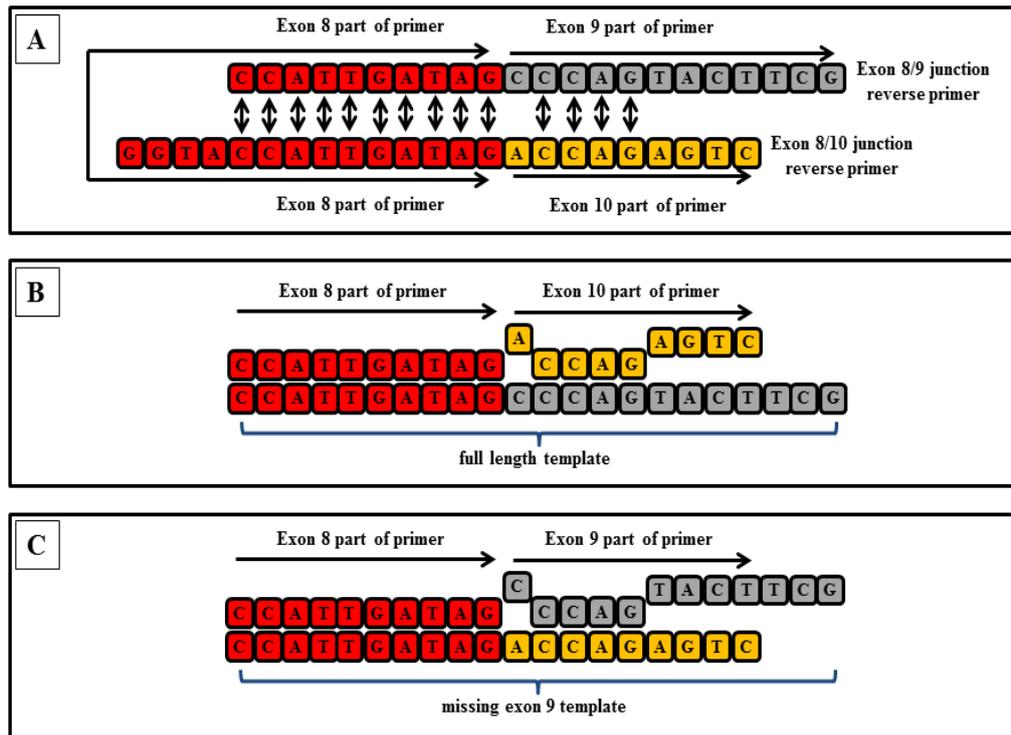
**Figure 2-63: schematic representation for possible cause for primers non-specificity.**
*A:* the similar nucleotide bases matching between both primer sets, *B:* possible explanation for missing exon 9 reverse primer binding to full length template part, *C:* possible explanation for full length reverse primer binding to missing exon 9 template part. Red boxes=exon 8, grey boxes=exon 9, yellow boxes= exon 10.

A further check for the common sense primer on exon 4 for non-specificity was carried out. First, TA clone plasmid templates with full length and missing exon 4 inserts were used for RT-PCR using both primer sets. This time, the forward primer was proved to be specific and could amplify only templates with exon 4 (Figure 2-50). In addition, sequence alignment analysis for exon 4 primer sequence was run against exons D-A of CHRFAM7A. The result again showed that exon 4 forward primer is unlikely to hybridise with any sequence on these exons, and thus it is specific for CHRNA7 sequences (Figure 2-51). This signaled for the presence of many difficulties facing the use of QRT-PCR technique for quantifying the two main CHRNA7 transcripts. Surprisingly,

during the course of search for other references that used QRT-PCR with CHRNA7, a recent paper published results for CHRNA7 (using primers designed to amplify the part between exons 9 & 10 of the gene) and compared it with CHRFAM7A (using primers designed to amplify exons D & C of the gene). A major limitation of this study is that it could not distinguish CHRNA7 and CHRFAM7A transcripts as they used exons 9-10 relative to exons D-C to compare the expression of the two genes (van der Zanden et al., 2012). This is due to the fact that exons 9-10 are expressed by both genes and its amplification signal is reflecting transcripts from both genes. Furthermore, exons D-C might not be the perfect choice for representing CHRFAM7A as they have been shown to be replicated on three loci on chromosome 15 between CHRNA7 and CHRFAM7A, one of these is just upstream of CHRNA7 duplicated part (exons 5-10) and forming the CHRFAM7A (see Figure 1-9 in section 1.3.3) (Araud et al., 2011). This common part between the two genes (exons 5-10) was used in another study which referred this part only to CHRNA7 and used QRT-PCR for amplifying part of exon 10 and interpreting the results as an effect of CHRNA7 solely (Schedel et al., 2011). As an alternative for QRT-PCR, a semi-quantitative RT-PCR method was used to analyse the percent expression of the main transcripts (full length and missing exon 9) of both genes in A549 cells. This was applied using a similar techniques described before for estimating the relative expression of two transcripts of CHRNA7 (Saragoza et al., 2003). RT-PCR reactions for CHRNA7 (exons 1-10 products) and CHRFAM7A (exons A-10 products) were carried out. Duplicate samples were used for the PCR and the samples were removed after each (from cycles 27-35) to determine the exponential and

plateau phases of PCR amplification and the relative percent expression of the two major transcripts for both genes (Figure 2-52 to Figure 2-56). This method helped in estimating that A549 expression for CHRNA7 transcripts was divided into around two thirds of full length and one third of missing exon 9. A similar ratio for CHRFAM7A transcripts was noted. Interestingly, this ratio was more than that reported for other CHRNA7 transcripts as missing exon 8 (García-Guzmán et al., 1995) including exon 9b (Saragoza et al., 2003) or including exon 4a (Severance et al., 2004). As missing exon 9 seems to affect both genes, and bearing in mind that exon 9 is within the common part of both genes, it might be worth considering that exon 9 is considered as a cassette exon in such cases and is not included in all transcripts of both genes. The results of PESE/PESS ratio for exon 9 might support this theory (=1) as it is less than that of constitutive exons (=5.5) and is more than that of pseudo exons (=0.63).

### 2.8.5    Using bio-informatics tools to predict the protein structure:

Using ORF finder shed the light on the possible start codons for the detected transcripts. Interestingly, such analysis showed that missing exon 9 was in-frame, and that CHRNA7 main transcripts full length and missing exon 9 has the same starting codon with missing exon 9 ending shorter than the full length one (Table 2-22). For CHRFAM7A transcripts, the ORF seemed to start at exon B for all the transcripts detected except missing TG transcript (Araud et al., 2011) (Table 2-23). When predicting the secondary protein structure of the detected transcripts using a number of bio-informatics software programs (section 2.6.12) the full length transcripts of both genes were almost identical,

with amine and carboxylic ends located extracellularly, all TMD and intracytoplasmic loop were almost identical. The transcripts missing exon 9 from both genes and all transcripts included exon 9b in CHRFAM7A transcripts showed almost similar pattern of losing the third TMD and some of intracytoplasmic loop with the inversion of the carboxylic end into intracellular localization (Table 2-22 & Table 2-23) (Figure 2-58 & Figure 2-59). It is well documented that the agonist binding domains that are located in the large amine terminus of the α7 protein and that are encoded by exons 1-4 of CHRNA7 (CHRFAM7A lacking this part) to be responsible for the α7 protein folding and assembly in the endoplasmic reticulum (ER) (Green and Millar, 1995). The second TMD is responsible for the ion channel pore lining, and the other TMD (1, 3 & 4) have a role in fixing the protein to the cell membrane (Bertrand D., 1995). The carboxy terminal end was proved to have no effect on changing the physical and pharmacological properties of the receptor (Bertrand et al., 2008). However, this study only tested for the role of carboxy terminal end in its extracellular position. The formation of functional α7 receptors requires several steps. First, the translation of the proper mRNA into α7 protein subunit is required. Then a number of post-translational modification steps take place in the ER to configure the ultimate receptor. These steps include folding of each subunit in such a way that allow the amine termini of the adjacent subunits to form the ACh binding sites and hold the members of the pentamer together (Brejc et al., 2001). This was proposed to be followed by assembly of 5 of the α7 subunits to form the α7 receptor that is set for transport to the cell surface. Then the TMD2 of all the 5 subunits will form the core of the ion channel pore (Bertrand D., 1995, Green and Millar, 1995, Bertrand et al.,

2008). In addition, the formation of three N-glycosylation sites and α-BTX binding sites are essential for the assembly of the receptors (Green and Millar, 1995, Chen et al., 1998). It is worth noting that the presence of RIC-3 protein is essential for proper folding and assembly of α7 receptors in many species including humans (Lansdell et al., 2005, Williams et al., 2005). Our results show that missing exon 9 transcript might cause the translation of truncated protein (towards its carboxy terminal end). This protein, if translated, might assemble with the wild type α7 protein, and thus might be trapped within the ER and not transported to the cell membrane. This is similar to the assumption that dup α7 proteins co-assemble with α7 subunits in the ER and form a heteropentamer that might interfere with its transport to the cell membrane or even render α7 nonfunctional (Araud et al., 2011, de Lucas-Cerrillo AM, 2011).

### 2.8.6   The α7 protein detection:

Previously published work on α7 protein was inconsistent and showed different sizes for the protein in different tissues. Different sizes of α7 protein were correlated with different conformations of the disulfide bonds (Drisdel and Green, 2000). However this explanation was only for larger sized bands than the original protein. Other researchers showed that different sizes could be detected for α7 protein besides the 57kDa predicted protein band. These were attributed to differential post-translational modification, different modification by RIC-3 protein or palmitoylation steps (Rakhilin et al., 1999, Drisdel and Green, 2000, Wang et al., 2001, Williams et al., 2005). Other studies explained such different protein bands detected to different glycosylation modification of

α7 protein in respiratory epithelial cells (Carlisle et al., 2004, Carlisle et al., 2007). Another group showed very weak expression for α7 protein in A549 cells using specific antibodies (Plummer et al., 2005). This result may be linked with the results of other studies showing that the expression of α7 mRNA was more than its protein in respiratory epithelial cells, possibly due to the effect of CHRFAM7A (Wang et al., 2001, Carlisle et al., 2004). We tried to use these same antibodies (as they were the only available antibodies for Western blot analysis specific for CHRNA7). Previous reports with this antibody showed the expression of double bands of 57 kDa size for α7 protein in respiratory epithelial cells (Plummer et al., 2005).

The test for this protein in A549 cells proved to be not a straight forward process and several adjustments were required in our protocol including changing the constituents of the polyacrylamide gel, the buffers used and the membrane for transfer. In addition, different signals were obtained with different blocking buffers and antibody concentrations. After optimizing the Western blot, the test for α7 protein showed the detection of multiple bands in A549 and other cell lines (Figure 2-60). This raised the following question: *Which part of the protein is the antibody targeting?* Then it seemed that the antibody is targeting the area around the first glycosylation site. This raised the need for a deglycosylation step prior to Western blotting step. The results then showed that two major bands (smaller than the original protein) were detected around 42 and 38 kDa sizes. Taking into account that the average molecular weight of an amino acid is about 110 Daltons, and that exon 9 constitutes around 36 amino acids of total of 502 for α7 protein, then by multiplying 36 by 110 Daltons give exon 9 a molecular weight contribution of around 4 kDa to

the total molecular weight of α7 protein. Subsequently, it was predicted that the two protein products detected of 42 and 38 kDa sizes coincides with the detected two major transcripts of full-length and missing exon 9, respectively. The protein sizes detected were different from that reported for α7 protein, and taking into account the different sizes of the reported proteins in different tissues and the addition of deglycosylation step to our Western blotting, collectively there seems to be fair evidence that the detected bands represent the detected mRNA transcripts full length and missing exon 9. However, there is a possibility that this effect of different protein size detected was due to the effect of proteolysis.

It is worth noting that other cell line were tested for α7 protein (mouse muscle protein as recommended by the antibody manufacturer's website, in addition to PBMC and BE (2)-c cells lysates). All these cells showed similar results for the α7 protein, with two bands around 56 kDa (Figure 2-62). Regarding PBMC and BE (2)-c cells, their mRNA showed the expression of full length and missing exon 9 transcripts similar to A549 (Figure 2-36 and Figure 2-38, respectively). This might indicate, at least in these two cell types, that the two main bands expressed might be representing the full length and missing exon 9 proteins, respectively, using the molecular weight of the amino acids for each protein as a guide (as explained above). Similar results for expression of double bands for α7 protein were reported in SH-SY5Y cells (bone marrow with neuroblastoma cells) (Charpantier et al., 2005) and from human macrophages and rat phaeochromocytoma cells (Wang et al., 2003).

It is worth noting that RIC-3 is a well-recognized chaperone that is required for α7 protein assembly and expression (Lansdell et al., 2005) and that the mRNA

transcripts for this protein were detected in A549 cells (that were used in our

study) an earlier study in our laboratory (Al-masmoum, 2011).

## 2.9   Conclusion:

In summary, the results of testing alternative splicing on the CHRNA7 and CHRFAM7A showed the expression of two novel transcripts that are reported for the first time. The first of these is expressed by both genes and is missing exon 9. This transcript preserves the reading frame but is lacking the TMD3 and with an intracellular carboxy-terminus. These changes are predicted to affect the function of the α7 protein, and might affect the subsequent ion channel formed from the resulting protein. It is well known that α7-nAChRs are homopentamers, but it is still possible that such a protein resulting from missing exon 9 to co-assemble with the full-length protein, leading again to an ion channel with altered conduction. This effect might alter the receptor function in the anti-inflammatory pathway, favoring inflammation.

Regarding the second novel transcript detected, it included an inserted sequence within intron 9 (so termed here as exon 9b). This transcript was shown to be expressed by CHRFAM7A by sequencing. To test for possible expression from both genes, further detailed work was carried out using the minigene methods (as will be discussed in detail in the next chapter).

The cells tested for the transcript expression from CHRNA7 and CHRFAM7A were airway epithelial cells and PBMC that are involved in anti-inflammatory pathways and could modulate the inflammatory response associated with COPD.

# 3  The effect of single nucleotide base changes on exon inclusion: exon 9b transcripts

`

This chapter will discuss the preparation of exon 9b minigene constructs from both genes to test for the possible effects of nucleotide base differences on the inclusion of exon 9b from both genes.

## 3.1  Introduction:

During the process of RNA splicing, the spliceosome assembly on the exons is dependent on the sum effect of enhancer and suppressor factors that either drive the spliceosome into or away from such an exon leading to its inclusion or exclusion, respectively (Fairbrother and Chasin, 2000). Among the factors that help establish the spliceosome on the exon are exon splice enhancers (ESEs). These will bind to SR proteins, and oppose the effects of exon splice suppressors (ESSs) binding to silencer elements. The balance of the ESEs and ESSs within a given exon seem to have a major role on exon inclusion or exclusion in the forthcoming mRNA transcript, with the ESEs predominance helping to shift towards exon inclusion while ESSs favours the opposite (Sun, 2000, Dewey et al., 2006, Zhang, 2004).


### 3.1.1  Exon inclusion:

The basic elements required for exon definition are the splice sites and the adenine branch point (Robberson et al., 1990). The exon is defined when the flanking splice sites and adenine branch point are recognised by the complex spliceosome machinery. However, failure to define an exon will have an impact on subsequent exon skipping, directing the splicing machinery to start screening for the next possible exon. The assembly of the spliceosome on a 'defined exon' is not merely driven by the splice sites and the branch point but

rather influenced by additional effects of other auxiliary factors. These include exon size (preferably less than 300 nucleotides), intron size (preferably shorter than 3 kb), enhancer elements represented by ESEs and ISEs (the more abundant, within an exon and its flanking introns, the more an exon is defined), ESSs and ISSs (opposite to ESEs and ISEs effects) (Hertel, 2008, Black, 2003, Reed and Maniatis, 1986). Sometimes two different SR proteins may act competitively to include or exclude an exon.

### 3.1.1.1 Intron size effect:

It is now evident that the sizes of flanking introns affect exon inclusion in the mRNA transcript. The splicing process starts with either exon or intron definition to help recognise the start and end points for the spliceosome machinery assembly. It was shown that the spliceosome might first use the intron definition approach looking for defining the splice sites and the branch point. However, this step is limited by a maximum intron length of 250 nucleotides. Once this limit is exceeded, which is the case most of the time, the spliceosome would switch to the exon definition mode (Hertel, 2008). For alternatively spliced exons, factors like the intron size and ESEs might affect splicing (as mentioned in chapter 2). It was shown that larger intron sizes, particularly those upstream to an exon, would shift the splicing away from such an exon, resulting in exon skipping (Berget, 1995, Hertel, 2008). In another study, it was shown that in addition to the intron size, other important factors, such as splice sites and ESEs, tend to compensate for enhancing splicing. This was illustrated in the form of increased ESEs density in exons flanked by rather

short introns (less than 1.5 kb size) while splice site signal strength was increased in larger size introns (Dewey et al., 2006).

### 3.1.2 Exon inclusion: the effect on certain diseases in human

Alternative splicing was shown to affect the development and progression of some diseases in humans. Among the common examples is spinal muscular atrophy (SMA). This disease is motor neuron disease that could cause death in some patients. There are two genes that encode for survival of motor neuron (SMN) protein: SMN1 and SMN2 genes. In SMA disease, most of the patients suffer SMN1 deletion, allowing SMN2 gene only to code for the SMN protein. This will result in a single nucleotide change from that encoded by SMN1 gene (C→T) change that leads to increasing disease severity (Cooper et al., 2009). This single nucleotide base change seems to affect a known ESE sequence causing exon 7 skipping, thus lowering the SMN protein concentrations. Another study suggested that the base change leads to suppression of SR protein binding rather than changing the ESE function into an ESS (Kashima and Manley, 2003). Whatever the real change is, it will affect the spliceosome assembly on exon 7 of SMN2 mRNA causing its skipping (Figure 3-1, A). Another example is the alternative splicing of Dystrophin gene causing mild form of Duchenne muscular dystrophy (DMD). The Dystrophin gene is liable to a wide range of effects from alternative splicing events due to the large size of the gene. Among these effects is a single nucleotide substitution within exon 31 that can create an ESS site near an ESE one. Such an effect is enough to cause skipping of exon 31 and aid in the production of partially functioning protein that cause the development of this disease form (Figure 3-1, B) (Disset

et al., 2006). Another common effect of alternative splicing on human disease development is the effect of alternative splicing on the tau gene. This gene codes for the tau protein, which is responsible for the assembly and morphogenesis of microtubules in the neurons. This protein has four microtubule binding domains encoded by exons 9-12. However, for a better function of the protein and to prevent neuronal degeneration, domain 2 encoded by exon 10 is omitted by alternative splicing of exon 10, thus keeping the ratio of exon 10 excluded/included transcripts almost equal. This effect is achieved via several nucleotide base changes that can affect an existing ESE site within exon 10 (Cooper et al., 2009). Some of these changes can increase the ESE strength and its binding to SR proteins, causing increased exon 10 inclusion. Such an effect will change the exon 10 excluded/included ratio, leading to FrontoTemporal Dementia with Parkinsonism on chromosome 17 (FTDP-17) (Figure 3-1, *C*) (Liu and Gong, 2008) .
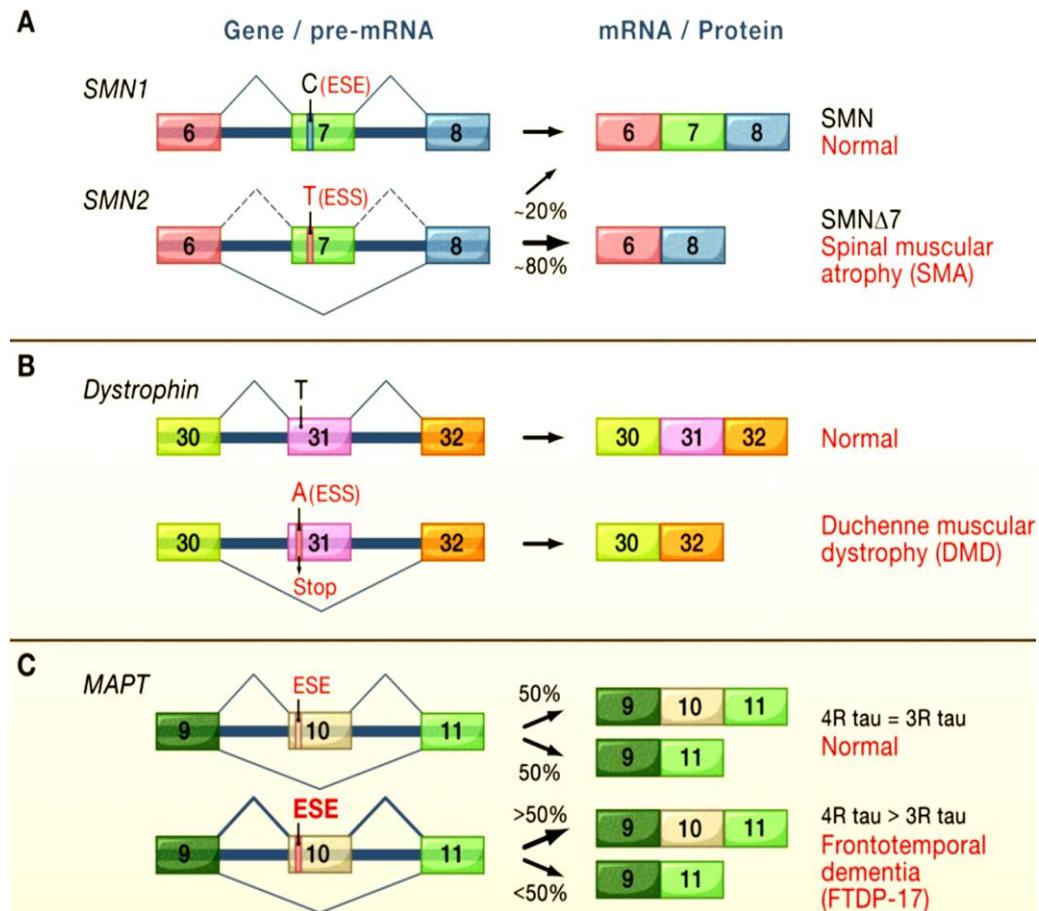
**Figure 3-1: Examples of ESE changes on human diseases.**
(A) The spinal muscular atrophy could result from single nucleotide base changes affecting the SMN protein production. The initial step includes deletion of the SMN1 gene, allowing for SMN2 (which has a C→T substitution in exon 7) to dominate the source of protein-producing mRNA. This will causes ESS to substitute an existing ESE leading to exon 7 skipping. (B) The Duchenne muscular Dystrophy (DMD): a mild form of the disease is precipitated by a single nucleotide substitution in exon 31 of the dystrophin gene (T→A). This change will introduce a premature stop codon (Stop) and an ESS, leading to exon 31 skipping. (C) Frontotemporal dementia: a single nucleotide base change can cause a change in the ratio of transcripts expressing or not exon 10 of the MAPT gene. This change can cause enhanced ESE site, which functions to increase exon 10 inclusion, thus changing the ratio of exon 10 transcripts. Taken from (Cooper et al., 2009)

The alternative splicing can affect the mode of the protein function. Such an effect was shown in the Apobec3 gene in mice, where transcripts missing exon 5 of the gene (due to polymorphisms) lead to the production of the APOBEC3 protein isoform. This protein isoform was shown to act more efficiently, helping in resisting viral infections better than the full-length protein (Li et al., 2012).

### 3.1.3   Exon 9b: a novel exon

In our earlier results, we showed that there was a transcript with a novel exon (exon 9b) from A549 cells in PCR products amplifying the common sequence from CHRNA7 and CHRFAM7A genes (exons 7-10 transcripts). However, DNA sequencing results later confirmed that these transcripts originated only from CHRFAM7A according to the reference sequence of both genes. When testing CHRFAM7A transcripts in A549 and BEAS2B cells, similar transcripts which included exon 9b were detected. These findings raised the following question: *do exon 9b transcripts originate only or mainly from CHRFAM7A?* And if this was true, do these two base differences have a possible role in this exon inclusion? To answer this question we decided to design minigene vectors that carry exon 9b sequences from both genes with only two nucleotide bases different between the two constructs.

### 3.1.4   Using the minigene technique:

The use of the minigene method was a key for exploring many ambiguous RNA splicing mechanisms. This method includes the generation of plasmid constructs that can harbor parts of or entire genes to study the process of

splicing and the related control mechanisms. The minigene construct preparation includes the use of vectors with a restriction enzyme recognition region that is flanked by multiple restriction sites. The latter sites can enhance the use of restriction enzymes that can recognize sequences on the vector and on the template amplified (Cooper, 2005). These sites are usually selected based on the sticky end method, and thus allowing for ligation of the insert with the plasmid or vector. This helps constructing a hybrid gene that can be introduced into different cell types for studying the splicing mechanism. Using the minigene method can help testing for cell, tissue, or species-specific splicing effects (Gromoll et al., 2007). Furthermore, minigene methods could test as small as ten nucleotides and up to several kb sequences. The use of the minigene method was employed for testing the splicing effect of nucleotide changes on exon inclusion/ exclusion (Smith, 1989, Cote et al., 1997, Kosaki et al., 1998, Gromoll et al., 2007, Vezain et al., 2010), splice site alterations (Reed and Maniatis, 1986, Cote J, 2001), ESEs/ESSs sites (Jumaa and Nielsen, 1997, Ibrahim et al., 2005) or even in treatment trials (Phelps et al., 1995).

Exon trap (also called pET01) is a cloning vector that allows the insertion of exons within the multiple cloning site of the vector. This vector represents a shuttle system that contains the necessary genetic factors for replication within the host cells. It can help positioning the exon (with its flanking intronic sequences) flanked by exon trap introns. In this way, any sequences within this site that are defined by the splicing machinery as 'exons' would be recognised once they are introduced into mammalian cells, and would be included within the mRNA of the vector, flanked by the exon trap 5` and 3` exons, respectively (Figure 3-2).
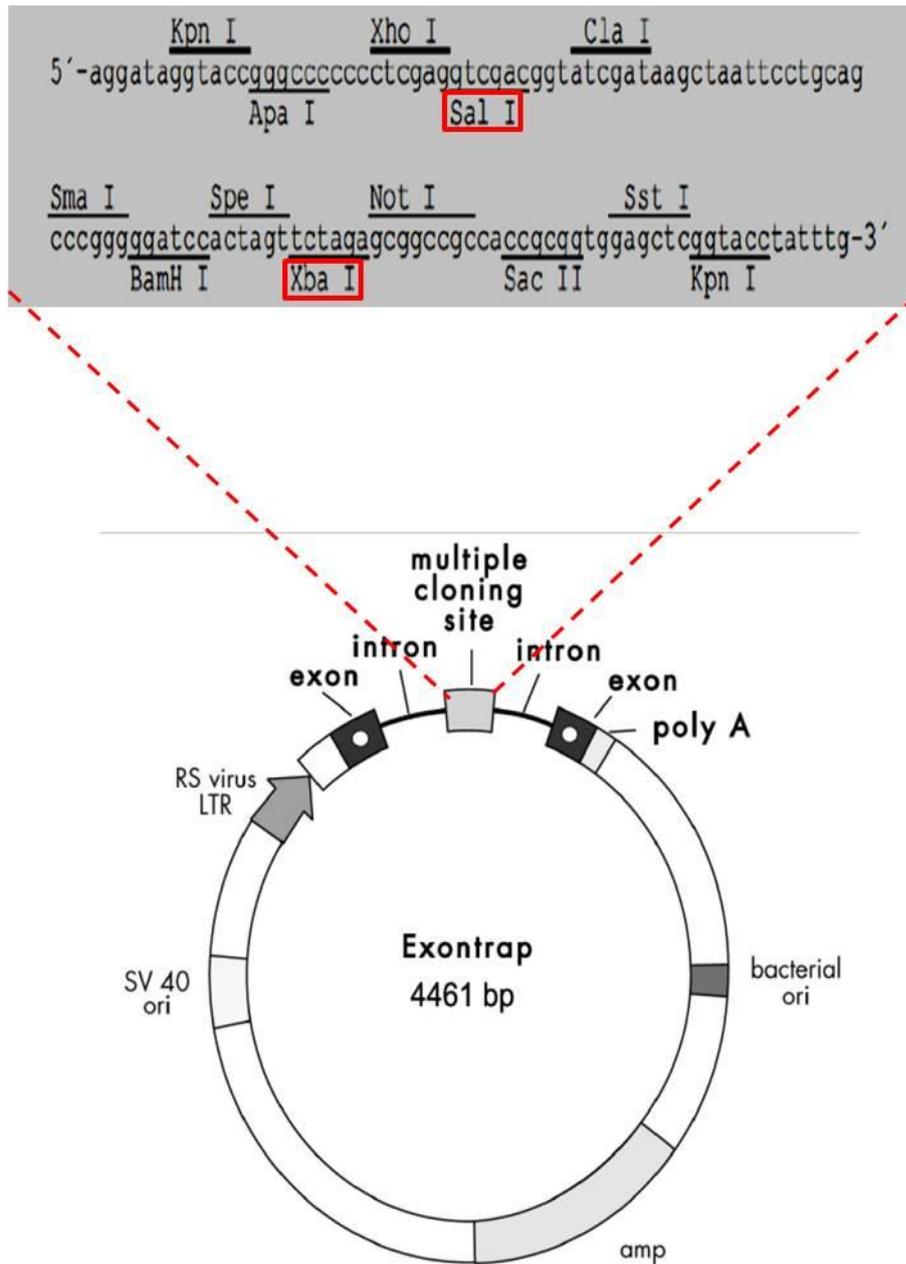
**Figure 3-2: Exon trap vector.**
Schematic representation of exon trap vector, Bottom: showing exon trap multiple cloning site enclosed between its 5` & 3` exons, Top: multiple cloning site showing restriction enzymes recognition sites, SalI and XbaI (within red boxes) were selected as target sites.

In summary, minigene experiments included preparing the two alleles of exon 9b from CHRNA7 and CHRFAM7A (G-G and C-A alleles) in genomic DNA samples by using TA cloning. This was followed by inserting these alleles into the exon trap vector for minigene construct preparations. These constructs were used as templates for preparing another two constructs using site-directed mutagenesis (G-A and C-G alleles). All the constructs were transfected into A549 and Cos7 cells, and the RNA was extracted from these cells for RT-PCR testing of exon 9b inclusion. The RT-PCR results were analysed on gel electrophoresis and then using AlphaDigiDoc software (Figure 3-3).
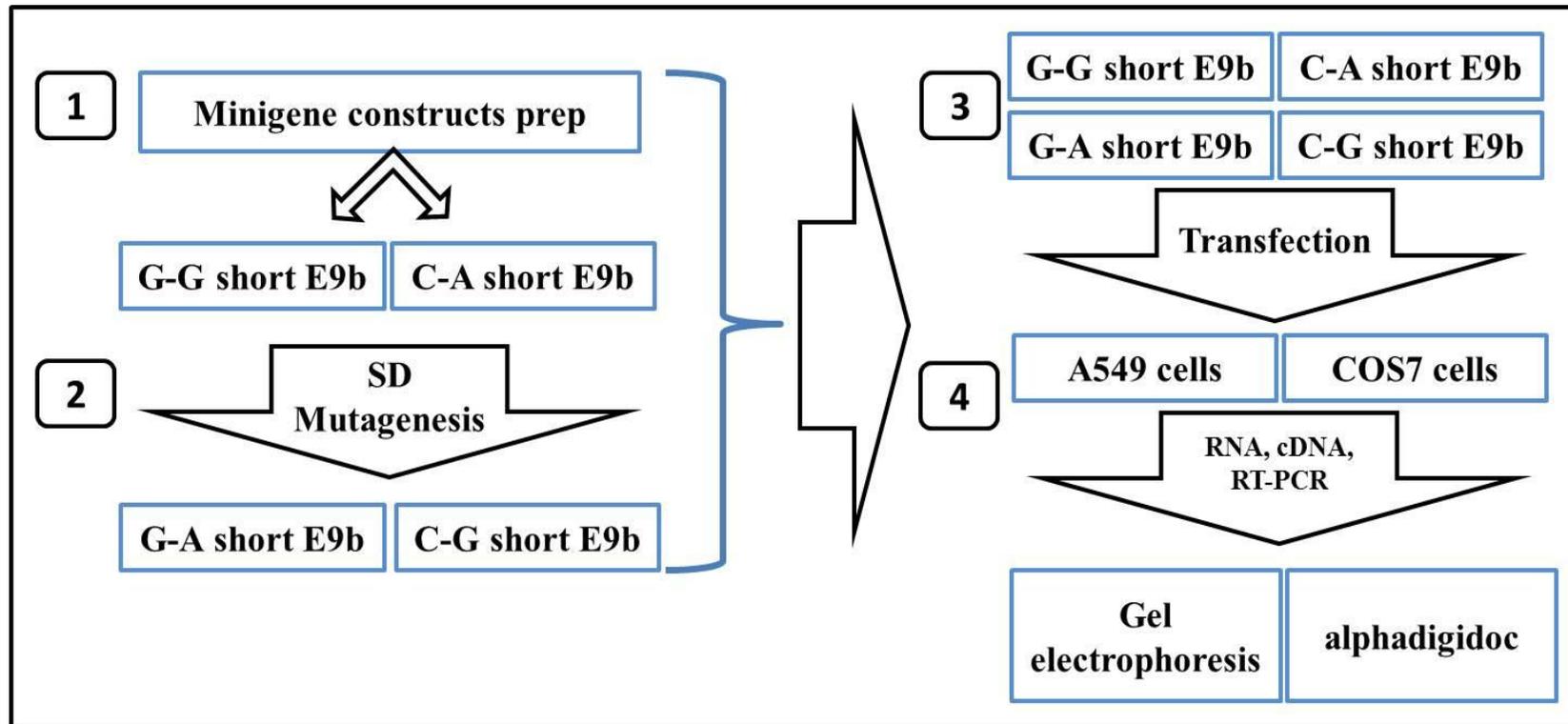
**Figure 3-3: Minigene experiments summary.**
SD: site-directed, alphaDigiDoc: software for densitometric analysis tool. Short: refers to short exon 9b fragment minigene constructs.

## 3.2  Aim:

The work of this chapter aimed at testing the effect of the two nucleotide bases (different between both genes within exon 9b sequence) on exon recognition, and which bases have the most influential effect on exon 9b inclusion.

To achieve this aim, minigene plasmid vectors carrying exon 9b sequences from CHRNA7 and CHRFAM7A genes were designed and prepared. Then these minigene constructs were transfected into A549 and Cos7 cells to test for the effect of the two nucleotide bases on exon 9b recognition.

## 3.3   Methods:

### 3.3.1   Bio-informatics analysis:

First, for confirming the sequence detected for exon 9b and the flanking introns, similar methods were used (as mentioned in section 2.6.12). This was true for extracting the reference sequence for intron 9 from both genes using *NCBI*, *UCSC* and *Ensembl* databases, and for using *ClustalW* sequence alignment to compare the resulting transcripts sequences. For testing exon inclusion within a transcript, *Exon Scan* software was used. This software uses the GT-AG base and RESCUE-ESE software for exon inclusion within a given sequence. For prediction of the possible restriction enzyme recognition sites within exon 9b and flanking introns, *Web Cutter* and *Neb Cutter* software programs were used. For primer design, *primer3(v.0.4.0)* software was used for designing the primers, *Primer-Blast/NCBI* software was used to check for the primer specificity and *SNPCheck* software was used to avoid including SNPs within the designed primers.  For exon definition, splice site prediction software *(BDGP)* was used (Reese et al., 1997). However, for predicting the ESE sequences within the detected exon 9b sequences, *ESE finder (release 3.0)* (Cartegni et al., 2003, Smith et al., 2006) and Splicing rainbow software programs were used in addition to other methods mentioned in section 2.6.12 (*RESCUE-ESE* and *Human splicing finder*). Further analysis for detecting possible SNPs within exon 9b were carried out using *Ensembl* and *dbSNP* software programs. The default thresholds of the above mentioned software programs were used as the reference values (Table 3-1).

**Table 3-1: Summary for the bio-informatic software programs and tools used for analysis of exon 9b transcripts.**

| Software program | URL |
|---|---|
| *Clustal W (sequence alignment)* | http://www.ebi.ac.uk/Tools/msa/clustalw2/ |
| *Exon Scan (exon definition)* | http://genes.mit.edu/exonscan/ |
| *Web cutter (restriction enzyme recognition sites)* | http://users.unimi.it/~camelot/tools/cut2.html |
| *Neb cutter (same as above)* | http://tools.neb.com/NEBcutter2/ |
| *Primer 3 (v.0.4.0) (PCR primer design)* | http://frodo.wi.mit.edu/ |
| *Primer blast/NCBI (sequence* | http://www.ncbi.nlm.nih.gov/tools/primer-blast/ |
| *SNP check (check presence of SNP within a designed primer sequence)* | https://ngrl.manchester.ac.uk/SNPCheckV3/snpcheck.htm;jsessionid=4228B00C8535DAEE9062415214505CC7 |
| *BDGP (splice site detection)* | http://www.fruitfly.org/seq_tools/splice.html |
| *ESE finder (release 3.0) (detect potential ESEs)* | http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home |
| *Splicing rainbow (predict  SR protein binding)* | Currently not available |

### 3.3.2    Preparation of the minigene constructs:

For preparing minigene constructs, exons 9b alleles (G-G & C-A) from short and long exon 9b fragments were amplified (using PCR) from genomic DNA samples of heterozygous subjects (details about these fragments will be discussed shortly after this summary). After preparing the amplified exon 9b (PCR products), these were introduced into the exontrap vector sequence (MoBiTec GmbH 2010).

### 3.3.2.1 Isolation of exon 9b alleles:

Prior to preparation of minigene constructs, exon 9b alleles were isolated from CHRNA7 and CHRFAM7A sequences. For this purpose, two exon 9b fragments were designed (short and long) using primer sets that amplify short (around 100 bp) or long (around 900 bp) intronic sequence upstream of exon 9b. Furthermore, as exon 9b and surrounding intronic sequences are almost exactly identical between CHRNA7 and CHRFAM7A genes, using the designed primers from either segment would amplify exon 9b from both genes. According to the NCBI reference database, there are only two bases within exon 9b sequence that are different between the two genes and these are at positions 77 and 154 of exon 9b sequence, respectively. These two bases are G-G in CHRNA7 allele (termed here as the G-G allele) and C-A in CHRFAM7A allele (termed here as the C-A allele). For this purpose, the term G77 during the text would refer to G nucleotide base at position 77 (in the G-G allele) while the term A154 would refer to A nucleotide at position 154 (in the C-A allele).

### 3.3.2.1.1  Exon 9b primers design:

For amplifying exon 9b from both genes, the reference sequence alignment of intron 9 parts of both genes was used as a target for designing the primers. While one reverse primer was designed (starting 122 nucleotide bases downstream of the exon 9b 3` end), two sets of forward primers were designed with the first primer (starting 133 nucleotide bases upstream of the exon 9b 5` end) generating "short exon 9b fragment" and the second primer (starting 994 nucleotide bases upstream of exon 9b 5` end) generating "long exon 9b fragment" (Figure 3-4). The reason for using two different segments surrounding exon 9b is to test for the regions of sequence differences between CHRNA7 and CHRFAM7A that are present within the long fragment according to the databases used. However, it could also prove to be useful when thinking of possible enhancer or suppressor elements within the intronic sequences upstream of exon 9b for possible stimulatory or inhibitory effect on inclusion in the transcripts (Cote et al., 1997).
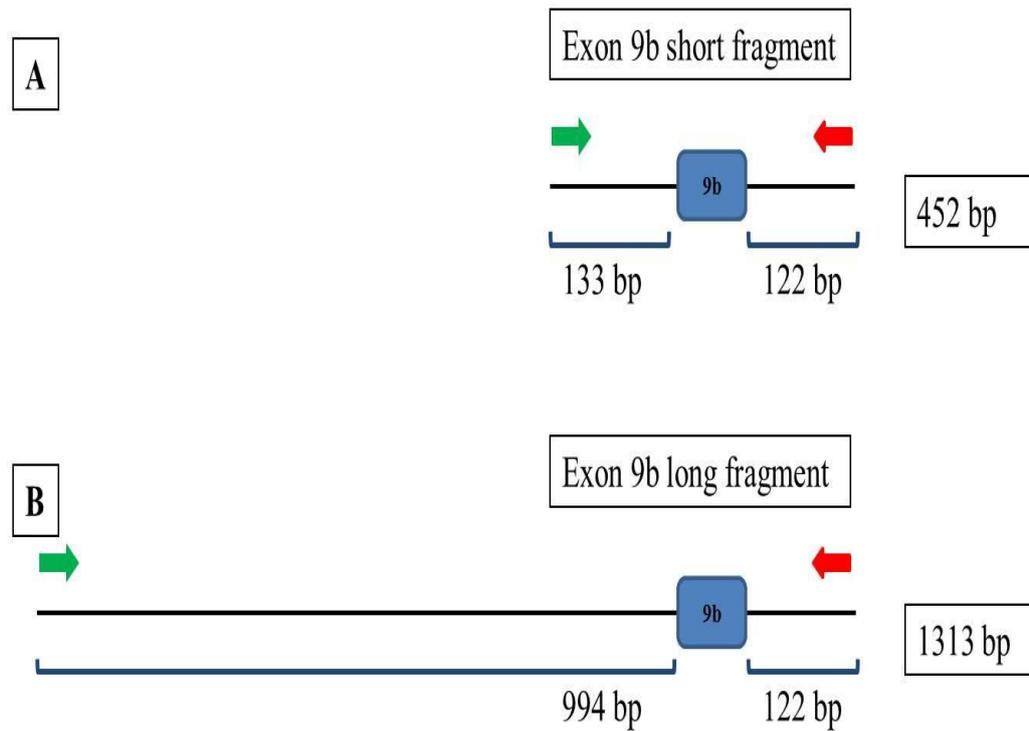
**Figure 3-4: Diagram showing location of primers used to amplify short and long exon 9b fragments.**

*A:* exon 9b short fragment, *B:* exon 9b long fragment. Blue boxes: exons, Green arrow: forward primer (including sequence for SalI restriction enzyme), Red arrow: reverse primer (including sequence for XbaI restriction enzyme), flanking gene (CHRNA7 or CHRFAM7A) introns (black lines) sizes are indicated below the introns; the expected size of the construct is shown in the boxes on the right. Exon 9b size= 197 bp.

For optimal results using the minigene method, the primers sets for generating the two allelic constructs were designed to amplify exon 9b with its flanking introns (at least ~ 0.1 kb of introns on either side) (Cooper, 2005). For all the primers, a restriction recognition site was added on 5` end (SalI site for forward primers and XbaI site for reverse primer) (Table 3-2).

**Table 3-2: Nucleotide sequences of the primers used for exon 9b transcripts detection.**

| Primer ID | Primer sequence (5`→3`) | Position |
|---|---|---|
| **E9b-shortS** | AATAAG<u>GTCGAC</u>TGCCTTAACAATGTAGCACCT | I 9 |
| **E9b-longS** | TCCAAG<u>GTCGAC</u>AGACCCAACCTTTCCAGGAC | I 9 |
| **E9bAS** | ACAGCT<u>TCTAGA</u>AAGGGGAAGGCTATCTGAGC | I 9b |
| All primers are from Eurogentec, Belgium. I= intron, S: Forward primer, AS: reverse primer.  E9b-short: upstream of exon 9b by ~ 0.1 kb, E9b-long: upstream of exon 9b by > 0.9 kb, E9bAS: downstream of exon 9b by ~ 0.1 kb. Restriction enzyme recognition sites for SalI (on E9b-shortS & E9b-longS) and XbaI (on E9bAS) are underlined. |||

### 3.3.2.1.2  Genomic DNA preparation:

Genomic DNA was extracted from A549 and BE (2)-c cells for PCR amplification applications. QIAamp DNA blood mini kit (QIAGEN) was used for this purpose according to the manufacturer's guidelines.

### 3.3.2.1.3   TA cloning of exon 9b PCR products:

The PCR products of exon 9b (from both short and long fragments and for both alleles (G-G & C-A)) were inserted into the pSC-A-Amp/kan vector using the TA cloning method (as mentioned in section 2.6.6).

### 3.3.2.1.4   Restriction digestion using Eco57I:

For initial screening of the allele constitution of the initially prepared TA clones, the restriction enzyme Eco57I (AcuI) (Fermentas Life Sciences) was used. The recognition site of this enzyme is located around the G77 of exon 9b. This means that this enzyme can cut the G-allele of exon 9b (from CHRNA7) but not the C-allele (from CHRFAM7A). The enzyme would cut at specific positions on the insert-plasmid construct and will show different product sizes on electrophoresis according to the allele digested (Figure 3-5) (Table 3-3).
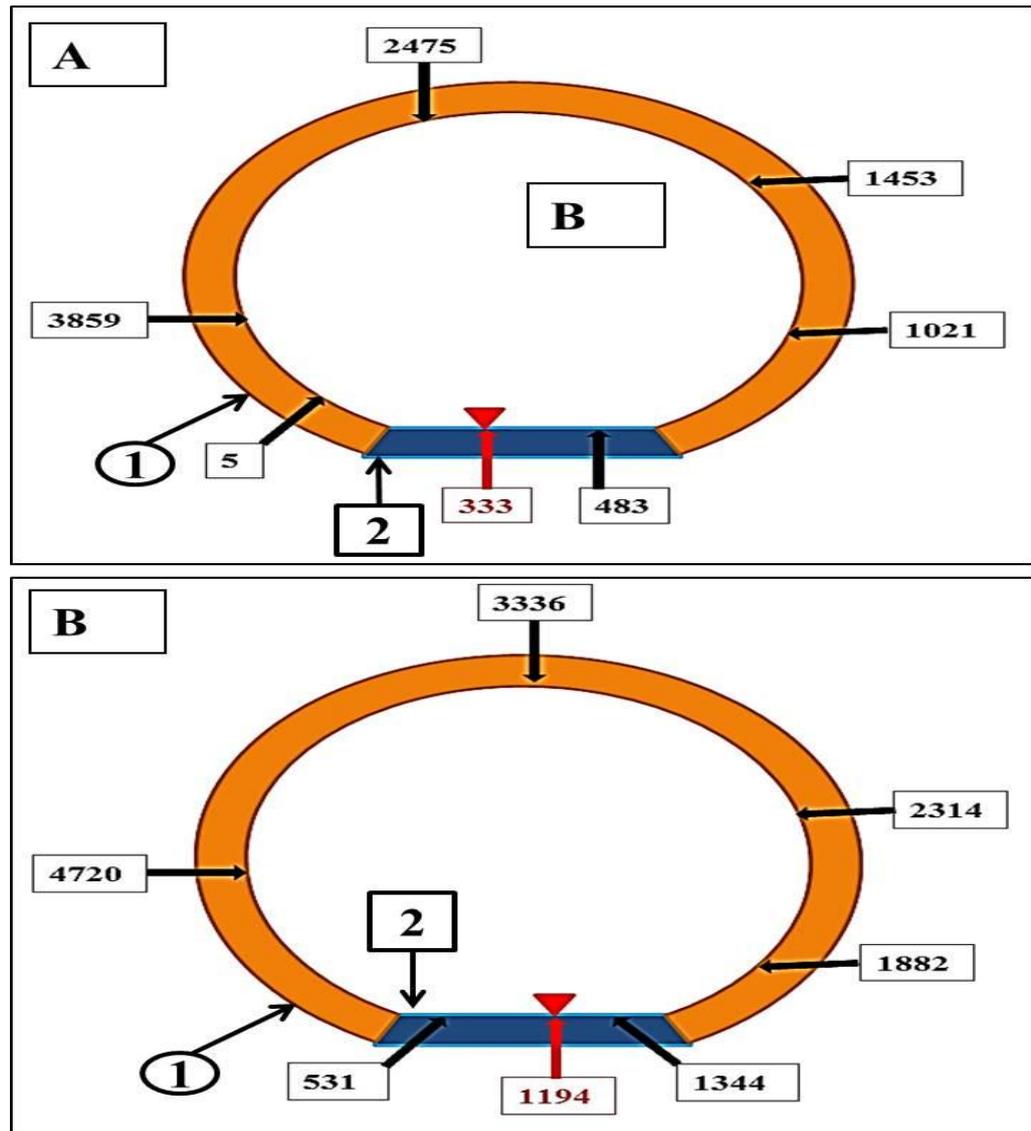
**Figure 3-5: Schematic representation of Eco57I enzyme recognition sites.**
Eco57I recognition sites (Black arrows) on pSC-A-Amp/Kan vector (orange boxed line) and inserts of exon 9b (blue boxed line): *A)* short fragments, *B)* long fragments. Red arrow: recognition site only for G at position 77 of exon 9b (indicated by inverted red triangle). The plasmid numbering starts at 1 (in circle) while the insert numbering starts at 2 (in square).

**Table 3-3: The product sizes resulting from restriction using Eco57I enzyme.**

| Fragment used | G-G allele-specific product sizes (bp) | C-A allele-specific product size (bp) | Allele non-specific product sizes (bp) |
|---|---|---|---|
| Exon 9b short fragment | 150 , 328 | 478 | 432, 538, 866, 1022, 1384 |
| Exon 9b long fragment | 150 , 663 | 813 | 432, 531, 538, 863, 1022, 1384 |

For restriction digestion using Eco57I, up to 1 µg of the target DNA was added with G and SAM buffers (final concentration 1x for both) and finally adding 2.5 units of Eco57I into a total reaction volume of 20 µl. The reaction mixture was then mixed gently and incubated for 60 minutes at 37°. This was followed by a heat inactivation step for the enzyme at 65° for 5 minutes. The digested DNA samples were subjected to gel electrophoresis (as mentioned in section 2.6.5 using 2% agarose gel).

The restriction using Eco57I helped to differentiate the G-G and C-A clones within each group of exon 9b fragments (short and long). This was followed by sequencing of six plasmid-inserts from each of the four different clones that were prepared (confirmed by sequencing using M13 forward and reverse primers). The plasmid-inserts picked up for subsequent preparation of minigene constructs were mainly four alleles: the G-G alleles (short and long) from CHRNA7 and the C-A alleles (short and long) from CHRFAM7A.

### 3.3.2.2 Preparation of minigene constructs:

First, each of the isolated exon 9b clones (mentioned above) were cut using restriction enzymes SalI and XbaI (as illustrated in Table 3-2). The restriction

of exontrap vector was applied in a separate step. This was followed by running the digested DNA fragments on gel electrophoresis (as mentioned in section 2.6.5). Using dark reader, the bands of expected size for each exon 9b allele (or the vector)-digested fragment were gel extracted and purified (as mentioned in section 2.6.7). This step was followed by ligation of each of the exon 9b alleles isolated with exontrap vector and the different plasmid-inserts prepared were used for subsequent cloning.

### 3.3.2.2.1   Restriction digestion using SalI and XbaI:

For the minigene constructs preparation, FastDigest SalI and XbaI enzymes were used (Fermentas Life Sciences). The clones that harbored exon 9b (inserted with pSC-A-Amp/kan) were subjected to restriction digestion using SalI and XbaI (Figure 3-6).
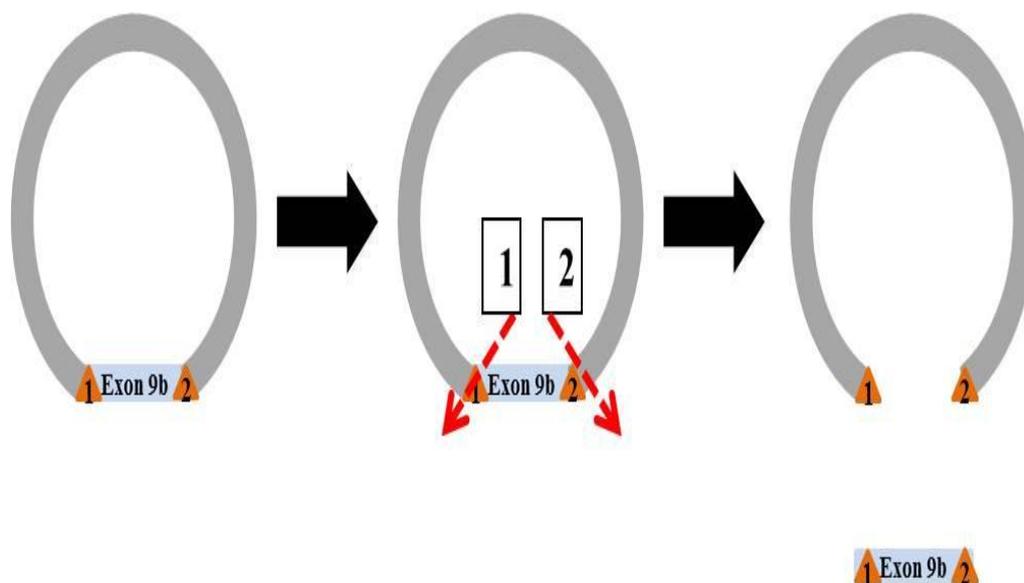
**Figure 3-6: Restriction using SalI & XbaI for minigene constructs.**
Exon trap-exon 9b (or pSC-A-Amp/kan-exon 9b) inserts are shown as grey circle
(exon trap or pSC-A-Amp/kan vector) and blue box (exon 9b). The junctions between
the two parts are SalI and XbaI recognition sites (orange triangles 1 & 2, respectively).
SalI (boxed 1) and XbaI (boxed 2) enzymes function to cut (interrupted red arrow) at
the specific sites, resulting in separating the two parts. Exon trap containing exon 9b
inserts were digested to check that only exon 9b inserts are included.

For the restriction reaction, the DNA targets (up to 1 µg) were added to the
reaction mixture in addition to the enzymes and green buffer (Fermentas Life
Sciences) according to the manufacturer's guidelines. For the exon trap vector
ligation, 2 units of alkaline phosphatase enzyme (Fast AP, Roche, GmbH,
Germany) was added to dephosphorylate the vector ends and prevent self-
religation of the vector. The resulting products were subjected to gel
electrophoresis (as mentioned in section 2.6.5) and according to the expected
product size (452 bp size for short exon 9b fragments, 1313 bp size for long
exon 9b fragments, and 4016 bp size for exon trap vector) were subjected to
gel extraction (as mentioned in section 2.6.7). After minigene construct

preparation, SalI and XbaI enzyme restriction was used as a screening test prior to sequencing.

### 3.3.2.2.2 Ligating the inserts with exon trap:

Following DNA quantification, one of exon 9b alleles (G-G or C-A) and exon trap vector were both added in a ligation reaction to insert the alleles within the vector sequence for subsequent steps. For this purpose, the vector: insert ratio used was 1:3 and was calculated using the following equation:

$\frac{vector\ (ng) \times insert\ size\ (bp)}{total\ vector\ size\ (bp)} \times 3 =$ ng of the insert required for the ligation

reaction.

This equation was used to calculate the insert from both exon 9b fragments required for the ligation, and the results are shown as follows:

$\frac{vector\ (100\ ng) \times insert\ size\ (452\ bp)}{total\ vector\ size\ (4461\ bp)} \times 3 = 30.39$ ng (for exon 9b short

fragment).

$\frac{vector\ (100\ ng) \times insert\ size\ (1313\ bp)}{total\ vector\ size\ (4461\ bp)} \times 3 = 88.29$ ng (for exon 9b long

fragment).

For monitoring purposes, a vector-only ligation reaction was run in parallel to check for the levels of background ligation (due to undigested and/or re-ligated vector molecules). The ligation reaction mixture included 100 ng of the exontrap vector, 1 unit of T4 DNA ligase and 1x ligation buffer (Roche, GmbH, Germany) and the total volume of 20 μl reaction was completed using RNase-free water. The reaction mixture was then incubated at 4° overnight.

### 3.3.2.2.3   Transformation into supercompetent cells:

Following ligation, about 50 ng of ligated DNA product was added into XL1-blue supercompetent cells (Agilent technologies) according to manufacturer's guidelines. This included adding 100 µl of XL1-blue cells on pre-chilled 1.5 ml microcentrifuge tubes. Then about 10 ng of the vector-insert DNA was added to the cells. The mixture was incubated on ice for 30 minutes and then subjected to heat shock at 42° for 45 seconds. The mixture was incubated on ice for another 2 minutes before adding pre-heated SOC (super optimal broth medium with glucose used to enhance the plasmid transformation efficiency) to a final volume of 1 ml. The mixture was then incubated at 37° with horizontal shaking for 60 minutes. Finally, 100 µl of reaction mixture was spread on plates containing pre-warmed CG agar (Circle Grow rich bacterial growth media, Anachem, Luton, UK) (containing 4 g/100 ml CG, 1.5 g/100 ml agar and 50 µg/ml ampicillin and 20% X-gal) and incubated overnight at 37°. Single colonies were inoculated into 5 ml CG bacterial growth media (including 50 µg/ml ampicillin) and incubated for 16-18 hours at 37° in a shaking incubator. On the next day, colonies were selected for purification and sequencing.

The vector-only ligation product was used as a control for monitoring the level of background ligation.

### 3.3.2.2.4  Plasmid-insert extraction:

Using endotoxin-free QIAprep Spin Miniprep kit (QIAGEN), plasmid was extracted from bacterial cultures according to manufacturer's guidelines. The DNA yield was then measured using a Nanodrop spectrophotometer.

### 3.3.2.2.5  Screening for insert-specific products:

To check for inclusion of the insert within the vector, SalI and XbaI (supposed to be located between the inserts and the vector sequences) were used for digesting the extraction products and to screen for the possible samples to send for DNA sequencing (using primers E9b-shortS, E9b-longS, & E9bAS primers, from Table 3-2) (as mentioned in section 2.6.8). The results of the sequencing were compared with the reference sequence of the corresponding exon 9b alleles.

### 3.3.2.2.6  Endotoxin-free midiprep preparation:

After confirming the insert DNA sequence is the correct sequence and is in the correct orientation, the minipreps DNA was further extracted by using QIAfilter Plasmid Midi kit (QIAGEN) according to manufacturer's guidelines.

### 3.3.2.3 Site-directed mutagenesis:

Using the Quick Change II Site-Directed Mutagenesis kit (Agilent technologies, USA) exon 9b generated alleles (G-G & C-A) were used as templates to generate another two alleles (G-A & C-G). First, two primer sets were designed using Agilent software recommended by the manufacturer.

The primer design aimed at targeting a single base within exon 9b, which was either G154 (on G-G allele) or the A154 (on C-A allele). Using this technique, the nucleotide base at position 154 of exon 9b in G-G and C-A minigene constructs (from exon 9b short fragment) were targeted for substitution. This included the generation of two more constructs to cover all four possible alleles of exon 9b (Table 3-4).

**Table 3-4: Site-directed mutagenesis summary for exon 9b minigene constructs.**

| Exon 9b allele | Base at position 154 | Substituted base | Resulting allele |
|---|---|---|---|
| G-G | G | A | G-A |
| C-A | A | G | C-G |

The characteristics and sequences of the primers are summarized in (Table 3-5).

**Table 3-5: primer design recommended guidelines-Agilent technologies.**

| Primer length (bp) | 25-45 |
|---|---|
| Base to be substituted-position | In the middle part of the primer |
| Primer last base | G or C |
| Primer GC contents | > 40% |
| Primers overlap | Both forward and reverse primer should overlap the region of substitution |
| AG-S* | TGTACAACCTCCCCCAGA**C**ACTCTTTGCGAATTTG |
| AG-AS* | CAAATTCGCAAAGAGTG**T**CTGGGGGAGGTTGTACA |
| CG-S‡ | TGTACAACCTCCCCCAG**G**CACTCTTTGCGAATTTG |
| CG-AS‡ | CAAATTCGCAAAGAGTG**C**CTGGGGGAGGTTGTACA |
| *: sequences of primers for generating the AG allele of exon 9b. |  |
| ‡: sequences of primers for generating the CG allele of exon 9b. |  |
| S: forward primer, AS: reverse primer. |  |
| The mutation base is highlighted in bold letter. |  |

The method was used according to the manufacturer's guidelines. This method involves the use of the designed primers in a PCR reaction. The templates used for the PCR included the exontrap vector with exon 9b insert (either G-G or C-A allele). Following amplification of the new (mutated) allele, the used of *Dpn-I* restriction enzyme aimed at degrading the old template and keeping the new template. This was followed by sequencing the mutated-insert and once the mutagenesis was proved successful, the same steps used for preparing the minigene constructs were followed (as mentioned in sections 3.3.2.2.6).

### 3.3.3   Ex vivo splicing testing:

### 3.3.3.1 Transfection:

Transient transfection experiments were applied for the generated constructs in two cell lines: A549 and Cos7 cells (fibroblast-like cells derived from monkey kidney tissue). The latter was used as a comparative cell line (due to its high transfection efficiency) (Cooper, 2005). Interestingly, the Cos7 cells were shown to be among the cells that do not express α7-nAChRs (Cooper and Millar, 1997). Three independent experiments were run for each construct in both cell lines. Empty exon trap vector was used in the transfection mixture compared to transfection with plasmid-inserts constructs. This step was included to confirm that  the origin of the expressed transcripts with exon 9b are due to an exon sequence inserted within the exontrap vector sequence and do not originate from the transfection per se.

### 3.3.3.2 Cell culture:

For A549 and Cos7 cells, the same media and conditions for growing the cells were used (as described in section 2.6.1). 24 hr prior to transfection, the target cells were seeded into 60 mm$^2$ dishes ($7.5 \times 10^5$ of A549 cells/dish and $3.75 \times 10^5$ of Cos7 cells/dish). The cells were tested using a light microscope for viability and confluency before and after each step, and only confluent cells (50-80%) were used for transfection. For each transfection experiment, 1 µg of the construct DNA was mixed with 9 µl of TRANS-FAST transfection reagent (Promega) and 2 ml of serum-free DMEM. After vortexing the mixture, it was incubated for 15 minutes at room temperature to be ready for use with the target cells. Then the old medium was removed from the cells and PBS cell washing was applied before adding the prepared transfection mixture. The latter was added slowly to the cells (to avoid excessive detachments, couple of cells detached were considered acceptable) and this was followed by examining the cells under the microscope for possible detachments. Once checked, the cells were incubated with the transfection mixture for 1 hr at 37°. After incubation, cells were examined before and after adding 4 ml of DMEM slowly. This was followed by incubating the cells for further 24 hr. On the next day, the RNA of transfected cells was prepared to allow for subsequent transcript testing.

### 3.3.3.3 RNA extraction and cDNA generation:

Following 24 hr of transfection, RNA was extracted from A549 and Cos7 cells (as mentioned in section 2.6.2). The extracted RNA was used for generation of cDNA as a template for RT-PCR (as mentioned in section 2.6.3).

### 3.3.3.4 RT-PCR:

Similar protocol and methods were used for RT-PCR mentioned in section

2.6.4, except for the primer sets used (Table 3-6).

Table 3-6: **Nucleotide sequences of the primers used for exon 9b minigene construct sequencing.**

| Primer ID | Primer sequence (5`→3`) | Position |
|-----------|-------------------------|----------|
| **PET01S** | GATCGATCCGCTTCCTG | 5` ET exon |
| **PET01AS** | CGGGCCACCTCCAGTG | 3` ET exon |
| All primers are from Eurogentec, Belgium. ET: exon trap vector, S: Forward primer, AS: reverse primer. PET01S & PET01AS are exontrap primers that anneal to the  5` and 3` exons of the vectors, respectively. | | |

### 3.3.3.5 Confirming the sequence of the amplified transcripts:

Following RT-PCR amplification, the detected products were gel extracted (as mentioned in section 2.6.7) and then sequenced (as mentioned in section 2.6.8) (Figure 3-7).
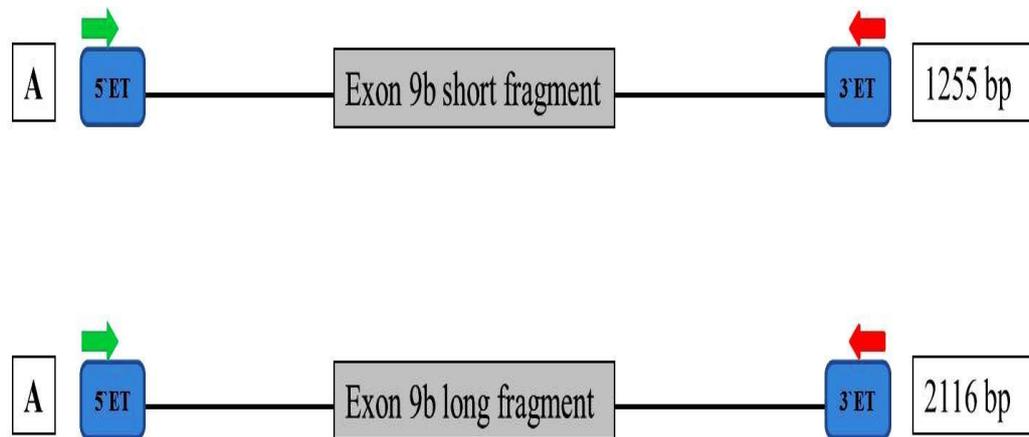


**Figure 3-7: Diagram showing the location of primers used for sequencing exon 9b minigene constructs.**
*A:* exon 9b short fragment. *B:* exon 9b long fragment. Blue boxes: exons (ET: exontrap exons: 5` and 3`), green arrows: forward primers, red arrows: reverse primers, black lines: flanking exontrap intron sequences, the expected sizes of the minigene constructs is shown on the right (boxed). Exon 9b size= 197 bp.

### 3.3.3.6 AlphaDigiDoc software and statistical analysis:

To evaluate the relative percentage of mRNA expression for the exon 9b transcripts, the gel electrophoresis pictures were analysed using AlphaDigiDoc1201 software (similarly to the method mentioned on section 2.6.11). The parameter used for comparisons was the band intensity for each transcript. To determine the relative percentage of expression of each transcript by A549 and Cos7 cells (for each gene transcripts), the ratio of band intensities of exon 9b-included/excluded transcripts was calculated and the correction for the size of products was applied. This was applied by dividing the intensity values of each band by the corresponding product size values (in bp).

## 3.4   Results:

Following the detection of exon 9b transcripts (as mentioned in chapter 2), the sequences of exon 9b within intron 9 of both genes (CHRNA7 and CHRFAM7A) were analysed for detecting exon defining elements (such as the GT-AG splice sites flanking the exon, ESEs and ESSs). This was followed by amplifying exon 9b alleles from a heterozygous control subject (that expresses both alleles) to prepare exon 9b alleles that represent both genes, CHRNA7 and CHRFAM7A. The isolated alleles of exon 9b were then used as template for preparing exon 9b minigene constructs.

### 3.4.1   Bioinformatics analysis:

The detected exon 9b sequence was tested using several bioinformatics software programs. First, using *Exon scan* software, exon 9b sequence (from both genes) could not define exon 9b as a constitutive exon. Similar results were obtained using *Gene scan* software program (Burge and Karlin, 1997). The next step included testing exon 9b sequences from both genes for potential ESE and ESS using *Human splicing finder*, *RESCUE-ESE, ESE finder* and *splicing rainbow* software programs. The reason for using different software programs is that to test for recognition of the two bases in exon 9b (at positions 77 and 154) to lie within ESEs or ESSs by one or more programs. A SNP check was applied using *dbSNP* and *Ensembl* databases. Finally, *ClustalW* software program was used for alignment of sequences from CHRNA7 and CHRFAM7A genes with those detected from transcripts of exon 9b.

### 3.4.1.1 Human splicing finder results:

Using *Human splicing finder* program, the potential splice sites, branching point, predicted ESEs (PESE), predicted ESSs (PESS), PESE/PESS ratio, enhancer and silencer elements within the tested sequence were predicted. The splice donor, splice acceptor and the branch point sites where detected at exactly the same positions in relation to exon 9b (Table 3-7 & Table 3-8).

**Table 3-7: Summary for Human splicing finder results of exon 9b.**

|  | G-G allele | C-A allele |
|---|---|---|
| **splice donor site position**[*] | -12 | -12 |
| **splice acceptor site position**[*] | +2 | +2 |
| **Branch point position**[*] | -31 | -31 |
| **PESE/PESS**[1] | 2.75 | 2.63 |
| **PESE/PESS**[2] | 1 | 0.98 |
| [*]: position in relation to exon 9b (-: refer to upstream sequences,+: refer to exon 9b sequences), PESE[1]: using exon 9b short fragment, PESE[2]: using exon 9b long fragment. | | |

**Table 3-8: Collective results of Human splicing finder results (experimental and other In-Silico analysis web-based) for analysis of exon 9b alleles with their flanking 100 bp inronic sequences (representing short exon 9b fragments).**

| In-Silico analysis | E9b G-G allele | | E9b C-A allele | |
|---|---|---|---|---|
| **-** | Position 77 | Position 154 | Position 77 | Position 154 |
| **RESCUE-ESE (Fairbrother et al., 2002)** | ctgaaG, tgaaGt | - | gaaCtg | - |
| **PESE (Zhang, 2004)** | cctgaaGt | - | ctgaaCtg | - |
| **EIE (Zhang et al., 2008)** | ctgaaG, gaaGtg | - | ctgaaC | cccagA, ccagAc, cagAca |
| **ESE motif from HSF- (SR 9G8)** | gaaGtg | - | - | ccagAc |
| **Silencer motif (Sironi et al., 2004)-motif2** | tgaaGtgc | - | - | - |
| **IIE (Zhang et al., 2008)** | tgaaGt, Gtgctt | - | - | - |
| **ESR elements (Goren et al., 2006)** | ctgaaG, aGtgct | cccagG | - | - |
| PESE=predicted ESE, EIE=exonic identity elements, IIE=inronic identity elements, ESR: exonic splicing regulating. Underlined= Silencing sequences, upper case letters: represents corresponding exon 9b base (77 or 154), lower case letters: flanking exon 9b sequence | | | | |

These results showed that the splice sites and branch points were identical for both exon 9b alleles. However, the predicted ESE/ESS ratio was higher for G-G allele in both short and long exon 9b fragments.

### 3.4.1.2 RESCUE-ESE results:

RESCUE-ESE results showed that exon 9b sequence from CHRNA7 (G-G allele) would harbour 16 ESEs while the presence of C allele at position 77 (C-A allele from CHRFAM7A) would predict one less ESE at this position making the total of 15 ESEs within exon 9b C-A allele (Figure 3-8).
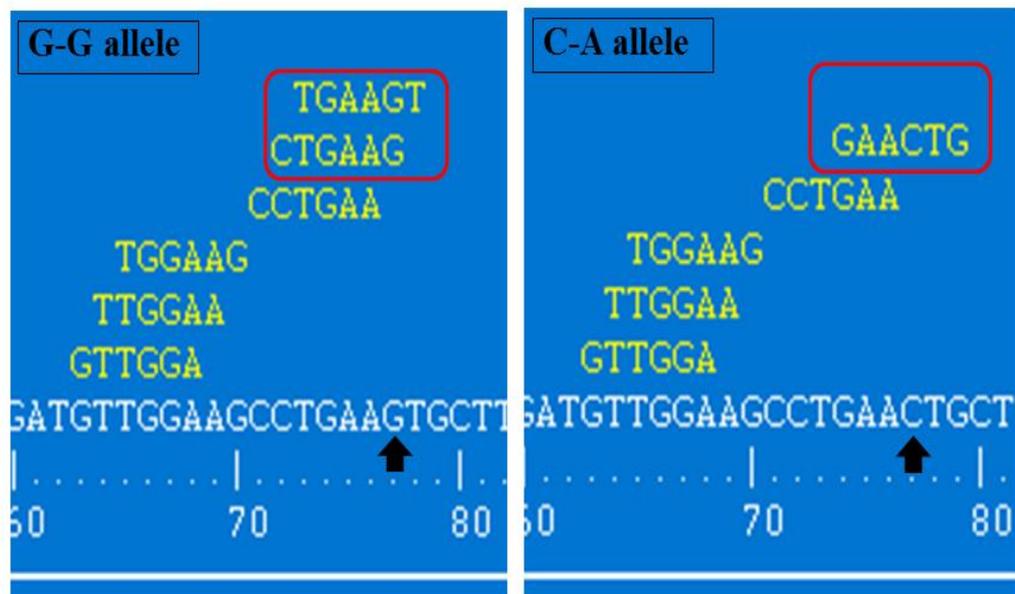


**Figure 3-8: RESCUE-ESE test for exon 9b sequences.**
Exon 9b sequences from CHRNA7 (G-G allele) and CHRFAM7A (C-A allele) were subjected to RESCUE-ESE analysis. At position 77 (arrowed) two potential ESEs for G-G allele were compared to single different ESE for C-A allele (within red boxes). The rest of the predicted ESEs were the same in both sequences (data not shown).

**3.4.1.3 ESE-finder results:**

For further test of ESE effects, *ESE finder* program was used. The results expressed possible potential binding of five SR proteins to ESE sequences within exon 9b (Table 3-9) (Figure 3-9).

**Table 3-9: ESE finder results for exon 9b alleles.**

| SR proteins (threshold value) | Exon 9b G-G allele | | Exon 9b C-A allele | |
| --- | --- | --- | --- | --- |
| | position77 | Position 154 | position77 | Position 154 |
| **SF2/ASF (1.956)** | ctgaaGt (2.3) | cccagGc (2.05) | - | - |
| **IgM    BRCA1 (1.867)** | ctgaaGt (2.44) | cccagGc (3.0) | - | ccccagA (2.4) |
| **Sc35 (2.383)** | - | - | - | - |
| **SRp40 (2.67)** | - | ccccagG (3.5) | - | - |
| **SRp55 (1.676)** | - | - | - | - |
| Sequence motifs for potential binding with SR proteins (shown on the left) are indicated for sequences at positions 77 & 154 of exon 9b from the two alleles (indicated). The threshold values (between brackets) are indicated for each motif detected and are considered 'high' when they are greater than the threshold value. Upper case letters: represents corresponding exon 9b base (77 or 154), lower case letters: flanking exon 9b sequence | | | | |
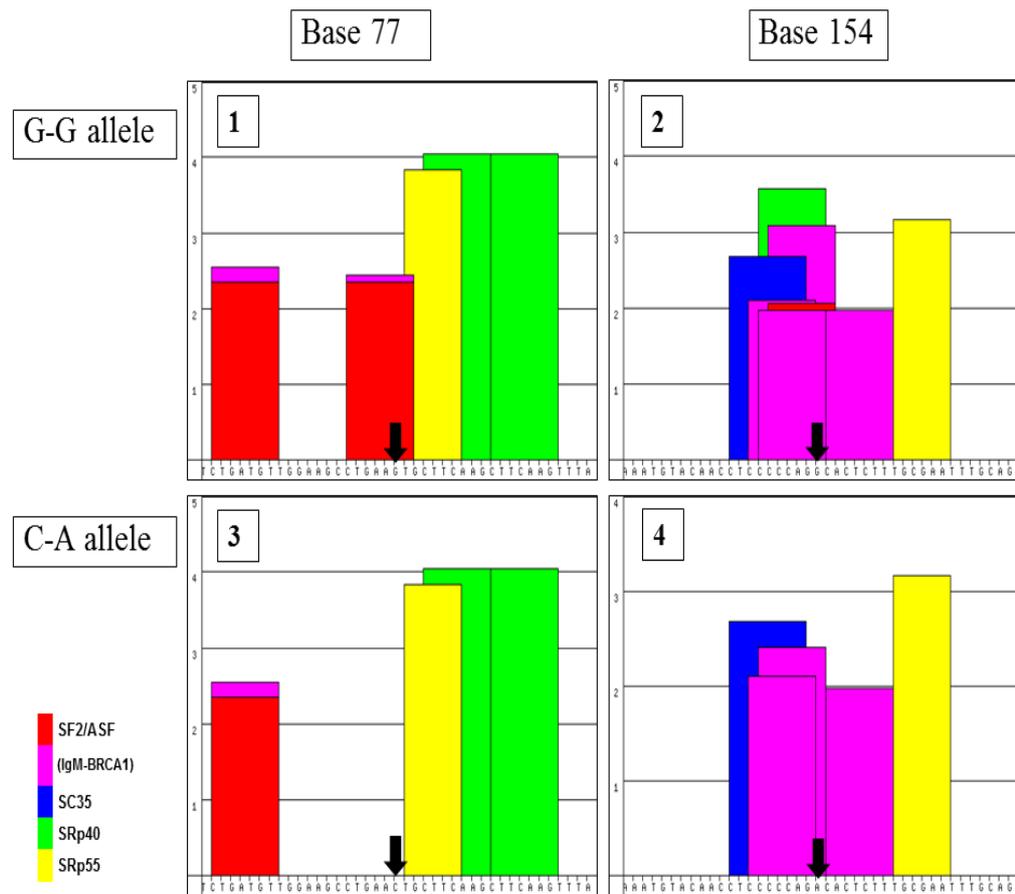
**Figure 3-9: ESE finder results for exon 9b alleles.**
Exon 9b alleles are shown on the left, and the two bases at positions 77 & 154 are indicated above (and arrowed). Potential SR protein binding (colour indicator is shown on left lower corner) to the alleles (exon 9b position) are: 1) G allele (77): for SF2/ASF & IgM BRCA1, 2) G allele (154) for SF2/ASF, IgM BRCA1 & SRp40, 3) C allele (77): for none; and 4) A allele (154) for IgM BRCA1. The nucleotide sequence tested is shown on the X-axis. The height of the bars indicate score values (represented on the Y-axis) while the width of the bars represent their nucleotide length and position within a given sequence (Cartegni et al., 2003).

### 3.4.1.4 Splicing rainbow results:

Similarly, Splicing rainbow program was used to test for the SR proteins and the splicing inhibitors (hnRNPs) potential binding sites within exon 9b allele sequences (Figure 3-10 & Figure 3-11).
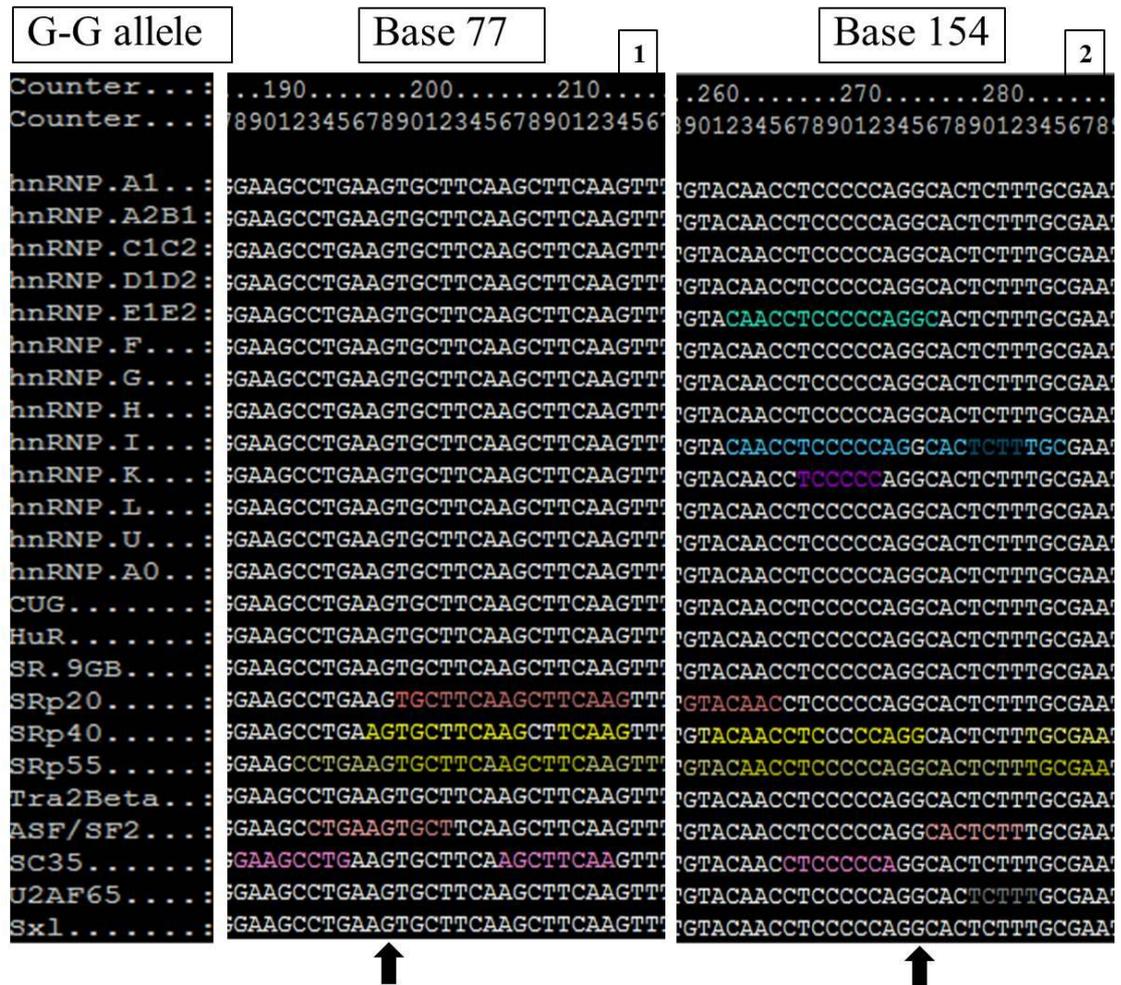
**Figure 3-10: Splicing rainbow results for exon 9b G-G allele.**
Bases at positions 77 & 154 of exon 9b are indicated on top and arrowed on bottom.

Potential SR protein binding sites for 1) G allele (77): SRp40, SRp55 & ASF/SF2, 2)

G allele (154): SRp40 & SRp55, Potential splicing inhibitor factors (hnRNP) binding

for 2) G allele (154): E1E2.
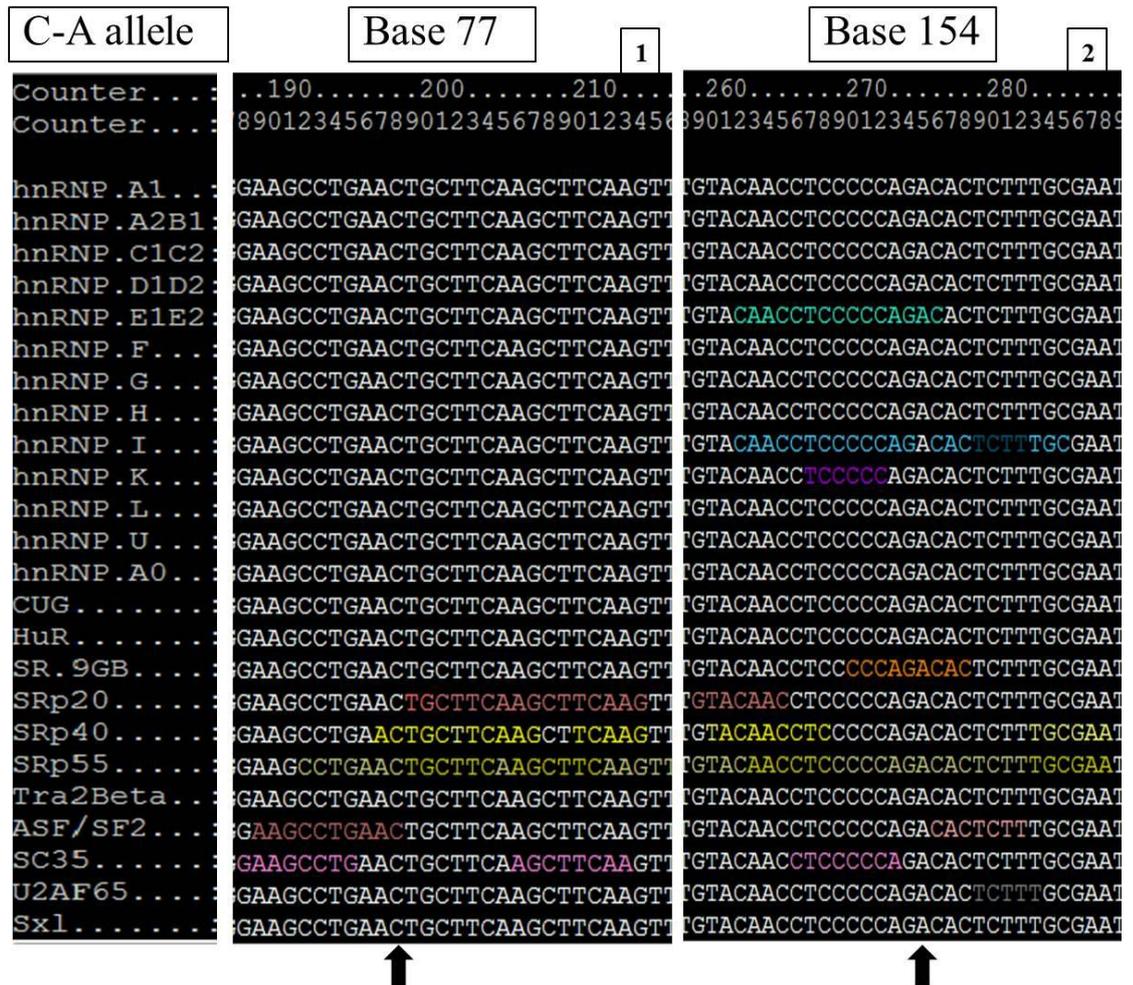
**Figure 3-11: Splicing rainbow results for exon 9b C-A allele.**

Bases at positions 77 & 154 of exon 9b are indicated on top and arrowed on bottom.

Potential SR protein binding sites for 1) C allele (77): SRp40, SRp55 & ASF/SF2, 2)

A allele (154): SRp55 & 9GB. Potential splicing inhibitor factors (hnRNP) binding for

2) A allele (154): E1E2.

A summary of the sequence motifs potential binding sites for enhancer and silencer elements at positions 77 and 154 of exon 9b from all used *In-Silico* analysis tools is represented in (Figure 3-12).
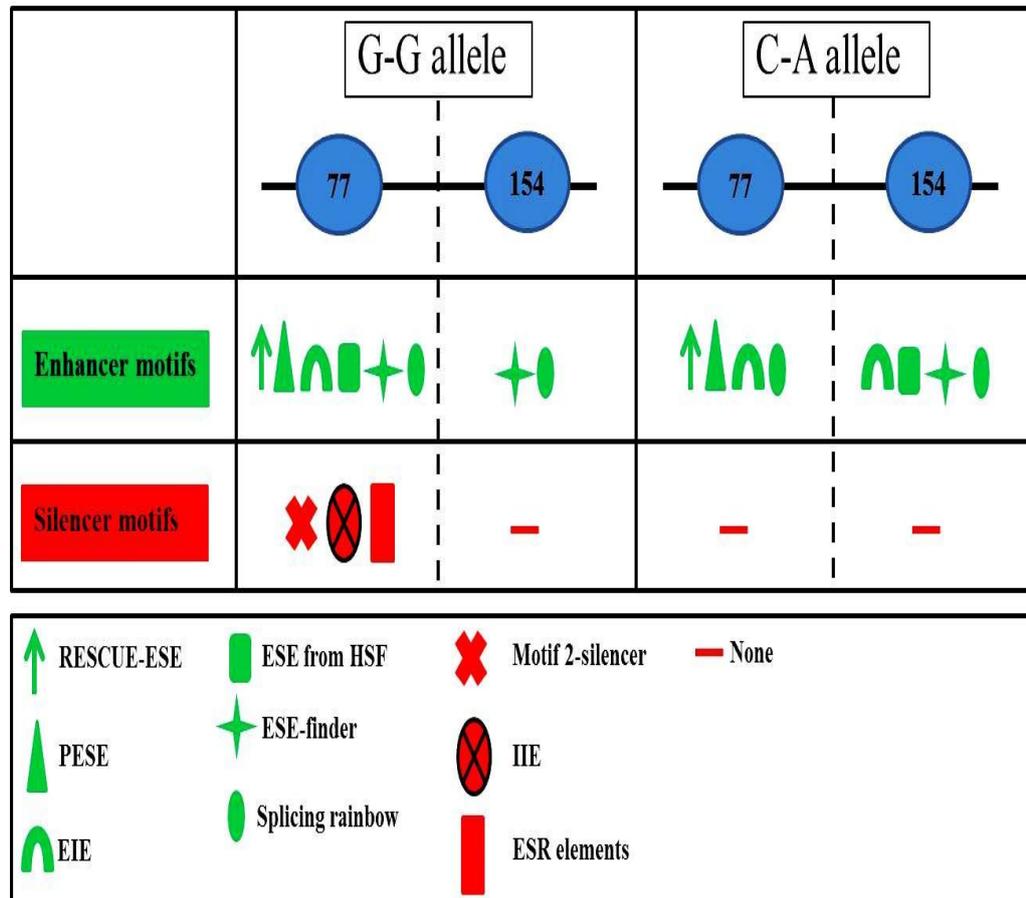


**Figure 3-12: Summary for enhancer and silencer potential sites at positions 77 and 154 of exon 9b.**
Exon 9b alleles are indicated above (G-G & C-A). Potential enhancer motif sites are shown in green shapes, while potential silencing sites are shown in red shapes. Human splicing finder web-based multi-software analysis tool included results for predicted enhancer motifs (RESCUE-ESE, PESE, EIE, and ESE from HSF & ESE-finder) and predicted silencer motifs (motif-2 silencer, IIE, PESS & ESR elements). In addition, splicing rainbow results for potential enhancer motifs are shown. The results are shown in table form to show enhancer or silencer sequence motifs that included each allele of exon 9b (at position 77 and 154).

**3.4.1.5 Screening for SNPs:**

To screen for possible SNPs located within exon 9b, *Ensembl, dbSNP* and *Hapmap* project databases were used. The two nucleotide base difference between the two genes exon 9b sequences were found to correspond to predicted positions for SNPs but with no genotype reported (rs286097 G/C and rs286098 G/A, for exon 9b bases at positions 77 & 154, respectively). However, very little is known about these two positions from the databases. These data indicates that each of the two positions was detected in few European individuals (Table 3-10).

**Table 3-10: dbSNP data for the two-nucleotide bases on exon 9b (at positions 77 & 154, respectively).**

| SNP detected | Genotype detected | Individuals tested |
|---|---|---|
| rs286097 | CG/CC/CC | 3 |
| rs286098 | AA/AG | 2 |

From the mentioned data above, it is evident that little is known about these two nucleotide bases and their ancestral origin is unknown. According to these results and as was shown previously (Makoff and Flomen, 2007, Flomen et al., 2008, Araud et al., 2011) that an individual might have a single or rarely double deletion of CHRFAM7A or rarely a single deletion of CHRNA7 (as mentioned in section 1.3.4.1), we concluded the following genotypes for exon 9b in a given population:

*A)* Individuals without SNPs affecting the two-nucleotide bases of exon 9b (at positions 77 & 154) would have double heterozygous alleles for exon 9b. This means (according to the reference sequence databases from *NCBI*) that individuals without SNPs affecting exon 9b bases at the above mentioned positions would have the G/C and G/A phenotypes at positions 77 and 154, respectively.

*B)* Individuals with a SNP affecting one of the two-nucleotide bases of exon 9b would have a single heterozygous genotype.

*C) & D)* Individuals missing both copies of CHRFAM7A or CHRNA7 would have single heterozygous genotype, respectively (Table 3-11).

**Table 3-11: Possible exon 9b allele genotypes in individuals:**

| Gene | Allele sequence (in exon 9b) | | |
|---|---|---|---|
| CHRNA7 copies | CHRFAM7A copies | At position 77 | At position 154 |
| 2 (or rarely 1) | 2 (or 1) | G/C | G/A |
| 2 (or rarely 1) | 0 (rarely) | G | G |
| 0 (rarely) | 2 (or 1) | C | A |
| This indicated that when an individual have the two genes, both exon 9b alleles could be detected (G-G and C-A). However, for individuals with only one of the gene, the corresponding exon 9b allele would be detected on genotyping exon 9b. | | | |

This indicates that, when both genes are present on chromosome 15, sequencing of exon 9b will show a double heterozygous peaks at positions 77 and 154 (G/C and G/A respectively).

**3.4.1.6 Differences between intronic sequences flanking exon 9b:**

This was tested using *ClustalW* software program. The sequences of intron 9 part upstream of exon 9b that were included in exon 9b fragments (short and long) were used for this purpose. The analysis detected three sites of differences in the sequence of intron 9 between the two genes and was given the sign (-) to indicate their upstream position in relation to exon 9b. These differences included single base at position -900, five bases at position -837 to -833, and at position -122 (Table 3-12) (Figure 3-13). However, when control subjects were tested for recognition of these differences (by using RT-PCR and sequencing) it showed that they actually do not exist and that there were no differences within intronic sequences from both genes at these positions (will be shown later).

**Table 3-12: The sequence differences between exon 9b and flanking intronic parts of long/short exon 9b fragments.**

| Gene/position | Exon 9b allele | Intron upstream of exon 9b | | |
|---|---|---|---|---|
| **Position**[1] | 77-154‡ | -900* | -837 to -833* | -122* |
| **CHRNA7** | G-G | A | TAAGA | A |
| **CHRFAM7A** | C-A | G | - | G |

[1]: position of nucleotide base related to exon 9b, *: nucleotide bases within intron 9 sequence (- : refers to upstream of exon 9b), ‡: nucleotide bases within exon 9b sequence. (CHRNA7 accession number NM_00074 and CHRFAM7A accession number NM_139320.1, NCBI).
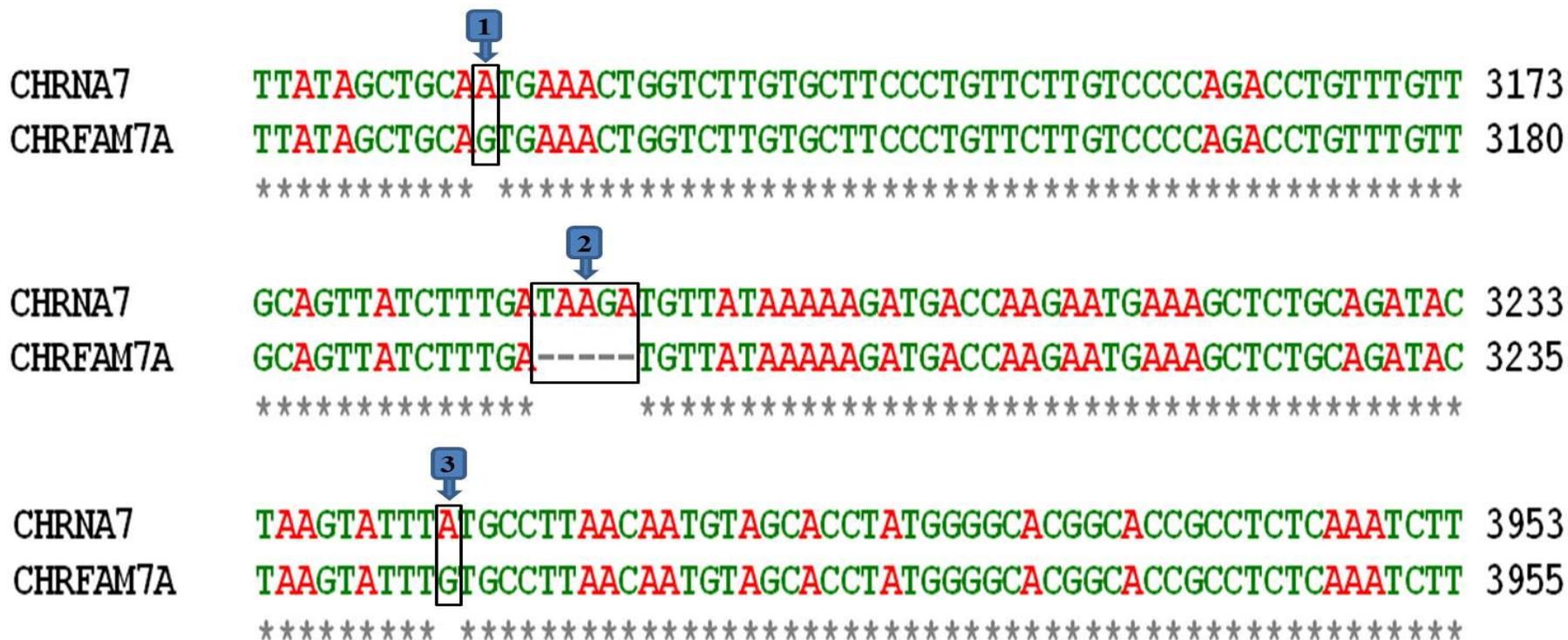
**Figure 3-13: Sequence alignment of long exon 9b fragment from both genes.**
Long exon 9b fragment corresponding sequence from intron 9 of CHRNA7 and CHRFAM7A was subjected to sequence alignment using Clustal W software.
Three differences were detected upstream of exon 9b: 1: at position -900, 2: at positions -837 to -833, 3: at position -122.
Numbers on the right side indicate the sequence within intron 9.

### 3.4.2    Preparation of minigene constructs:

### 3.4.2.1 Screening genomic DNA for heterozygous exon 9b:

For this purpose, "short exon 9b" primers were used using PCR. Screening of

11 control subjects, in addition to the genomic DNA of A549 and BE (2)-c

cells, showed the detection of exon 9b in all of the subjects (Figure 3-14).
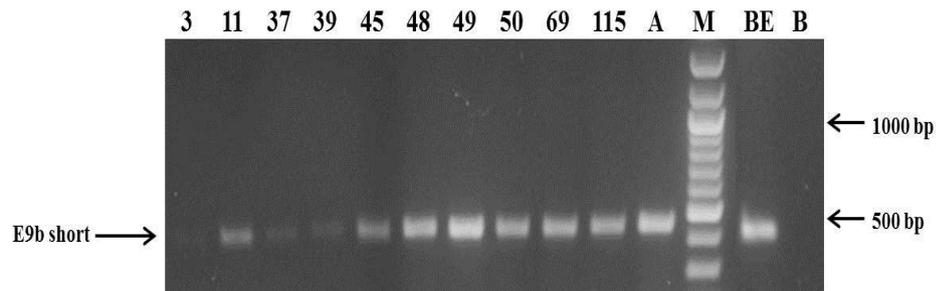


**Figure 3-14: Agarose gel electrophoresis for short exon 9b fragment amplification using RT-PCR.**
Testing genomic DNA samples for short exon 9b fragment (expected size =452 bp);
*Lanes 3-115:* DNA of control subjects, *A:* DNA of A549 cells, *BE:* DNA of BE (2)-c cells. *M:* 100 bp DNA ladder, *B:* blank.

The PCR products were sequenced and the results showed that control subjects

48 & 50, in addition to A549, were of double heterozygous genotype (for exon

9b bases at position 77 and 154), while all the other samples tested showed a

single homozygous genotype (at position 77) and with only BE (2)-c cells

showing double homozygous genotype (at both positions) (Table 3-13).

**Table 3-13: sequencing results for exon 9b genotype of the tested samples.**

| Exon 9b allele genotype | (G/C)* & (G/A)‡ | C* & (G/A)‡ | C*/A‡ |
|---|---|---|---|
| **Samples** | 48, 50, A549 | 3, 11, 37, 39, 45, 49, 69, 115 | BE (2)-c |
| Nucleotide base of exon 9b at position: * (77) and ‡ (154). | | | |

These results showed that BE (2)-c cells showing only exon 9b of C-A genotype that is matching that of CHRFAM7A. This is different from our earlier results where we could detect CHRNA7 more than CHRFAM7A in these cells (as mentioned in section 2.7.3). Thus, the detection of the C-A allele in BE (2)-c cells might suggest possible nucleotide polymorphisms within the two bases of exon 9b from CHRNA7 (G→C at position 77 and G→A at position 154).

Control subject 48 was selected for subsequent analysis to generate the two alleles of exon 9b: G-G and C-A.

Similarly, "long exon 9b fragment" was tested in controls 48, 49, A549 and BE (2)-c cells (Figure 3-15).
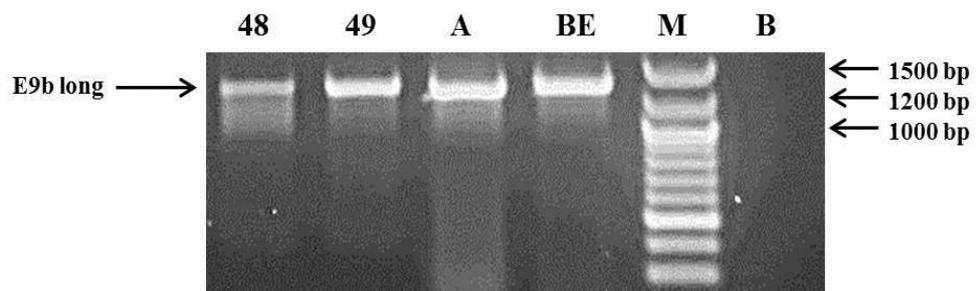


**Figure 3-15: Agarose gel electrophoresis for long exon 9b fragment amplification using RT-PCR.**
Testing genomic DNA samples for long exon 9b fragment (expected size =1313 bp); *Lanes 48-49:* DNA of control subjects, *A:* DNA of A549 cells, *BE:* DNA of BE (2)-c cells. *M:* 100 bp DNA ladder, *B:* blank.

These products were sequenced using exon 9b long primer set and the results are summarized in (Table 3-14), (Figure 3-16) and (Figure 3-17).

**Table 3-14: Sequencing summary for the DNA samples tested.**

| DNA tested | Intron upstream of exon 9b | | | Exon 9b genotype | Expected gene of origin |
|---|---|---|---|---|---|
| **Position[1]** | -900[*] | -837 to -833[*] | -122[*] | 77 & 154[‡] | - |
| **Subject 48** | A | TAAGA | A | G/C & G/A | CHRNA7 & CHRFAM7A |
| **Subject 49** | A | TAAGA | A | C & G/A | CHRNA7 & CHRFAM7A |
| **A549** | A | TAAGA | NA | G/C & G/A | CHRNA7 & CHRFAM7A |
| **BE (2)-c** | A | TAAGA | NA | C &   A | CHRNA7 & CHRFAM7A |

[1]: position of nucleotide base related to exon 9b, [*]: nucleotide bases at intron 9 sequence (-: refer to upstream of exon 9b), [‡]: nucleotide bases within exon 9b sequence. (CHRNA7 accession number NM_00074 and CHRFAM7A accession number NM_139320.1, NCBI). No differences were found between the two alleles of exon 9b, mentioned in the reference sequence databases mentioned above, at positions -900, -837 to -833 and -122 of intron 9 upstream of exon 9b.

NA: non-applicable comparison, as the sequences at this position were not amplified or sequenced.
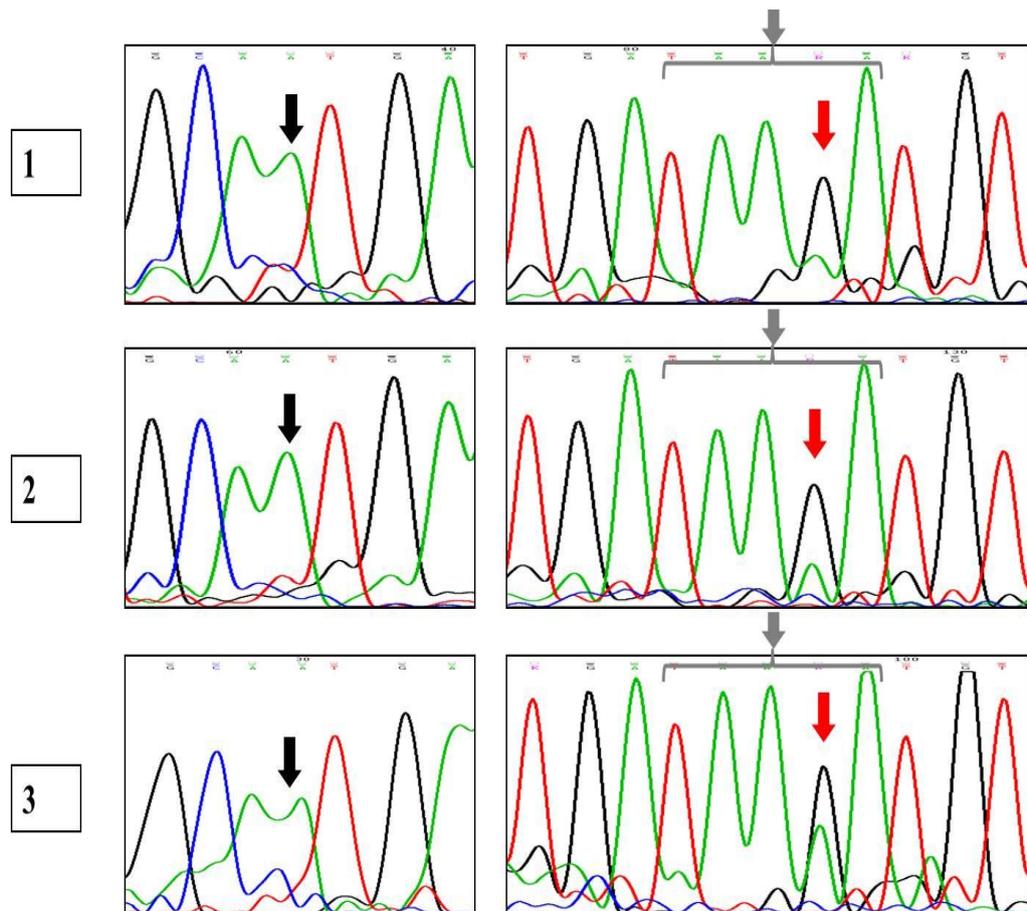
**Figure 3-16: Sequencing result for intronic sequences upstream exon 9b.**
Sequencing for intronic sequences upstream of exon 9b (Black arrow: nucleotide base at position -900, grey arrow: nucleotide bases at positions -837 to -833). Genomic DNA tested from: (1) A549 cells, (2) BE (2)-c cells and (3) control subjects 48 & 49. The sequences at these positions were A (at -900) and TAAGA (at positions -837 to -833) which match the CHRNA7 database reference sequence. Base colour reference: A=green, T=red, G=black, C=blue. Note that an apparent A peak in parts 2) and 3) is due to the flanking high peaks of A surrounding the G peak in the middle

Surprisingly, these results were different from what was expected based on the reference sequences used while designing the primers (Figure 3-13). This meant that both genes, CHRNA7 & CHRFAM7A, had no differences at positions suggested by the reference sequence.
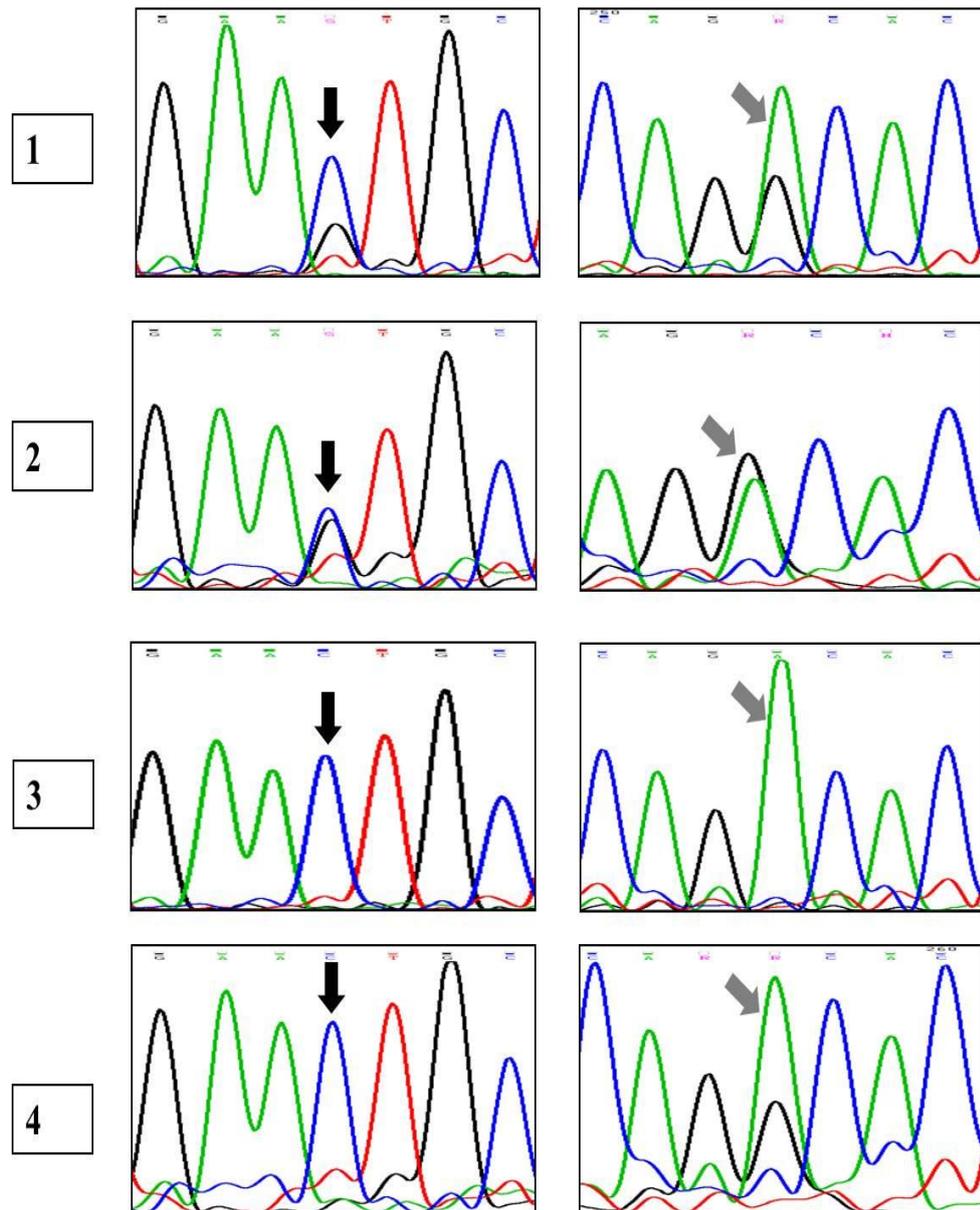
**Figure 3-17: sequencing results for exon 9b sequences.**
Sequencing for exonic sequences of exon 9b (Black arrow: nucleotide base at position 77, grey arrow: nucleotide base at position 154). Genomic DNA tested from: (1) A549 cells, (2) subject 48, (3) BE (2)-c cells and (4) other subjects (including subject 49). A549 and subject 48 showed double heterozygous expression at positions 77 & 154 of exon 9b, single heterozygous expression (at position 154 of exon 9b) could be detected in other subjects, while homozygous genotype was detected from BE (2)-c cells. Base colour reference: A=green, T=red, G=black, C=blue.

The sequencing results could not detect the sequence difference between the two genes (in the intron upstream of exon 9b) at positions -900 and -837 to -833 (upstream of exon 9) as shown in the reference sequence (Table 3-12) (Figure 3-13). Meanwhile, regarding exon 9b sequence, A549 cells and subject 48 showed double heterozygous genotype (with no possible polymorphic changes at either exon 9b bases) while BE (2)-c cells showed homozygous genotype (possibly due to polymorphisms at both exon 9b bases) and other subjects showed only single heterozygous genotype (at position 154) (possibly due to single polymorphism at position 77) (Figure 3-17).

### 3.4.2.2  TA cloning for exon 9b PCR products:

Exon 9b fragments (long and short) from control subject 48 (that were amplified using PCR (Figure 3-14 & Figure 3-15)) were inserted into the pSC-A-Amp/Kan vector using TA cloning technique. This was followed by screening for G-G & C-A alleles constitution prior to extraction using the restriction enzyme Eco57I. The most striking digested fragment for G-G allele (helped to differentiate it from C-A allele) was of 150 bp size.

This screening test was followed by sequencing several G-G and C-A alleles of exon 9b fragments (short and long) throughout the whole fragment sequence to ensure the compatibility of the fragments with the reference sequences and that the only difference between the two constructs was the two bases at positions 77 and 154. The sequence of the clones matched exactly the reference sequence that were used for minigene construct subsequent preparation.

### 3.4.2.3 SalI and XbaI restriction:

The restriction digestion using SalI and XbaI showed expected band size of corresponding exon 9b fragment (short=452 bp, long =1313bp) (Figure 3-18).
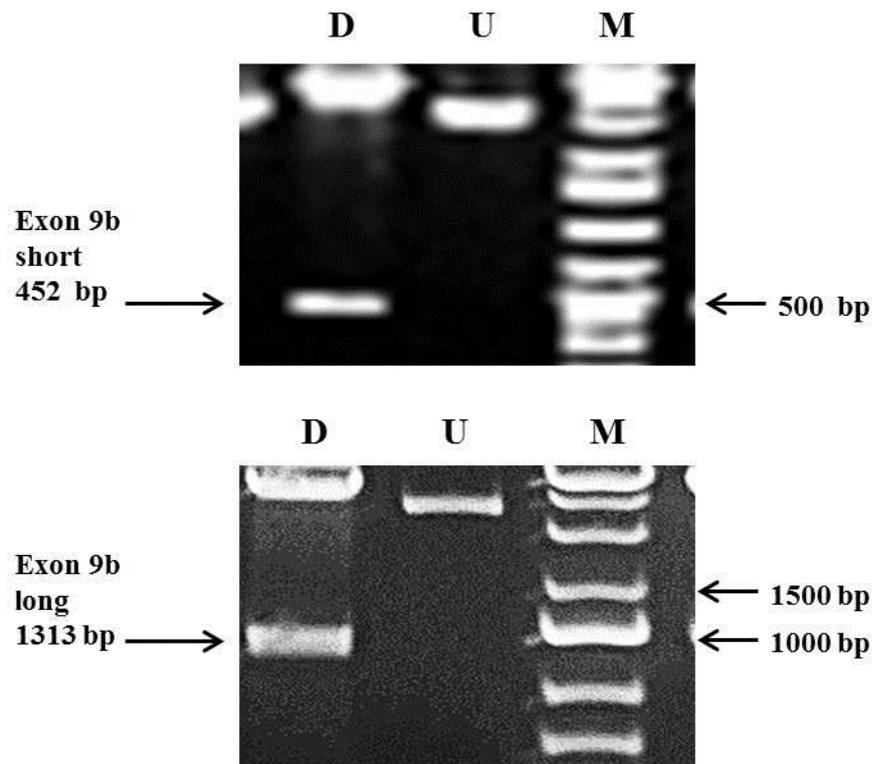


**Figure 3-18: Agarose gel for restriction digestion using SalI & XbaI for exon 9b short & long fragments.**
*Upper panel:* exon 9b short fragment (452 bp size product-arrowed), *Lower panel:* exon 9b long fragment (1313 bp size product-arrowed). D: digested DNA, U: undigested DNA. Similar results were detected for both exon 9b alleles. M: Upper panel= 100 bp DNA ladder, lower panel=1kb DNA ladder.

Once the sequence was confirmed, each allele was subjected to further DNA extraction step using midiprep system to generate minigene constructs in amounts required for the transfection step.

### 3.4.2.4 Site directed mutagenesis:

Once prepared, these constructs were first confirmed to match the reference sequence target (except for the substituted base) by sequencing, and then they were subjected to midiprep system for generating extra amounts for transfection experiments (Figure 3-19).
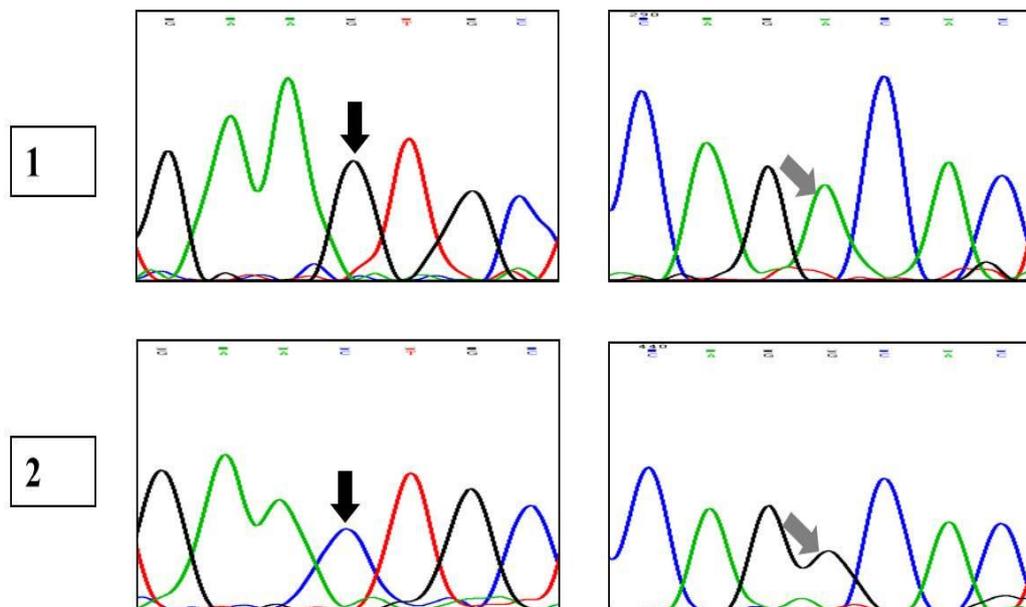


**Figure 3-19: sequencing results for minigene constructs generated by site-directed mutagenesis.**
Sequencing at positions 77 & 154 of exon 9b (Black arrow& grey arrow, respectively). 1) G-A allele, 2) C-G allele. Base colour reference: A=green, T=red, G=black, C=blue.

### 3.4.3   Transfection:

Following preparation of all four minigene constructs of exon 9b, they were used for transfection experiments (Figure 3-20).
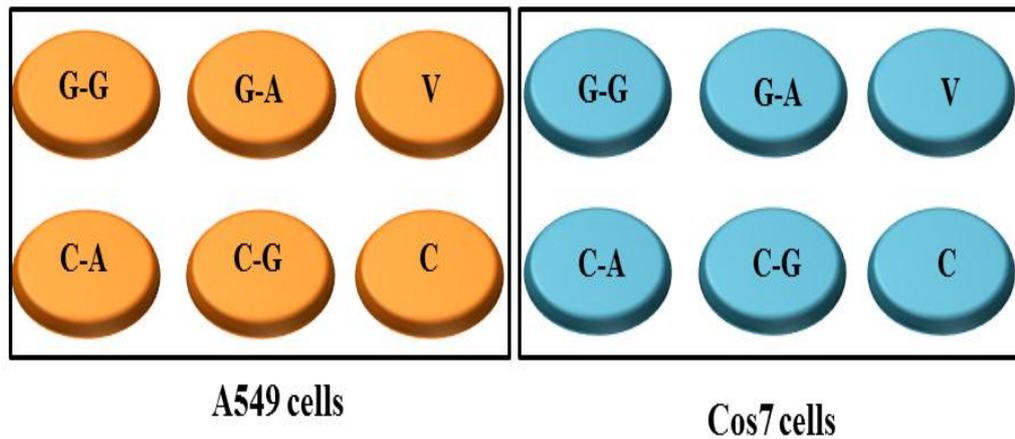
**Figure 3-20: Summary for transfection of minigene constructs of exon 9b.**
A549 and Cos7 cells were transfected using the four minigene constructs prepared (G-G, C-A, G-A & C-G), in addition empty vector (V) and non-transfected cells as a control (C).

### 3.4.3.1 Testing for transfection results:

For this purpose, RNA was extracted from the transfected cells 24 hr following transfection start. This was followed by cDNA preparation to use as a template for RT-PCR with exon trap primers (Table 3-6).

### 3.4.3.2 RT-PCR results:

During the initial runs, it was noticed that the results prediction was affected by the position of the PCR product applied in the electrophoresis tank as more ethidium bromide seemed to be concentrated in the upper half of the tank affecting the brightness of the products bands compared to the lower half with less ethidium bromide and hence less brightness of all bands that affected the AlphaDigiDoc analysis results (Figure 3-21).
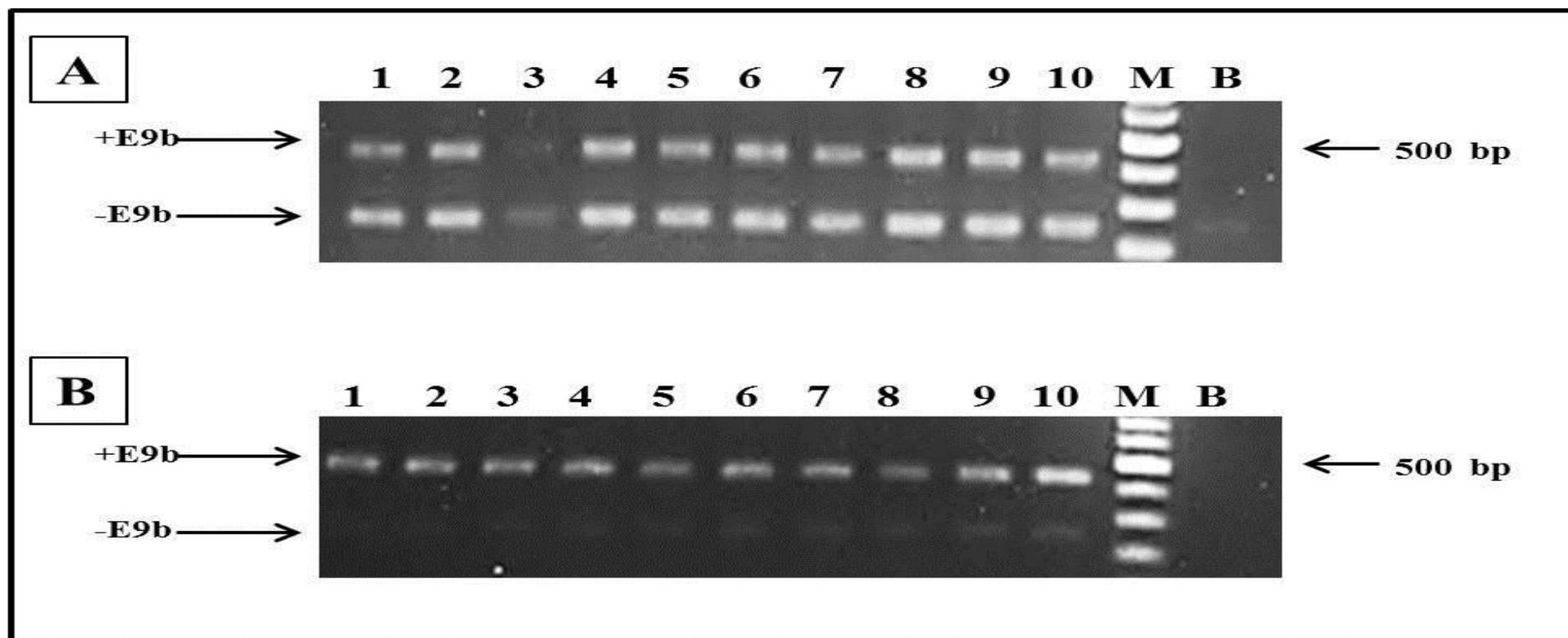
**Figure 3-21: Agarose gel electrophoresis for optimising RT-PCRs using exon trap primers.**
Two alleles were compared (G-G and C-A) from two exon 9b fragments minigene constructs short. These constructs were transfected into A549 cells. Product size (arrowed): exon 9b-included transcripts (+E9b) =439 bp and exon 9b-excluded transcripts (-E9b) =242 bp.  10 replicates of each allele were tested using exon trap primers in RT-PCR. A) Testing G-G allele (in the upper half of electrophoresis tank using the upper sample row) & B) Testing C-A allele (using the lower half and row). M: 100 bp DNA ladder, B: blank.

It is clear from these results that the upper half of the gel picture seems more bright and thus might affect the AlphaDigiDoc analysis as the lower row showed only very intensely stained bands (exon 9b including transcripts) compared to the bottom bands (missing exon 9b). This led to the preference of using only the upper row of samples for gel electrophoresis for all the subsequent analysis to help maintain the consistency of the results.

First, exon 9b long and short fragments minigene constructs (G-G & C-A) were tested for exon 9b inclusion in A549 and Cos7 cells in three independent experiments (Figure 3-22).
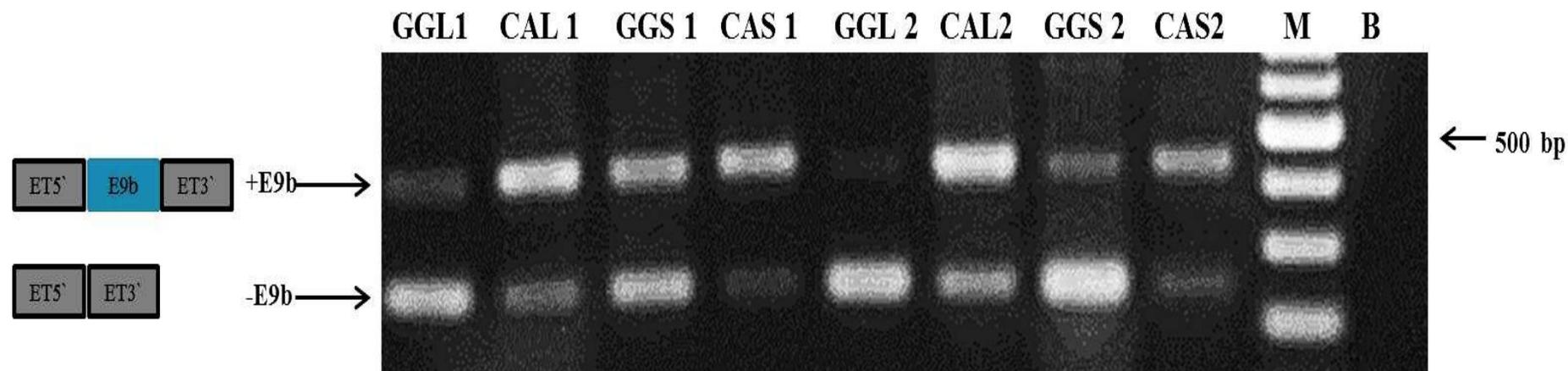
**Figure 3-22: Agarose gel electrophoresis for first RT-PCRs using exon trap primers.**
Two alleles were compared (G-G (GG) and C-A (CA)) from two exon 9b fragments minigene constructs: long (L) and short (S). These constructs were transfected into two cell lines: A549 (1) and Cos7 (2). Product size (arrowed): exon 9b-included transcripts (+E9b) =439 bp and exon 9b-excluded transcripts (-E9b) =242 bp. ET: exontrap exons 5` & 3`. M: 100 bp DNA ladder, B: blank.

Exon 9b was included in all PCR products but the C-A allele constructs in both cell lines showed more remarkable expression of exon 9b transcripts than the G-G allele constructs. It is worth noting that there was no noticeable difference that could affect the expression of exon 9b between the exon 9b short and long fragments.

The next step included testing exon 9b short minigene constructs (G-G, C-A, G-A & C-G) to test the effect on these nucleotide base changes on exon 9b inclusion. These minigene constructs were tested in two cell lines (A549 and Cos7 cells) in three independent experiments (Figure 3-23).
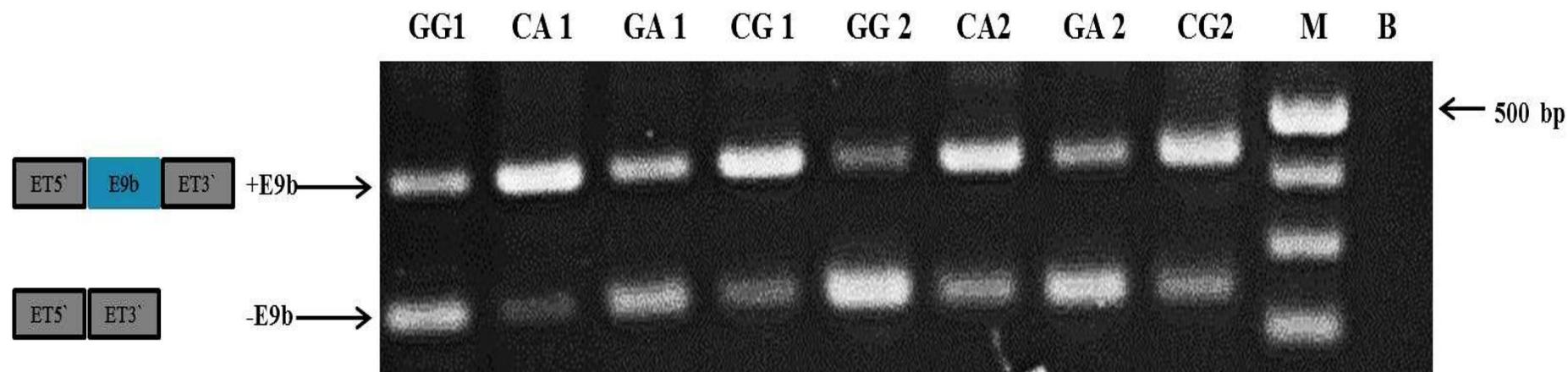
**Figure 3-23: Agarose gel electrophoresis for second RT-PCRs using exon trap primers.**
Four alleles were compared (G-G (GG), C-A (CA), G-A (GA) & C-G (CG)) from exon 9b short fragment minigene constructs. These constructs were transfected into two cell lines: A549 (1) and Cos7 (2). Product size (arrowed): exon 9b-included transcripts (+E9b) =439 bp and exon 9b-excluded transcripts (-E9b) =242 bp. ET: exontrap exons 5` & 3`. M: 100 bp DNA ladder, B: blank.

Again, exon 9b was included in all PCR products but the C-A and C-G allele constructs showed more remarkable expression of exon 9b transcripts than the other two alleles (G-G and G-A).

All the products from the RT-PCR were first gel extracted and then were sequenced and all the results came back matching the reference sequences expected.

Following transfection of the four exon 9b short fragment constructs, the same cDNA extracted from A549 cells was tested for CHRNA7 and CHRFAM7A main transcripts. This test aimed at testing a possible effect of transfection on expression of the two main transcripts for both genes (full length and missing exon 9). The results matched the previous results obtained for A549 cells, showing a relatively similar ratio of expression of full length: missing exon 9 for both genes, and a much more CHRFAM7A compared to those of CHRNA7 under the same conditions (Figure 3-24).
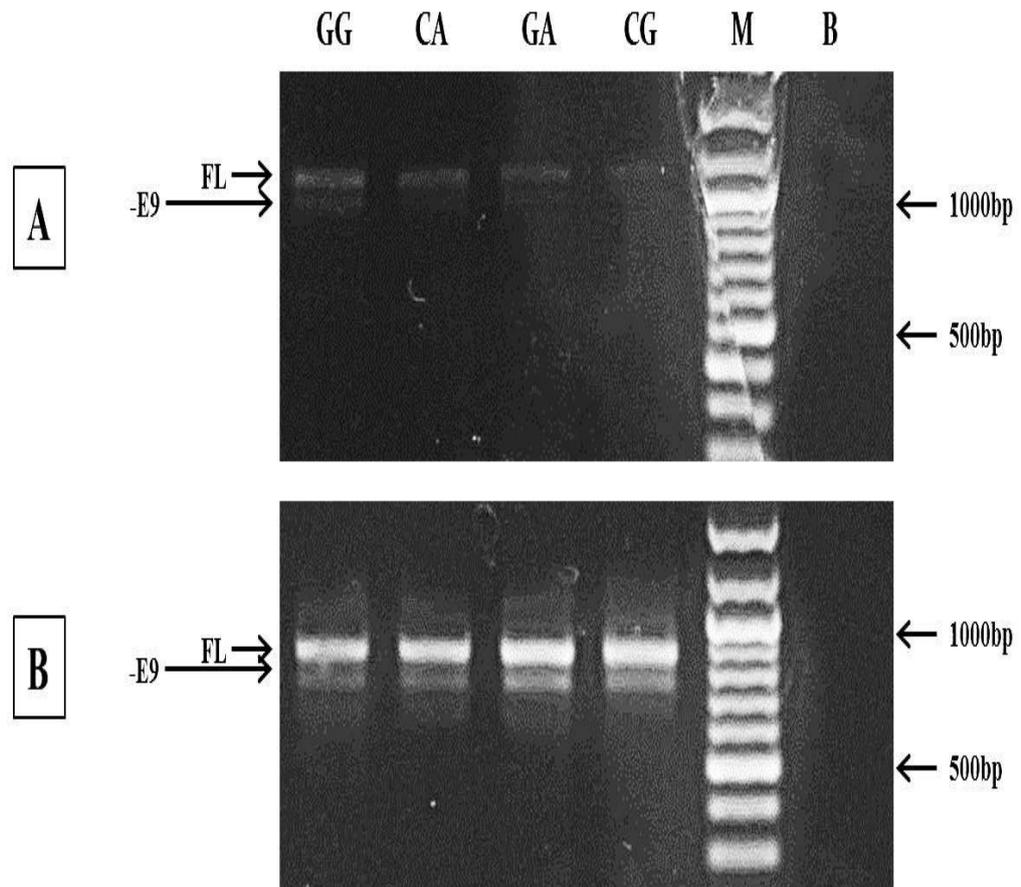
**Figure 3-24: Comparison of the relative expression of CHRNA7 and CHRFAM7A main transcripts in exon 9b-transfected A549 cells.**
*A)* CHRNA7 main transcripts. *B)* CHRFAM7A main transcripts. RT-PCR was run using cDNA from exon 9b-transfected A549 cells as template. Four alleles were compared (G-G (GG), C-A (CA), G-A  (GA) & C-G (CG)) from exon 9b short fragment minigene constructs using specific primers (CHRNA7=exons 1-10; CHRFAM7A=exons A-10). Two major transcripts were detected, full length (FL) and missing exon 9 (-E9) (arrowed). M: 100 bp DNA ladder, B: blank.

### 3.4.3.3 AlphaDigiDoc analysis:

This software analysis is a type of densitometric analysis for semi quantitative measurement of the band intensities on agarose gel electrophoresis. Following each RT-PCR run, the products were subjected to agarose gel electrophoresis and the visualized bands were subjected later to AlphaDigiDoc software analysis. The band intensity was the main parameter used and the exon 9b-excluded transcripts (lower bands on gel electrophoresis) were used as an internal control for each minigene construct. The analysis included calculation of the ratio of transcripts with/without exon 9b inclusion adjusted for the size of the bands used (as the upper band, with exon 9b-included, is of 439bp size while the lower one, without exon 9b, is of 242 bp size) (Figure 3-25).
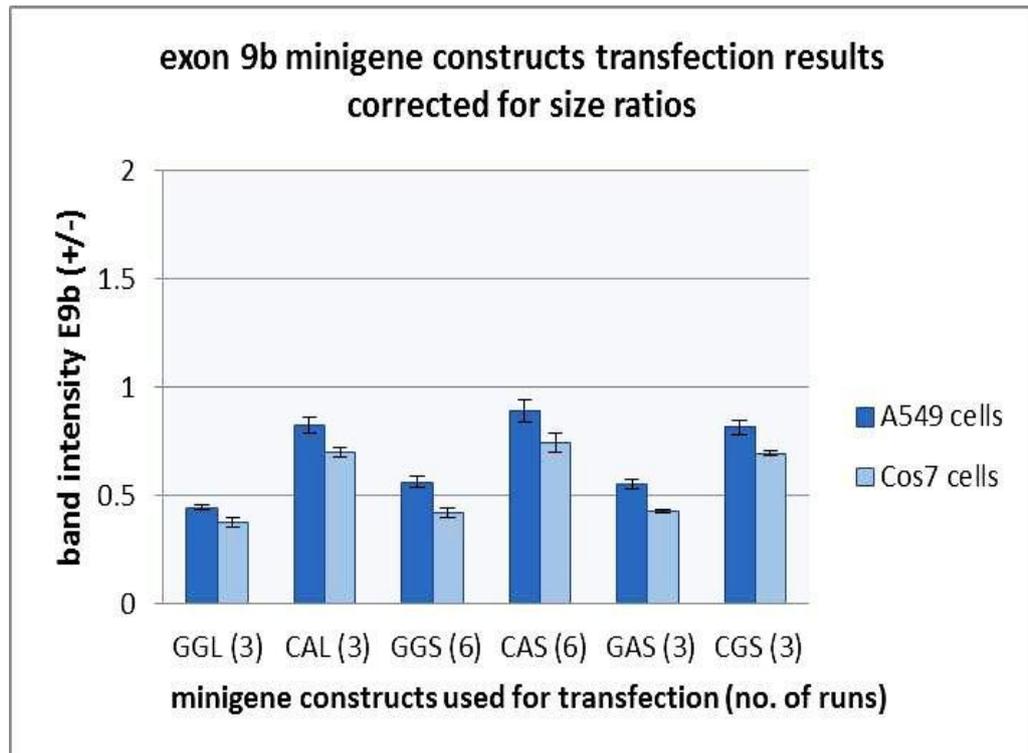
**Figure 3-25: AlphaDigiDoc analysis for the exon 9b inclusion/exclusion transcript ratios of minigene transfection experiments.**
The ratio of band intensity for exon 9b-included/excluded (E9b (+/-)) transcripts resulted from transfection of minigene constructs into A549 and Cos7 cells. The minigene constructs used were G-G (GG), C-A (CA), G-A (GA) & C-G (CG) from exon 9b long (L) or short (S) fragments with number of replicates indicated (between brackets). The ratio corrected for size is shown. For all the data presented, the average of three independent experiments was used and the standard deviation and standard error of the mean were calculated (the latter is shown on the bars).

The results showed that position 77 seem to be the key element in transcript determination and that the presence of C allele at this position enhanced the inclusion of exon 9b compared to the correspondent G allele. This can be shown as follows: for A549 cells, the band intensity ratio was increasing from 0.45 for G-G allele (GGL) to nearly double up to 0.83 for the corresponding C-A allele (CAL). Similarly, the G-G allele (GGS) and G-A allele (GA) showed a 0.6 ratio that increased up to 1.5-1.25 fold (up to 0.9 and 0.8) when the corresponding C-A allele (CAS) and C-G allele (CG) were compared. Similar results were shown for Cos7 cells, with more consistent results showing almost double- fold increase in the band intensity ratio from 0.4 for all G-G alleles tested to 0.7 for all C-A alleles tested (Table 3-15).

**Table 3-15: Summary of AlphaDigiDoc results for minigene constructs of exon 9b.**

| Minigene constructs | Cells tested | |
|---|---|---|
| | **A549** | **Cos7** |
| **GGL** | 0.45 | 0.4 |
| **CAL** | 0.83 | 0.7 |
| **GGS** | 0.6 | 0.42 |
| **CAS** | 0.9 | 0.74 |
| **GA** | 0.6 | 0.43 |
| **CG** | 0.8 | 0.7 |
| The values shown represent the band intensity as tested by using AlphaDigiDoc software analysis adjusted for size of the RT-PCR product (intensity of exon 9b-inserted RT-PCR products/intensity of exon trap exons RT-PCR products. | | |

## 3.5  Discussion:

The work in this chapter focused on testing the effect of the two nucleotide bases (at positions 77 and 154) on CHRNA7 and CHRFAM7A transcripts. Our results suggested that the G77 allele harboured more silencing elements than the C77 allele. This was followed by preparing exon 9b minigene constructs and both exon 9b alleles (G-G from CHRNA7 and C-A allele from CHRFAM7A) were expressed in A549 and Cos7 cells. However, the expression of C-A allele was more than that of the G-G allele. We next aimed at testing single nucleotide effects. The results again were similar with all exon 9b alleles expressed in the same cells, but with C-containing alleles being expressed more than G-containing comparatives at position 77. These results explains why exon 9b transcripts were detected after amplifying CHRFAM7A but not CHRNA7A transcripts in cells tested (A549, BEAS2B, and BE (2)-c).

### 3.5.1  Earlier results on exon 9b transcripts:

In the current study, I tested for the possible effect of a single nucleotide base change on exon inclusion in transcripts of CHRNA7 and its duplicated gene, CHRFAM7A. Earlier results showed that transcripts with inserted sequence between exons 9 and 10 could be detected when amplifying exons 7-10 (Figure 2-27, section 2.7.2). As this part is common between CHRNA7 and CHRFAM7A, it was possible to originate from either gene. However, after using sequence alignment and comparing the detected sequence with intron 9 parts from both genes, the inserted sequence was confirmed to match the reference sequence of intron 9 of CHRFAM7A using NCBI and UCSC

genome browser databases (Figure 2-31, section 2.7.2.1.2). CHRFAM7A is a partial duplication of CHRNA7 that has exons 5-10 almost completely identical (Gault et al., 1998). However, there are still some differences within the introns connecting these exons when aligning such sequences from both genes. The main difference between the two genes within exon 9b sequence was the two bases at positions 77 and 154 of the exon. These bases were both G in CHRNA7 sequence and C & A (respectively) in CHRFAM7A sequence. This fact led to the following question: *Do these two bases affect exon 9b inclusion?*

### 3.5.2    The effects of exonic sequences on exon inclusion:

A proper RNA splicing will include all constitutive exons and remove all introns and pseudo exons (Robberson et al., 1990). However, sometimes splicing might include cassette (alternatively spliced) exons that are not present in all the mRNA transcripts, and that might be included during specific cellular requirements (as a part of a control mechanism) (Wang et al., 2006).

There are a group of factors that might help to direct the spliceosome towards (or away from) a specific exon. These factors are known collectively as the exonic splicing regulatory elements (ESRs) and are grouped into two parts, based on their influencing effect, into enhancers or silencers (Zhang et al., 2008). During RNA splicing, the splicing factors start screening the pre-mRNA sequence for sequence motifs that can bind to enhancers and silencer elements. This is supposed to be followed by the binding of corresponding enhancer and silencer proteins and the net effect are directed towards the more prevalent group (enhancer or silencer) (Sironi et al., 2004). This might explain the reason

behind including constitutive exons rather than pseudo exons in spite of the latter overwhelming the former in more than 10- fold.

From a splicing point of view, exons were supposed to be classified into constitutive, alternatively spliced and pseudo exons (Zhang, 2004). By comparing ESE and ESS within each type of these exons, it was evident that when the ratio of ESE to ESS (PESE/PESS) is higher, the chances of exon inclusion are increased (Table 3-16).

**Table 3-16: Results for PESE/PESS among different exon types.**

| Exon type | PESE/PESS |
|---|---|
| Constitutive | 5.5 |
| Alternatively spliced | 3.6 |
| Pseudo exon | 0.64 |
| Summarised from (Zhang, 2004) | |

In addition to enhancer and silencer elements, other factors play an important role in defining exons such as splice sites, intron sizes and SR proteins (Dewey et al., 2006) and the overall interaction of these factors with the dominant effector (enhancer or silencer) interact together.

There is a growing body of evidence that enhancers and silencers compete for settling in or getting rid of the spliceosome, respectively. This seems clearer on changing a single nucleotide base within ESE or ESS sequences, an effect that can change the entire function of such a sequence in the opposite direction (Sironi et al., 2004, Wang, 2004, Zhang, 2004). These minor changes can have a huge impact on changing exon status in either way from inclusion to exclusion or vice versa (Kashima and Manley, 2003, Buratti et al., 2004,

Cooper et al., 2009, Vezain et al., 2010). It was supposed that enhancer elements act by recruiting spliceosome and antagonising effects of silencer elements, influencing exon inclusion (for constitutive exons), whereas the silencers predominate the overall effects when pseudo exons are enrolled for spliceosome scanning (Fairbrother and Chasin, 2000).

### 3.5.3   Using bio-informatics tools to test for the effects of exon 9b bases:

*Exon Scan* software was used which could not predict exon 9b as possible transcripts of both genes. This result confirms the other results tested this sequence and could not detect a constitutive exon sequence using PESE/PESS ratio (Table 3-7). The next step included checking for a possible SNP site within the two base sites. The results came back with scarce information (with no ancestral origin and the fact that the results included test in few individuals) rendering the information less important and could not name these two positions as SNPs (Table 3-10).

Other differences within intron 9 (upstream of exon 9b) were detected using NCBI and UCSC genome browser databases (Table 3-12).

The next step was using *In-Silico* analysis using multiple software analysis tools for exploring enhancer and silencer motifs surrounding exon 9b positions 77 and 154. The two main sequences compared were exon 9b short fragments of G-G and C-A alleles. First, the *Human Splicing Finder* tool was used, which comprises more than eight based-in software analysis tools that apply multiple functions. The first parameters compared between the two alleles were the splice sites and branch points and these were exactly matching between both alleles. The next part included analysis for enhancer and silencer elements within exon 9b and their flanking 100 bp intronic sequences. This analysis

showed that the G77 allele was more enriched with ESE sequences than C77 allele based on analysis from *RESCUE-ESE, ESE-finder* and other *Human splicing finder*-based experimental tools (Table 3-8). Similar results were obtained when using RESCUE-ESE, ESE-finder software program independently via their specific website. In addition, splicing rainbow software analysis tool was used and it showed similar results for both alleles in terms of enhancers and silencers around the two bases compared. However, Human splicing finder is wealthier with analysis tools, and the silencer testing showed that more silencer elements were clustered around G77 site (Table 3-8 and Figure 3-12). Taking into account all these results, it seems that G77 site is attracting more silencer elements than C77 site within exon 9b, although G77 seem to have more potential binding preferences to enhancer elements than C77. The *In-Silico* analysis tools were confirmed of high accuracy in predicting ESE sequence motifs but were also suggested not to include all potential ESE sites. The same can be true for silencer elements, increasing the challenge for predicting exactly enhancer and silencer elements for a given sequence prior to experimental testing for the effect of a specific sequence on splicing (Goren et al., 2006). It is worth mentioning that the ratio of predicted ESE/ESS (PESE/PESS) was higher for G-G allele than for C-A allele in the short exon 9b fragments (Table 3-7). However, this ratio was almost equal when long exon 9b fragment sequences of the two alleles were compared, indicating that more silencer elements for G-G allele or more enhancer elements for C-A allele were included. Our results showed that there was no single difference within the intron 9 part included in exon 9b long fragment apart from bases 77 and 154 of exon 9b (Table 3-14). Similarly, RESCUE-ESE detected two

different ESE at G77 compared to one at C77 (Figure 3-8). The fact that a single nucleotide mutation can change ESE sequence into an ESS or even create a new silencer element should be borne in mind (Fairbrother et al., 2002, Zhang et al., 2005, Kashima and Manley, 2003). Another important observation is that A154 seemed to be a potential binding site for SR-9G8 when Splicing Rainbow software program was used) (Figure 3-11). This protein was detected as a silencer in one study on human tissue, but in other studies it was shown to be enhancing transport to the cytoplasm for some viral mRNA infecting human body (Hurst et al., 2013, Boyd, 1997, Albuquerque et al., 2009). Based on the results found in this study, A154 had a minor effect on including exon 9b, but it might have a secondary role to C77 in helping such an mRNA transcript to be transported to the cytoplasm for subsequent translation.

### 3.5.4   The preparation and testing of exon 9b minigene constructs:

After completing the bioinformatics analysis, the aim was to prepare the two alleles of exon 9b. For this purpose, genomic DNA samples from control subjects were screened for exon 9b using a specific primer set that anneals to exon 9b flanking introns (around 100 bp each side) to generate short exon 9b fragment. Furthermore, an additional forward primer was designed upstream of exon 9b farther than the first forward designed primer by about 850 bp to generate exon 9b long fragments (Table 3-2).

Only two control subjects were shown to have the two alleles of exon 9b, in addition to A549 cells, while the rest of the control subjects were shown to express a single heterozygote exon 9b at position 154 and only BE (2)-c cells expressed double homozygote exon 9b (Table 3-13). It is worth noting that BE

(2)-c cells showed higher expression of CHRNA7 than CHRFAM7A transcripts (Figure 2-38, section 2.7.3). The detection of C-A rather suggests possible polymorphism at the two base positions of exon 9b (77 & 154).

These results led to the following question: *Are either of these alleles at positions 77 and 154 the result of a SNP? Therefore, which allele is the original sequence and which one is possibly polymorphic. Moreover, if there was a polymorphism, did that SNP change took place before or after CHRNA7 duplication?* This might be explained in part by the detection of C only at position 77 in 8 out of 10 control subjects. In addition, BE (2)-c cells expressed transcripts from both genes (CHRNA7 and CHRFAM7A) but were expressing double homozygote allele at positions 77 and 154 of exon 9b, adding a further evidence that the ancestral allele is probably C-A.

Earlier results in this study showed that the TG at positions 67-68 of exon 6 of CHRFAM7A were missing from three cells tested (A549 and PBMC) (as mentioned in section 2.8.2). This suggested that the orientation of this gene in relation to that of CHRNA7 is in a tail-to-head arrangement (Araud et al., 2011, Flomen et al., 2008). In addition, each of the cells tested could have no, one or both CHRFAM7A copies (Flomen et al., 2006) which might explain the different peaks for G-G and C-A alleles at positions 77 and 154 of exon 9b (Figure 3-17). This figure summarised the results of sequencing of exon 9b short fragment amplified in A549 and BE (2)-c cells in addition to 10 control subjects. The C-A peaks for A549 were higher than those of G-G peaks (in a ratio of 2:1). This could suggest that there might be only a single copy of CHRNA7 in these cells (that is already rarely occurring). The other possible explanation is that the C-A was the ancestral allele and that a G-G double

mutation within CHRNA7 sequence took place in only one copy of the gene, explaining the peak difference. For BE (2)-c cells, the opposite is true and there was no mutation at these positions reflecting only one dominant C-A allele (earlier results showed that these cells express both CHRFAM7A and CHRNA7, but the latter was expressed much more than the former, a finding that was opposite to all other tested cells). Regarding the control subjects, they were clustered into two main groups:  the double heterozygote exon 9b, with the two alleles (G-G and C-A) expressed equally, and the single heterozygote exon 9b, with a single peak for C allele at position 77, and a double peak at position 154 (higher peak for A than for G in a ratio of 2:1). The double heterozygote control subjects showed that they contain almost equal copies of the two genes (this was used as a template for subsequent preparation of minigene constructs). The single heterozygote group suggested a single mutation at position 154 of CHRNA7 took place and involved only one of the two gene copies (explained by the small peak height).

The other differences within intron 9, upstream of exon 9b, that were included in long exon 9b fragment (Table 3-12) couldn't be detected from amplified RT-PCR products of A549, BE (2)-c cells and two control subjects (representing single and double heterozygotes at positions 77 and 154 of exon 9b) (Table 3-13 & Figure 3-16). This meant the different bases within intron 9 of both genes could represent an error in the database. Apart from the two bases at positions 77 and 154, the rest on intron 9 part included in exon 9b constructs was completely matching between the two alleles. This meant that intronic sequences embedded within the databases were less reliable than the coding sequences, and possible errors might exist within such sequences.

The next step included preparation of the minigene constructs. The double heterozygote samples were selected as the target for construct preparation (as the two alleles of exon 9b were detected in almost equal proportion). First, exon 9b short fragment constructs were prepared, and then exon 9b long fragments were prepared. The fragments were then inserted into exon trap vector for subsequent transfection assays. The prepared exon 9b alleles were inserted into the multiple cloning site of exon trap flanked by intronic sequences followed by an exon on both sides. The only difference between the short and long exon 9b fragment constructs was about 850 additional bases of intronic sequences upstream of exon 9b. However, the transfection results showed that all the constructs expressed exon 9b transcripts in both cell lines used (A549 and Cos7 cells), although in different ratios (Figure 3-22). The C-A allele constructs expressed more exon 9b than the G-G allele highlighting a possible effect attributed to the base constitution at these two positions (77 and 154) of exon 9b.

*Is it position 77 or 154 that has the major influence on exon 9b inclusion?* To answer this question, site directed mutagenesis was used to generate constructs with additional possible exon 9b alleles: G-A and C-G alleles. As there was no noticeable difference when comparing exon 9b short and long fragments, the former was used as a template for the mutagenesis preparation, targeting the nucleotide at position 154. Thus a new set of ex vivo splicing experiments was carried out (using the same cell lines used above) and showed similar results, with C-A and C-G alleles including more exon 9b than G-G and G-A alleles (Figure 3-23). These results suggested a major role for the C-allele at position

77 of exon 9b to enhance its inclusion in the mRNA transcript. The DNA sequence for the constructs used was identical except at positions 77 and 154.

Next, we aimed to test for possible effects of transfection of cells with exon 9b minigene constructs on the transcription process within the transfected cells using RT-PCR. The results matched those detected earlier on A549 cells, with the two main transcripts and the two genes expressed at similar ratios (Figure 3-24). To quantify exon 9b expression between the different constructs used, the UV picture of the gel electrophoresis of the RT-PCR products was subjected to densitometric analysis. For this purpose, AlphaDigiDoc1201 software was used, taking into consideration the density of upper band brightness (that included exon 9b) as the comparator. In addition, the density of the lower band brightness (with no exon 9b) was used as an internal control, and accordingly the ratio of exon 9b-included/excluded transcripts was measured and compared across all constructs used. It is worth mentioning that exon 9b-included transcripts measured larger than exon 9b-excluded transcripts (439 bp sizes compared to 242 bp sizes, respectively). This raised the need for using a correction for the size of the two bands accordingly to omit the size difference factor that might influence more ethidium bromide chelation with larger size products. However, the results were almost matching with and without correcting for size difference, with all the results showing again a major effect of C-allele at position 77 (Figure 3-25).

Taking together the ex vivo splicing assays and the *In-Silico* analysis data, it is clear that the potential silencer elements that could bind to G77 can be more powerful than the enhancer elements competing for the same site resulting in lower overall inclusion ratio of exon 9b G-G allele. Considering the same

concept, it is possible to conclude that the absence of potential silencer elements binding to C77 is less interfering with C-A allele on exon 9b inclusion, increasing its inclusion rate.

The effect of the inclusion of exon 9b on secondary protein structure was tested on secondary protein structure using PSIPRED software analysis (as shown in section 2.7.6.2.4). The results showed that inclusion of exon 9b within both predicted proteins would have the same effect as missing exon 9 (Figure 2-59, *C & D*), with the protein having three TMDs and an intracellular, rather than an extracellular, carboxy terminal end. Surprisingly, the translation was shown to be ended by a premature stop codon just before position 77 of exon 9b indicating that exon 9b would not be translated completely into the subsequent protein. This result shed the light on the importance of such a transcript in the control mechanism that CHRFAM7A has under normal conditions to suppress expression of CHRNA7 and its protein.

## 3.6   Conclusions

The results shown in this chapter confirm that the nucleotide base at position 77 of exon 9b has a dramatic effect on exon inclusion within the mRNA transcripts. In the case of the C77 allele, the chances of including the exon 9b sequence were increased compared to the G77 allele. These results suggest that a C-base at this position favours exon 9b inclusion. Taking into consideration that most bio-informatic analysis tools of exon 9b sequences showed that the G77 can act as a greater potential enhancer sequence than C77, it is worth noting that C77 escaped binding to many silencer sequences. Taking all of these results together including the minigene construct transfection results, it seems likely that G77 acts as a competition site for enhancer and silencer factors which ultimately drive exon 9b G-G allele into a low expression ratio that could be undetectable (as evident from earlier results of CHRNA7 transcripts). In contrast, it seems likely that C77 acts as a more potent site than G77 in stabilising the spliceosome on exon 9b (no silencer potential binding sites were detected). This indicates that position 77 of exon 9b is crucial for inclusion within the mRNA transcripts depending on the nucleotide base included. However, there is still a room for suggesting a possible role of this transcript in the control mechanism theory of CHRFAM7A over CHRNA7 during the normal (non-inflammatory) condition. This means that increased expression of exon 9b transcripts from CHRFAM7A can interfere with $\alpha$7 mRNA (from CHRNA7) affecting its expression.

# 4   General discussion

## 4.1   CHRNA7 and CHRFAM7A transcripts:

There is a range of human diseases that are called "chronic inflammatory diseases". Among these is Chronic Obstructive Pulmonary Disease (COPD). These diseases are characterised by a chronic inflammation that involve one (or more) tissues. On the other hand, the immune response in patients with these diseases sometimes seems to be defective and not counteracting the inflammatory changes. A normal response by the immune system to an injury or infection to human body tissues would be in the form of start of inflammation (Tracey, 2002). This reaction would involve activation of different cells that participate in one way or another to produce inflammatory mediators. Such mediators can help to repair the damaged tissues and suppress a threatening pathogen. However, this inflammatory response is regulated by the Central Nervous System (CNS) to be limited in time and magnitude to prevent further tissue damage, spread of inflammation to other body systems, and a possible change to chronic inflammatory status. The immune system is the active arm of the CNS and can limit the effect of inflammation via many pathways. Among these is the "cholinergic anti-inflammatory pathway" which is mediated by the vagus nerve activation (Borovikova et al., 2000). The vagus nerve is well known for innervating a wide range of body organs and thus its role in limiting inflammation is vital to prevent progression of inflammation to sepsis (Wang et al., 2003, Wang et al., 2004). Although the vagus nerve is the key arm of the cholinergic anti-inflammatory pathway, its stimulatory signal requires another vital component to be delivered to the target organs. This

component is "the alpha7-nicotinic acetylcholine receptor (α7-nAChR). It was shown that the absence of this receptor in an experimental model of inflammation could render the vagus nerve ineffective against inflammation and subsequent sepsis (Borovikova et al., 2000). The vagus nerve stimulates the α7-nAChR via acetylcholine (ACh), but nicotine can produce a similar effect to ACh and thus suppress inflammation. It was shown that chronic exposure to nicotine could lead to suppression of α7-nAChRs, which means that the immune response is partially impaired. Such an effect of nicotine was suggested to be the basis for a defective anti-inflammatory response in chronic smokers and in COPD patients (Kawashima et al., 2012).

### 4.1.1   Do α7-nAChRs have a role in preventing lung inflammation?

This was answered by using a human inflammatory lung model where α7-nAChRs were shown to mediate the anti-inflammatory pathway and suppress inflammation (Kox et al., 2011). This means that factors that can affect α7-nAChRs expression or function might play a role in the development or progression of COPD. For this reason, it was necessary to try to understand the mechanisms that control α7-nAChRs. This receptor is encoded by the CHRNA7 gene on chromosome 15, with its partial duplication (CHRFAM7A) located 1.6 Mb upstream of it (Gault et al., 1998). The complex regulatory mechanism between the two genes was uncovered recently and it shows a correlation with inflammation. It was shown that during the resting condition (no tissue injury or inflammation), CHRFAM7A is down regulating CHRNA7 and thus limiting the receptor role in the anti-inflammatory pathway. However, once inflammation starts (in response to injury or infection) the released inflammatory mediators start down regulating CHRFAM7A and thus allowing

for an indirect CHRNA7 up regulation (Araud et al., 2011, Benfante et al., 2011, de Lucas-Cerrillo AM, 2011). This effect will end up with expression of α7-nAChRs and a subsequent suppression of inflammation. This suppression would help to limit the effect of inflammation to the tissue involved and prevent a possible systemic spread and subsequent sepsis (Wang et al., 2003). This effect of CHRNA7 helps in its own regulation, and by suppressing the inflammatory mediators that blocked an earlier down regulation of CHRFAM7A (on CHRNA7) this would help the tissue to go back to the resting condition.

There are a number of factors that could affect the regulatory role of CHRFAM7A on CHRNA7 which may interfere with the anti-inflammatory regulation of involved tissues. These include the copy number of both genes and the factors that affect the assembly and function of the α7-nAChRs. It was shown that CHRFAM7A is a recent evolutionary event that is unique to the human species and could not be detected in animals (Locke et al., 2003). However, the gene could not always be identified, and some individuals lack one or both copies of the gene (Riley, 2002, Araud et al., 2011). It was proved in many studies that the locus on chromosome 15 which harbours the CHRNA7 and CHRFAM7A (15q13-14) is subjected to a range of deletion and duplication events that affected the restructuring of this locus. Deletion and duplication events that affected the CHRNA7 helped creating the CHRFAM7A (Makoff and Flomen, 2007). Similarly, a whole deletion event might affect one both copies of CHRNA7 leading to different neuropsychiatric disorders (Shinawi et al., 2009, Hoppman-Chaney et al., 2013). Collectively the resulting deletion of either gene could affect the anti-inflammatory pathway.

The other factor that can affect the α7-nAChRs function significantly is a defective assembly. This receptor is a homopentamer that form by assembly of five subunits of α7-nAChRs around a central channel pore and changes to this architecture can affect the channel function and subsequent effect on inflammation (Galzi et al., 1992). The alternative RNA splicing could also affect the assembly of the α7-nAChRs possibly by producing different protein isoforms of the receptor that could co-assemble with the wild type isoform. This can give rise to a somewhat heteropentameric structure of the receptor that was proved to affect the receptor and channel functions. Examples of such alternative transcripts were missing exon 8 (García-Guzmán et al., 1995), and inserted exon 9b (Saragoza et al., 2003). However, the co-expression of dupα7 protein (translated from CHRFAM7A transcripts) gave similar results to co-expression of alternative transcripts and all lead to a different co-assembly of the α7-nAChRs. This effect changed the receptor into a non-functioning status. This is not necessarily always true, and the co-expression of transcripts with inserted exon 4a did not have an inhibitory effect on the receptor function (Severance and Yolken, 2008). Exon 4a transcripts were suggested to have no effect on the function of the receptor when co-assembled with the wild type subunits.

Our results showed that CHRNA7 and CHRFAM7A both express a novel transcript that is missing exon 9. At the start of the work, we suggested that this effect might be due to tissue specific splicing factors that could be inactive and thus causing skipping of exon 9. As both genes have almost 100 % similarity in the sequences of exons 5-10 (Gault et al., 1998), it might be accepted to suggest that exon 9 splicing from both genes is under control of the same

splicing factors and SR proteins. If this is true, it might explain the reason for detecting missing exon 9 transcripts from both genes in the same cells. However, a possible effect due to tissue specific splicing should be ruled out as the same transcript missing exon 9 was detected from both genes in different human cell lines (A549, BEAS2B, BE (2)-c) and primary human cells (PBMC).

It is important to remember that CHRFAM7A has a dominant negative effect on CHRNA7 during the resting condition (non-inflammatory). This effect is kept until the start of inflammation, which would reverse the roles of both genes ending up with expression of α7-nAChRs to start localising inflammation (Benfante et al., 2011, de Lucas-Cerrillo AM, 2011). Bearing in mind this fact, it might be possible to suggest the production of alternative transcripts (such as missing exon 9 transcript) as part of the regulatory mechanism of CHRFAM7A on CHRNA7. On the other hand, it might be also possible to suggest that CHRFAM7A is keeping CHRNA7 down regulated via enhancing production of missing exon 9 transcripts. These transcripts have no frame shift, and if translated (predicted using software programs, like PSIPRED) can form a protein similar to the wild type, except  that they are missing TMD 3 and have an intracellular carboxy terminus (rather than extracellular one in the wild type). If such transcripts were translated, they are expected to co-assemble with the wild type α7 protein, as was shown in other studies (discussed above). This could lead to the formation of non-functioning receptors and ion channels.

In the present study, another novel transcript was detected from CHRFAM7A, which is inserted exon 9b. This transcript was detected from CHRFAM7A

transcripts only and not from those of CHRNA7. By referring to the reference sequence of both genes, we showed that exon 9b is actually part of the intron 9 sequence. Interestingly, there were only two bases within exon 9b sequence (in both genes) different between both genes at positions 77 and 154 (from the 5` end of the exon). These nucleotide bases were both G in CHRNA7 sequence while they were C and A at positions 77 and 154, respectively for CHRFAM7A sequence. This meant that the exon 9b allele was G-G in CHRNA7 and C-A in CHRFAM7A at these two positions. Considering these two bases, these nucleotide bases might affect the spliceosome assembly on exon 9b in both genes differently leading to exon 9b expression from CHRFAM7A gene only. To test this effect, bio-informatics analysis showed that G77 (in CHRNA7) is harbouring more silencer sequence motifs than C77 (in CHRFAM7A). Following these results, exon 9b minigene constructs were prepared to test for effects of the two bases on transcription. Four different exon 9b alleles were prepared (G-G, C-A, G-A, C-G) to cover the two bases within exon 9b from both genes for a possible splicing enhancer or suppressor effect. These constructs were transfected into A549 and Cos7 cells. Our results showed that the C-A exon 9b allele was expressed more than G-G allele in both cell lines. Further results revealed that the C77 have the superior effect on all other nucleotide bases in enhancing exon 9b inclusion in splicing. These results showed that a single nucleotide base might affect splicing and change inclusion or exclusion of exon 9b.

It is tempting to suggest that exon 9b transcripts from CHRFAM7A could also play a role in the control of CHRNA7 transcripts via interfering with α7

mRNA. However, all the suggested effects of missing exon 9 or inserted exon 9b are still a matter of speculation.

The α7-nAChR has a central role in the anti-inflammatory pathway that participated in making this receptor the main target for many studies aiming at further understanding or treating inflammation. Among these studies is targeting the receptor by using nicotine treatment in pancreatitis to decrease the incidence of individuals developing type 1 diabetes mellitus (Mabley et al., 2002). Using the same principle, nicotine treatment targeting α7-nAChR was shown to decrease the incidence of endotoxemia and shock preventing sepsis and death (Wang et al., 2004). Similar approach was used by stimulating the vagus nerve. This effect is mediated by the production of acetylcholine (ACh) that would act as agonist for these receptors and starting anti-inflammatory effect (Borovikova et al., 2000, Wang et al., 2003). The same principle was used for treatment of chronic inflammatory diseases (Pavlov et al., 2009). Recent and ongoing treatment trials tried successfully the use of α7-agonists in the treatment of inflammatory diseases, such as rheumatoid arthritis (van Maanen et al., 2009, Olofsson et al., 2012). Further treatment trials involved targeting mild forms of obesity with α7-agonist treatment (Cancello et al., 2012).

Interestingly, α7-nAChR was the focus of interest for understanding neuropsychiatric disorders. This might be best presented by the finding of decreased CHRNA7 expression in schizophrenia patients versus control subjects (Gault et al., 2003) and α7-nAChR was linked with schizophrenia in subsequent studies (Araud et al., 2011). One of the theories explaining this link was based on the defective neuronal transmission (mediated by the α7-nAChR)

that could affect cognition (Weiland et al., 2000). This was supported by the fact that vagus nerve stimulation was approved for the treatment of refractory cases of schizophrenia and depression (Ulloa, 2005).

Finally, exploring the role of α7-nAChRs in COPD might shed some light on the mechanisms that control the development and progression of the disease. This is based on the anti-inflammatory role that this receptor could play in a chronic inflammatory disease (as COPD). It is worth mentioning that in a rodent model for acute lung injury, α7-nAChRs was detected on the surface of alveolar macrophages and neutrophils, and was found to mediate anti-inflammatory response to the lung injury (Su et al., 2007).

The findings from my study showed that there is altered mRNA expression of CHRNA7 and CHRFAM7A in airway cells, which could affect α7-nAChR expression and/or function in these cells. This effect could lead to a dysfunction of the cholinergic anti-inflammatory pathway in airway cells and a possible exaggerated inflammatory response in COPD. Whether this is true or not, these results warrant the need for further studies to explore this phenomenon and explain its role in the inflammatory process in COPD.

## 4.2   Limitations of work:

This work helped in the detection of many transcripts for both CHRNA7 and CHRFAM7A genes, some of which were novel transcripts. However, there are still many limitations facing the exploration of transcription for these two genes, which is mainly the duplicated part. The fact that exons 5-10 have the same sequence in both genes has a major role in interfering with many methods aimed at testing each gene transcripts. For instance, it was non-applicable to quantify the full-length and missing exon 9 transcripts of CHRNA7 using QRT-PCR (as mentioned in section 2.7.4).

Furthermore, another important obstacle when exploring the intronic part of these genes was the errors in the sequences deposited in the main databases. These errors can lead to unexpected results if the planned work was greatly based on such sequences.

Another limitation is that I could not test for the functional effect of inserted exon 9b or missing exon 9.

## 4.3 Future work:

From our results, we can conclude the following targets for a future work on these two genes:

1) To test for the cigarette smoke extract (CSE) work on airway epithelial cells to detect the effect of smoking on inflammatory mediators in the presence or absence of (knocked down) α7 receptors.

2) Generate plasmid constructs from the full length mRNA of both genes (CHRNA7 & CHRFAM7A). Such constructs could be used in transfection studies that involve inflammatory models lacking α7 receptors (such as Cos7 cells). Additional plasmid constructs could be generated from the full length ones that are matching missing exon 9 and with inserted exon 9b to test for their effects on anti-inflammatory response.

3) To test a group of COPD patients and control subjects for CHRNA7 and CHRFAM7A main transcripts to correlate the expression of transcripts missing exon 9 or with inserted exon 9b from both genes. The transcript expression could be correlated with inflammatory mediator levels and with the disease severity.

# 5 Appendix

Part of our work was not completed due to the limitation of time. This part included the start of functional analysis for the effects of CHRNA7 and CHRFAM7A in inflammatory models of airway epithelial cells. For this purpose, A549 cells were used as inflammatory model. The cells were grown on plates as mentioned in section 2.6.1. The initial optimization steps included testing for the effect of apoptosis on the cells. Different concentrations of cigarette smoke extract (CSE) were prepared for incubation with A549 cells. Similar work was applied by using CSE on alveolar macrophages (Kojima et al., 2013). This was followed by testing first for semi-quantitative RT-PCR (as mentioned in section 2.6.10). The initial results showed similar results for expression of full length and missing exon 9 transcripts from A549 after stimulation with 1, 5 and 10 % CSE for 0, 8 and 24 hr. Due to limitation of time, these results could not be replicated and further inflammatory mediators measurements were suspended.

## 5.1 Methods and results:

### 5.1.1 Cigarette smoke extract (CSE) preparation:

Using CSE-preparation apparatus, CSE was prepared (Figure 5-1).
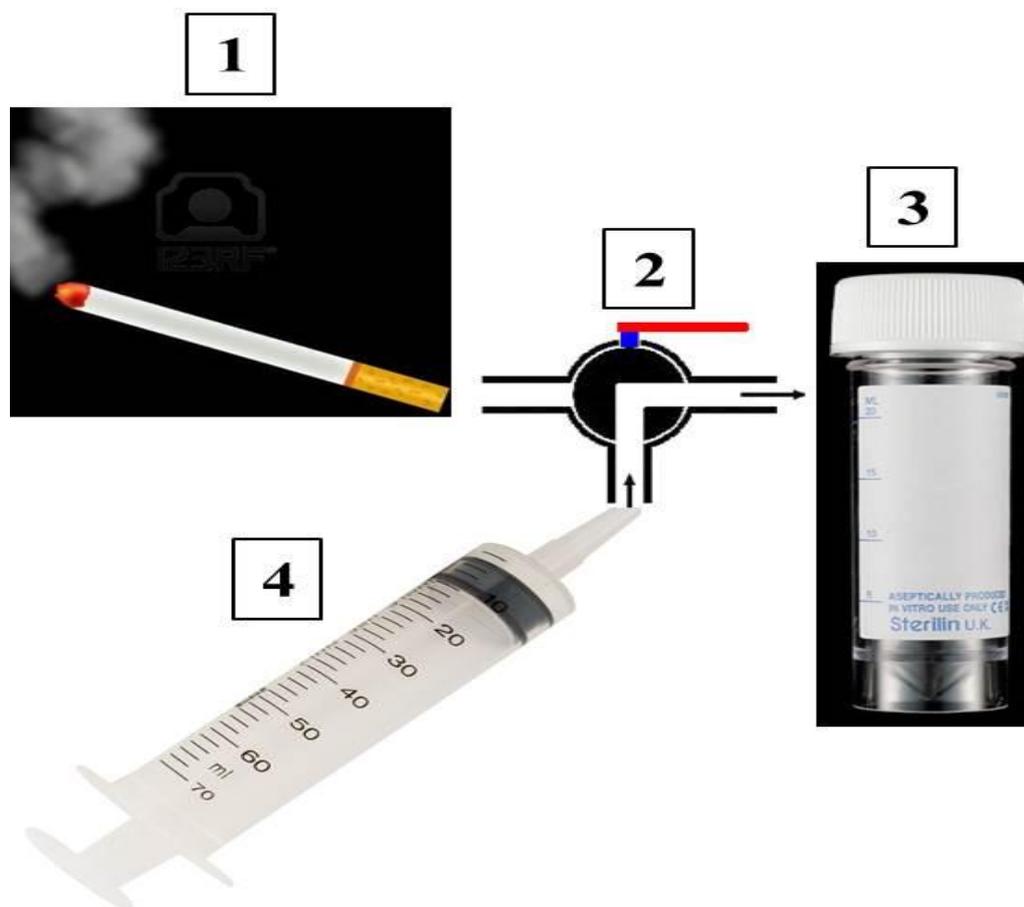
**Figure 5-1: CSE-apparatus.**
For preparing CSE, one cigarette, 10 ml of phenol-red free and 50 ml syringe were attached together using rubber tubes and three-way valve. The valve helped control aspirating smoke from the cigarette at a step and then pushing the smoke to the medium in the next step. These steps were repeated until the whole cigarette was completed.

The method included the lighting up of a single cigarette at a time. The smoke of the cigarette is then aspirated and introduced into a medium to make it dissolve within. This preparation step was applied using a three-way valve tube system that connects between the cigarette, medium used and a 50-ml syringe. The syringe helps to aspirate the smoke from the cigarette and then push it to the medium for subsequent dissolution. The medium used was phenol red-free DMEM (similar to that used with A549, mentioned in section 2.6.1) to avoid interfering with Nanodrop spectrometric measurements. Following the

preparation, different dilutions of the prepared conditioned smoke-medium (CSM) were prepared and optical densities were detected using Nanodrop spectrometer (using protein A-280 property, at 320 nm wave length). From the dilution readings (that represents their CSE saturation status), an optical density of 0.15 was considered as equal to 100% CSE concentration (Figure 5-2).

This calculation was used to prepare 100% CSE, which can be diluted later to the required CSE concentrations (1, 5 & 10%).

## CSE % prepared

$y = 2.1569x - 0.0179$
$R^2 = 0.9999$
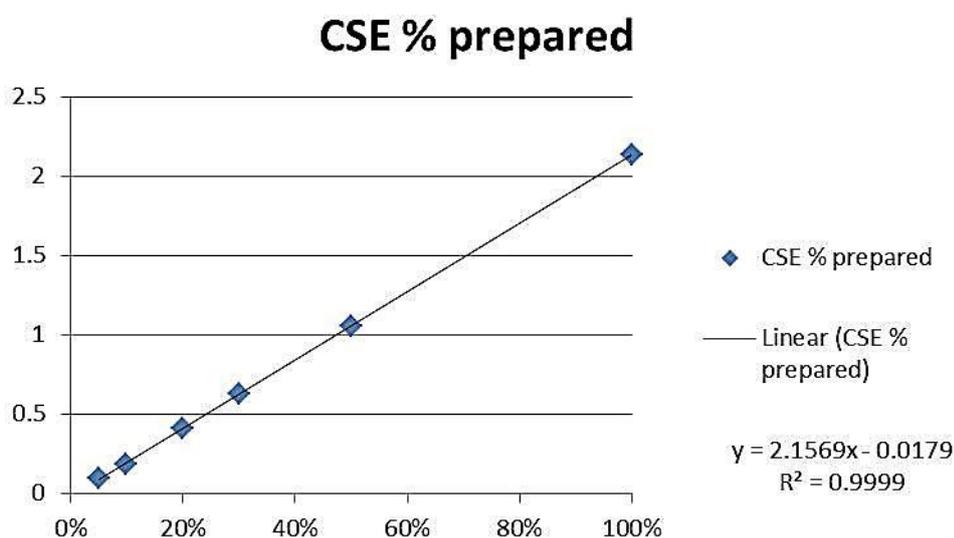
**Figure 5-2: CSE dilutions Nanodrop spectrometric readings.**
Several dilutions from the prepared CSM were prepared and optical densities were measured using Nanodrop spectrometry (using protein-A280 property at wavelength 320 nm). The dilutions tested were 5, 10, 20, 30, 50, and 100% of the prepared CSM. Note the increase in intensities with increased CSM concentrations.

### 5.1.2  Cell counting:

The A549 cells were prepared and grown as mentioned in section 2.6.1. To ensure that relatively equal number of cells was used in each stimulatory condition, cell counting proceeded adding the prepared CSE. This was applied

by adding $1*10^6$ cells/flask used. The aim of cell counting was, in addition to ensuring the comparison of equal cell numbers within different CSE % used, is to test for possible apoptosis effect with increased time. This was tested using trypan blue dye to stain (that can permeate dead but not live cells) following 0, 8 and 24 hr of incubation at conditions used for A549 cells (as mentioned in section 2.6.1). The zero time was regarded as a control status of the cells at the start of seeding and the next two time points (8 & 24 hr) were compared with the former one. The final calculations used were total cell, live cell an dead cell numbers and percentages (Figure 5-3).
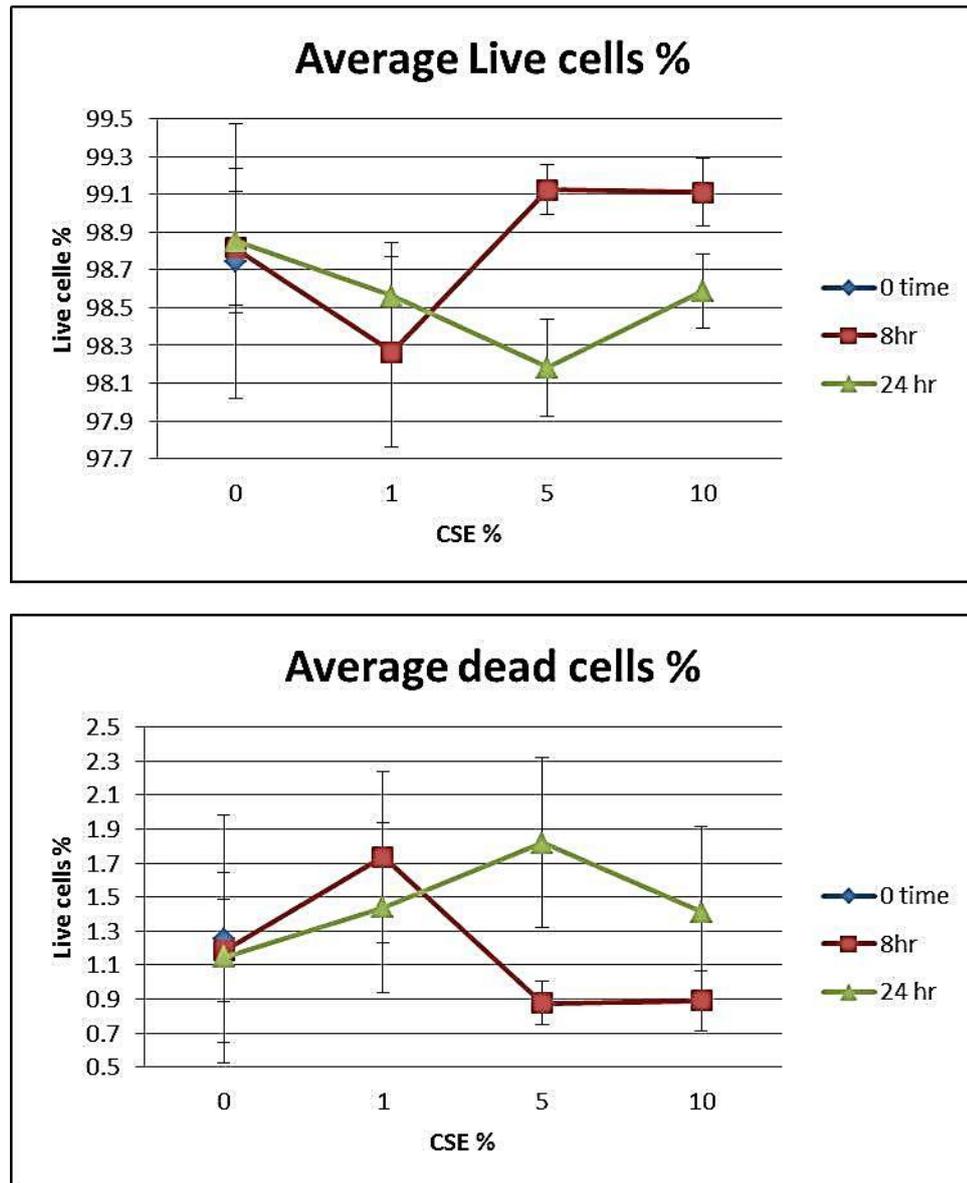
**Figure 5-3: The average percentages of live and dead A549 cells.**
*Top:* average percentage of live cells grown over different time points and using different CSE %. *Bottom:* corresponding percentages for dead cells. Note that the cell numbers were almost equal at 0% CSE for all time points. However, for 8h incubation with CSE preps, cells shifted towards proliferation and fewer dead cells could be detected. When incubation time was increased to 24 hr, fewer live cells and more dead cells were detected at 5% & 10% CSE preps used (compared to corresponding 8h results). The error bars represent SEM for two sets of experiments.

The next step included extracting RNA from the A549 cells in different time points and CSE conditions (as mentioned in section 2.6.2) and generating cDNA (as mentioned in section 2.6.3) for testing with RT-PCR (as mentioned in section 2.6.10). The results were similar to those shown with non-stimulated A549 cells (as shown in section 2.7.5). However, these results are based on single run, and thus no complete interpretation could be based by using these results.

However, further work on measurement of inflammatory mediators was not completed due to limitation of time.

# REFERENCES

AL-MASMOUM, H. A. 2011. Cloning and Expression of CHRNA7 isoforms.

ALBUQUERQUE, E. X., PEREIRA, E. F., ALKONDON, M. & ROGERS, S. W. 2009. Mammalian nicotinic acetylcholine receptors: from structure to function. *Physiol Rev,* 89**,** 73-120.

ALEXANDER, J. K., GOVIND, A. P., DRISDEL, R. C., BLANTON, M. P., VALLEJO, Y., LAM, T. T. & GREEN, W. N. 2010. Palmitoylation of nicotinic acetylcholine receptors. *J Mol Neurosci,* 40**,** 12-20.

AMOS, C. I., WU, X., BRODERICK, P., GORLOV, I. P., GU, J., EISEN, T., DONG, Q., ZHANG, Q., GU, X., VIJAYAKRISHNAN, J., SULLIVAN, K., MATAKIDOU, A., WANG, Y., MILLS, G., DOHENY, K., TSAI, Y. Y., CHEN, W. V., SHETE, S., SPITZ, M. R. & HOULSTON, R. S. 2008. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet,* 40**,** 616-22.

ARAUD, T., GRAW, S., BERGER, R., LEE, M., NEVEU, E., BERTRAND, D. & LEONARD, S. 2011. The chimeric gene CHRFAM7A, a partial duplication of the CHRNA7 gene, is a dominant negative regulator of alpha7*nAChR function. *Biochem Pharmacol,* 82**,** 904-14.

AVRAMOPOULOU, V., MAMALAKI, A. & TZARTOS, S. J. 2004. Soluble, oligomeric, and ligand-binding extracellular domain of the human alpha7 acetylcholine receptor expressed in yeast: replacement of the hydrophobic cysteine loop by the hydrophilic loop of the ACh-binding protein enhances protein solubility. *J Biol Chem,* 279**,** 38287-93.

BARNES, P. J., SHAPIRO, S. D. & PAUWELS, R. A. 2003. Chronic obstructive pulmonary disease: molecular and cellular mechanisms. *Eur Respir J,* 22**,** 672-88.

BENFANTE, R., ANTONINI, R. A., DE PIZZOL, M., GOTTI, C., CLEMENTI, F., LOCATI, M. & FORNASARI, D. 2011. Expression of the alpha7 nAChR subunit duplicate form (CHRFAM7A) is down-regulated in the monocytic cell line THP-1 on treatment with LPS. *J Neuroimmunol,* 230**,** 74-84.

BERGET, S. M. 1995. Exon recognition in vertebrate splicing. *J Biol Chem,* 270**,** 2411-4.

BERTRAND, D., BERTRAND, S., CASSAR, S., GUBBINS, E., LI, J. & GOPALAKRISHNAN, M. 2008. Positive allosteric modulation of the alpha7 nicotinic acetylcholine receptor: ligand interactions with distinct binding sites and evidence for a prominent role of the M2-M3 segment. *Mol Pharmacol,* 74**,** 1407-16.

BERTRAND, D., GALZI, J. L., DEVILLERS-THIERY, A., BERTRAND, S. & CHANGEUX, J. P. 1993. Mutations at two distinct sites within the channel domain M2 alter calcium permeability of neuronal alpha 7 nicotinic receptor. *Proc Natl Acad Sci U S A,* 90**,** 6971-5.

BERTRAND D., A. C., JP. 1995. Nicotinic receptor: an allosteric protein specialized for intercellular communication. *seminars in THE NEUROSCIENCES,* 7**,** 75–90.

BEUTLER, B., GREENWALD, D., HULMES, J. D., CHANG, M., PAN, Y. C., MATHISON, J., ULEVITCH, R. & CERAMI, A. 1985a. Identity of

tumour necrosis factor and the macrophage-secreted factor cachectin. *Nature,* 316**,** 552-4.

BEUTLER, B., MILSARK, I. W. & CERAMI, A. C. 1985b. Passive immunization against cachectin/tumor necrosis factor protects mice from lethal effect of endotoxin. *Science,* 229**,** 869-71.

BLACK, D. L. 2003. mechanisms of alternative pre-messenger RNA splicing. *Ann. Rev. Biochem.,* 72**,** 291-336.

BLUMENTHAL, E. M., CONROY, W. G., ROMANO, S. J., KASSNER, P. D. & BERG, D. K. 1997. Detection of functional nicotinic receptors blocked by alpha-bungarotoxin on PC12 cells and dependence of their expression on post-translational events. *J Neurosci,* 17**,** 6094-104.

BOROVIKOVA, L. V., IVANOVA, S., ZHANG, M., YANG, H., BOTCHKINA, G. I., WATKINS, L. R., WANG, H., ABUMRAD, N., EATON, J. W. & TRACEY, K. J. 2000. Vagus nerve stimulation attenuates the systemic inflammatory response to endotoxin. *Nature,* 405**,** 458-62.

BOYD, R. T. 1997. The molecular biology of neuronal nicotinic acetylcholine receptors. *Critical Reviews in Toxicology,* 27**,** 299-318.

BREJC, K., VAN DIJK, W. J., KLAASSEN, R. V., SCHUURMANS, M., VAN DER OOST, J., SMIT, A. B. & SIXMA, T. K. 2001. Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature,* 411**,** 269-76.

BUIST, A. S., MCBURNIE, M. A., VOLLMER, W. M., GILLESPIE, S., BURNEY, P., MANNINO, D. M., MENEZES, A. M., SULLIVAN, S. D., LEE, T. A., WEISS, K. B., JENSEN, R. L., MARKS, G. B., GULSVIK, A. & NIZANKOWSKA-MOGILNICKA, E. 2007. International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study. *Lancet,* 370**,** 741-50.

BURATTI, E., MURO, A. F., GIOMBI, M., GHERBASSI, D., IACONCIG, A. & BARALLE, F. E. 2004. RNA folding affects the recruitment of SR proteins by mouse and human polypurinic enhancer elements in the fibronectin EDA exon. *Mol Cell Biol,* 24**,** 1387-400.

BURGE, C. & KARLIN, S. 1997. Prediction of complete gene structures in human genomic DNA. *J Mol Biol,* 268**,** 78-94.

CANCELLO, R., ZULIAN, A., MAESTRINI, S., MENCARELLI, M., DELLA BARBA, A., INVITTI, C., LIUZZI, A. & DI BLASIO, A. M. 2012. The nicotinic acetylcholine receptor alpha7 in subcutaneous mature adipocytes: downregulation in human obesity and modulation by diet-induced weight loss. *Int J Obes (Lond),* 36**,** 1552-7.

CAPORALE, L. H. 2006 The implicit genome. Cary, NC, USA Oxford University Press

CARLISLE, D. L., HOPKINS, T. M., GAITHER-DAVIS, A., SILHANEK, M. J., LUKETICH, J. D., CHRISTIE, N. A. & SIEGFRIED, J. M. 2004. Nicotine signals through muscle-type and neuronal nicotinic acetylcholine receptors in both human bronchial epithelial cells and airway fibroblasts. *Respir Res,* 5**,** 27.

CARLISLE, D. L., LIU, X., HOPKINS, T. M., SWICK, M. C., DHIR, R. & SIEGFRIED, J. M. 2007. Nicotine activates cell-signaling pathways through muscle-type and neuronal nicotinic acetylcholine receptors in non-small cell lung cancer cells. *Pulm Pharmacol Ther,* 20**,** 629-41.

CARRASCO-SERRANO, C., CAMPOS-CARO, A., VINIEGRA, S., BALLESTA, J. J. & CRIADO, M. 1998. GC- and E-box motifs as regulatory elements in the proximal promoter region of the neuronal nicotinic receptor alpha7 subunit gene. *J Biol Chem,* 273**,** 20021-8.

CARTEGNI, L., WANG, J., ZHU, Z., ZHANG, M. Q. & KRAINER, A. R. 2003. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res,* 31**,** 3568-71.

CASTALDI, P. J., CHO, M. H., COHN, M., LANGERMAN, F., MORAN, S., TARRAGONA, N., MOUKHACHEN, H., VENUGOPAL, R., HASIMJA, D., KAO, E., WALLACE, B., HERSH, C. P., BAGADE, S., BERTRAM, L., SILVERMAN, E. K. & TRIKALINOS, T. A. 2010. The COPD genetic association compendium: a comprehensive online database of COPD genetic associations. *Hum Mol Genet,* 19**,** 526-34.

CELEDON, J. C., LANGE, C., RABY, B. A., LITONJUA, A. A., PALMER, L. J., DEMEO, D. L., REILLY, J. J., KWIATKOWSKI, D. J., CHAPMAN, H. A., LAIRD, N., SYLVIA, J. S., HERNANDEZ, M., SPEIZER, F. E., WEISS, S. T. & SILVERMAN, E. K. 2004. The transforming growth factor-beta1 (TGFB1) gene is associated with chronic obstructive pulmonary disease (COPD). *Hum Mol Genet,* 13**,** 1649-56.

CELLI, B. R., HALBERT, R. J., NORDYKE, R. J. & SCHAU, B. 2005. Airway obstruction in never smokers: results from the Third National Health and Nutrition Examination Survey. *Am J Med,* 118**,** 1364-72.

CHANG, Y. F., IMAM, J.S. AND WILKINSON, M.F. 2007. the nonsense-mediated decay. *RNA surveillance pathway,* 76**,** 51-74.

CHARPANTIER, E., WIESNER, A., HUH, K. H., OGIER, R., HODA, J. C., ALLAMAN, G., RAGGENBASS, M., FEUERBACH, D., BERTRAND, D. & FUHRER, C. 2005. Alpha7 neuronal nicotinic acetylcholine receptors are negatively regulated by tyrosine phosphorylation and Src-family kinases. *J Neurosci,* 25**,** 9836-49.

CHEN, D., DANG, H. & PATRICK, J. W. 1998. Contributions of N-linked glycosylation to the expression of a functional alpha7-nicotinic receptor in Xenopus oocytes. *J Neurochem,* 70**,** 349-57.

CHINI, B., RAIMOND, E., ELGOYHEN, A. B., MORALLI, D., BALZARETTI, M. & HEINEMANN, S. 1994. Molecular cloning and chromosomal localization of the human alpha 7-nicotinic receptor subunit gene (CHRNA7). *Genomics,* 19**,** 379-81.

COOPER, S. T. & MILLAR, N. S. 1997. Host cell-specific folding and assembly of the neuronal nicotinic acetylcholine receptor alpha7 subunit. *J Neurochem,* 68**,** 2140-51.

COOPER, T. A. 2005. Use of minigene systems to dissect alternative splicing elements. *Methods,* 37**,** 331-40.

COOPER, T. A., WAN, L. & DREYFUSS, G. 2009. RNA and disease. *Cell,* 136**,** 777-93.

CORRINGER, P. J., LE NOVERE, N. & CHANGEUX, J. P. 2000. Nicotinic receptors at the amino acid level. *Annu Rev Pharmacol Toxicol,* 40**,** 431-58.

COTE, G. J., HUANG, E. S., JIN, W. & MORRISON, R. S. 1997. Sequence requirements for regulated RNA splicing of the human fibroblast growth factor receptor-1 alpha exon. *J Biol Chem,* 272**,** 1054-60.

COTE J, D. S., JIANG ZH, WU JY 2001. caspase-2 pre-mRNA alternative splicing: identification of an inronic element containing a decoy 3' acceptor site. *PNAS,* 98**,** 983-943.

COUTURIER, S., BERTRAND, D., MATTER, J. M., HERNANDEZ, M. C., BERTRAND, S., MILLAR, N., VALERA, S., BARKAS, T. & BALLIVET, M. 1990. A neuronal nicotinic acetylcholine receptor subunit (alpha 7) is developmentally regulated and forms a homo-oligomeric channel blocked by alpha-BTX. *Neuron,* 5**,** 847-56.

DANI, J. A. & BERTRAND, D. 2007. Nicotinic acetylcholine receptors and nicotinic cholinergic mechanisms of the central nervous system. *Annu Rev Pharmacol Toxicol,* 47**,** 699-729.

DE JONGE, W. J. & ULLOA, L. 2007. The alpha7 nicotinic acetylcholine receptor as a pharmacological target for inflammation. *Br J Pharmacol,* 151**,** 915-29.

DE LUCAS-CERRILLO AM, M. M., ARNALICH F, RENART J ATIENZA G, SERANTES R, ET AL 2011. function of partially duplicated human alpha7 nicotinic receptor subunit CHRFAM7A gene. *JBC,* 7**,** 594-606.

DE SERRES, F. J. 2002. Worldwide racial and ethnic distribution of alpha1-antitrypsin deficiency: summary of an analysis of published genetic epidemiologic surveys. *Chest,* 122**,** 1818-29.

DEWEY, C. N., ROGOZIN, I. B. & KOONIN, E. V. 2006. Compensatory relationship between splice sites and exonic splicing signals depending on the length of vertebrate introns. *BMC Genomics,* 7**,** 311.

DISSET, A., BOURGEOIS, C. F., BENMALEK, N., CLAUSTRES, M., STEVENIN, J. & TUFFERY-GIRAUD, S. 2006. An exon skipping-associated nonsense mutation in the dystrophin gene uncovers a complex interplay between multiple antagonistic splicing elements. *Hum Mol Genet,* 15**,** 999-1013.

DRISDEL, R. C. & GREEN, W. N. 2000. Neuronal alpha-bungarotoxin receptors are alpha7 subunit homomers. *J Neurosci,* 20**,** 133-9.

DRISDEL, R. C., MANZANA, E. & GREEN, W. N. 2004. The role of palmitoylation in functional expression of nicotinic alpha7 receptors. *J Neurosci,* 24**,** 10502-10.

DUJARDIN, G., LAFAILLE, C., PETRILLO, E., BUGGIANO, V., GOMEZ ACUNA, L. I., FISZBEIN, A., GODOY HERZ, M. A., NIETO MORENO, N., MUNOZ, M. J., ALLO, M., SCHOR, I. E. & KORNBLIHTT, A. R. 2012. Transcriptional elongation and alternative splicing. *Biochim Biophys Acta.*

EL-HAJJ, R. A., MCKAY, S. B. & MCKAY, D. B. 2007. Pharmacological and immunological identification of native alpha7 nicotinic receptors: evidence for homomeric and heteromeric alpha7 receptors. *Life Sci,* 81**,** 1317-22.

ENDRIS, V., HACKMANN, K., NEUHANN, T. M., GRASSHOFF, U., BONIN, M., HAUG, U., HAHN, G., SCHALLNER, J. C., SCHROCK, E., TINSCHERT, S., RAPPOLD, G. & MOOG, U. 2010. Homozygous loss of CHRNA7 on chromosome 15q13.3 causes severe encephalopathy with seizures and hypotonia. *Am J Med Genet A,* 152A**,** 2908-11.

FAIRBROTHER, W. G. & CHASIN, L. A. 2000. Human genomic sequences that inhibit splicing. *Mol Cell Biol,* 20**,** 6816-25.

FAIRBROTHER, W. G., YEH, R. F., SHARP, P. A. & BURGE, C. B. 2002. Predictive identification of exonic splicing enhancers in human genes. *Science,* 297**,** 1007-13.

FAUSTINO, N. A. & COOPER, T. A. 2003. Pre-mRNA splicing and human disease. *Genes Dev,* 17**,** 419-37.

FINLAY-SCHULTZ, J., CANASTAR, A., SHORT, M., EL GAZZAR, M., COUGHLAN, C. & LEONARD, S. 2011. Transcriptional repression of the alpha7 nicotinic acetylcholine receptor subunit gene (CHRNA7) by activating protein-2alpha (AP-2alpha). *J Biol Chem,* 286**,** 42123-32.

FLOMEN, R. H., COLLIER, D. A., OSBORNE, S., MUNRO, J., BREEN, G., ST CLAIR, D. & MAKOFF, A. J. 2006. Association study of CHRFAM7A copy number and 2 bp deletion polymorphisms with schizophrenia and bipolar affective disorder. *Am J Med Genet B Neuropsychiatr Genet,* 141B**,** 571-5.

FLOMEN, R. H., DAVIES, A. F., DI FORTI, M., LA CASCIA, C., MACKIE-OGILVIE, C., MURRAY, R. & MAKOFF, A. J. 2008. The copy number variant involving part of the alpha7 nicotinic receptor gene contains a polymorphic inversion. *Eur J Hum Genet,* 16**,** 1364-71.

FU, X. W., LINDSTROM, J. & SPINDEL, E. R. 2009. Nicotine activates and up-regulates nicotinic acetylcholine receptors in bronchial epithelial cells. *Am J Respir Cell Mol Biol,* 41**,** 93-9.

GALZI, J. L., DEVILLERS-THIERY, A., HUSSY, N., BERTRAND, S., CHANGEUX, J. P. & BERTRAND, D. 1992. Mutations in the channel domain of a neuronal nicotinic receptor convert ion selectivity from cationic to anionic. *Nature,* 359**,** 500-5.

GANROT, P. O., LAURELL, C. B. & ERIKSSON, S. 1967. Obstructive lung disease and trypsin inhibitors in alpha-1-antitrypsin deficiency. *Scand J Clin Lab Invest,* 19**,** 205-8.

GARCÍA-GUZMÁN, M., SALA, F., SALA, S., CAMPOS-CARO, A., STÜHMER, W., GUTIÉRREZ, L. M. & CRIADO, M. 1995. alpha-Bungarotoxin-sensitive nicotinic receptors on bovine chromaffin cells: molecular cloning, functional expression and alternative splicing of the alpha 7 subunit. *Eur J Neurosci,* 7**,** 647-55.

GAULT, J., HOPKINS, J., BERGER, R., DREBING, C., LOGEL, J., WALTON, C., SHORT, M., VIANZON, R., OLINCY, A., ROSS, R. G., ADLER, L. E., FREEDMAN, R. & LEONARD, S. 2003. Comparison of polymorphisms in the alpha7 nicotinic receptor gene and its partial duplication in schizophrenic and control subjects. *Am J Med Genet B Neuropsychiatr Genet,* 123B**,** 39-49.

GAULT, J., ROBINSON, M., BERGER, R., DREBING, C., LOGEL, J., HOPKINS, J., MOORE, T., JACOBS, S., MERIWETHER, J., CHOI, M., KIM, E., WALTON, K., BUITING, K., DAVIS, A., BREESE, C., FREEDMAN, R. & LEONARD, S. 1998. Genomic organization and partial duplication of the human alpha7 neuronal nicotinic acetylcholine receptor gene (CHRNA7). *Genomics,* 52**,** 173-85.

GEHLE, V. M., WALCOTT, E. C., NISHIZAKI, T. & SUMIKAWA, K. 1997. N-glycosylation at the conserved sites ensures the expression of properly folded functional ACh receptors. *Brain Res Mol Brain Res,* 45**,** 219-29.

GOLD. 2013. *Global strategy for diagnosis, management and prevention of chronic obstructive pulmonary disease* [Online]. Available:

http://www.goldcopd.org/uploads/users/files/GOLD_Report_2013_Feb2 0.pdf 2013].

GOREN, A., RAM, O., AMIT, M., KEREN, H., LEV-MAOR, G., VIG, I., PUPKO, T. & AST, G. 2006. Comparative analysis identifies exonic splicing regulatory sequences--The complex definition of enhancers and silencers. *Mol Cell,* 22**,** 769-81.

GOVIND, A. P., VEZINA, P. & GREEN, W. N. 2009. Nicotine-induced upregulation of nicotinic receptors: underlying mechanisms and relevance to nicotine addiction. *Biochem Pharmacol,* 78**,** 756-65.

GREEN, W. N. & MILLAR, N. S. 1995. Ion-channel assembly. *Trends Neurosci,* 18**,** 280-7.

GREEN, W. N., ROSS, A. F. & CLAUDIO, T. 1991. Acetylcholine receptor assembly is stimulated by phosphorylation of its gamma subunit. *Neuron,* 7**,** 659-66.

GROMOLL, J., LAHRMANN, L., GODMANN, M., MULLER, T., MICHEL, C., STAMM, S. & SIMONI, M. 2007. Genomic checkpoints for exon 10 usage in the luteinizing hormone receptor type 1 and type 2. *Mol Endocrinol,* 21**,** 1984-96.

GROSSO, A. R., GOMES, A.Q., BARBOSA-MORAIS, N.L., CALDERIA, S., THORNE, N.P., ET AL 2008. tissue-specific splicing factor gene expression signatures. *Nucleic Acids Res,* 36**,** 4823-4832.

GUO, X. & WECKER, L. 2002. Identification of three cAMP-dependent protein kinase (PKA) phosphorylation sites within the major intracellular domain of neuronal nicotinic receptor alpha4 subunits. *J Neurochem,* 82**,** 439-47.

GUSLANDI, M. 1999. Nicotine treatment for ulcerative colitis. *Br J Clin Pharmacol,* 48**,** 481-4.

HALEVI, S., MCKAY, J., PALFREYMAN, M., YASSIN, L., ESHEL, M., JORGENSEN, E. & TREININ, M. 2002. The C. elegans ric-3 gene is required for maturation of nicotinic acetylcholine receptors. *EMBO J,* 21**,** 1012-20.

HALEVI, S., YASSIN, L., ESHEL, M., SALA, F., SALA, S., CRIADO, M. & TREININ, M. 2003. Conservation within the RIC-3 gene family. Effectors of mammalian nicotinic acetylcholine receptor expression. *J Biol Chem,* 278**,** 34411-7.

HERSH, C. P., DEMEO, D. L., LANGE, C., LITONJUA, A. A., REILLY, J. J., KWIATKOWSKI, D., LAIRD, N., SYLVIA, J. S., SPARROW, D., SPEIZER, F. E., WEISS, S. T. & SILVERMAN, E. K. 2005. Attempted replication of reported chronic obstructive pulmonary disease candidate gene associations. *Am J Respir Cell Mol Biol,* 33**,** 71-8.

HERTEL, K. J. 2008. Combinatorial control of exon recognition. *J Biol Chem,* 283**,** 1211-5.

HOLTON, T. A. & GRAHAM, M. W. 1991. A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. *Nucleic Acids Res,* 19**,** 1156.

HOPPMAN-CHANEY, N., WAIN, K., SEGER, P., SUPERNEAU, D. & HODGE, J. 2013. Identification of single gene deletions at 15q13.3: further evidence that CHRNA7 causes the 15q13.3 microdeletion syndrome phenotype. *Clin Genet,* 83**,** 345-51.

HUBERT, H. B., FABSITZ, R. R., FEINLEIB, M. & GWINN, C. 1982. Genetic and environmental influences on pulmonary function in adult twins. *Am Rev Respir Dis,* 125**,** 409-15.

HUNG, R. J., MCKAY, J. D., GABORIEAU, V., BOFFETTA, P., HASHIBE, M., ZARIDZE, D., MUKERIA, A., SZESZENIA-DABROWSKA, N., LISSOWSKA, J., RUDNAI, P., FABIANOVA, E., MATES, D., BENCKO, V., FORETOVA, L., JANOUT, V., CHEN, C., GOODMAN, G., FIELD, J. K., LILOGLOU, T., XINARIANOS, G., CASSIDY, A., MCLAUGHLIN, J., LIU, G., NAROD, S., KROKAN, H. E., SKORPEN, F., ELVESTAD, M. B., HVEEM, K., VATTEN, L., LINSEISEN, J., CLAVEL-CHAPELON, F., VINEIS, P., BUENO-DE-MESQUITA, H. B., LUND, E., MARTINEZ, C., BINGHAM, S., RASMUSON, T., HAINAUT, P., RIBOLI, E., AHRENS, W., BENHAMOU, S., LAGIOU, P., TRICHOPOULOS, D., HOLCATOVA, I., MERLETTI, F., KJAERHEIM, K., AGUDO, A., MACFARLANE, G., TALAMINI, R., SIMONATO, L., LOWRY, R., CONWAY, D. I., ZNAOR, A., HEALY, C., ZELENIKA, D., BOLAND, A., DELEPINE, M., FOGLIO, M., LECHNER, D., MATSUDA, F., BLANCHE, H., GUT, I., HEATH, S., LATHROP, M. & BRENNAN, P. 2008. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature,* 452**,** 633-7.

HURST, R., ROLLEMA, H. & BERTRAND, D. 2013. Nicotinic acetylcholine receptors: from basic science to therapeutics. *Pharmacol Ther,* 137**,** 22-54.

HUSTON, J. M., OCHANI, M., ROSAS-BALLINA, M., LIAO, H., OCHANI, K., PAVLOV, V. A., GALLOWITSCH-PUERTA, M., ASHOK, M., CZURA, C. J., FOXWELL, B., TRACEY, K. J. & ULLOA, L. 2006. Splenectomy inactivates the cholinergic antiinflammatory pathway during lethal endotoxemia and polymicrobial sepsis. *J Exp Med,* 203**,** 1623-8.

IBRAHIM, E. C., SCHAAL, T. D., HERTEL, K. J., REED, R. & MANIATIS, T. 2005. Serine/arginine-rich protein-dependent suppression of exon skipping by exonic splicing enhancers. *Proc Natl Acad Sci U S A,* 102**,** 5002-7.

IZAURRALDE, E., LEWIS, J., MCGUIGAN, C., JANKOWSKA, M., DARZYNKIEWICZ, E. & MATTAJ, I. W. 1994. A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell,* 78**,** 657-68.

JANSEN, M., BALI, M. & AKABAS, M. H. 2008. Modular design of Cys-loop ligand-gated ion channels: functional 5-HT3 and GABA rho1 receptors lacking the large cytoplasmic M3M4 loop. *J Gen Physiol,* 131**,** 137-46.

JONES, D. T. 2007. Improving the accuracy of transmembrane protein topology prediction using evolutionary information. *Bioinformatics,* 23**,** 538-44.

JUMAA, H. & NIELSEN, P. J. 1997. The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation. *EMBO J,* 16**,** 5077-85.

JUNG, M. Y., LORENZ, L. & RICHTER, J. D. 2006. Translational control by neuroguidin, a eukaryotic initiation factor 4E and CPEB binding protein. *Mol Cell Biol,* 26**,** 4277-87.

KALSHEKER, N. & CHAPPELL, S. 2008. The new genetics and chronic obstructive pulmonary disease. *COPD,* 5**,** 257-64.

KASHIMA, T. & MANLEY, J. L. 2003. A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy. *Nat Genet,* 34**,** 460-3.

KAWASHIMA, K., FUJII, T., MORIWAKI, Y. & MISAWA, H. 2012. Critical roles of acetylcholine and the muscarinic and nicotinic acetylcholine receptors in the regulation of immune function. *Life Sci,* 91**,** 1027-32.

KELEMEN, O., CONVERTINI, P., ZHANG, Z., WEN, Y., SHEN, M., FALALEEVA, M. & STAMM, S. 2012. Function of alternative splicing. *Gene*.

KHIROUG, S. S., HARKNESS, P. C., LAMB, P. W., SUDWEEKS, S. N., KHIROUG, L., MILLAR, N. S. & YAKEL, J. L. 2002. Rat nicotinic ACh receptor alpha7 and beta2 subunits co-assemble to form functional heteromeric nicotinic receptor channels. *J Physiol,* 540**,** 425-34.

KIMURA, K., WAKAMATSU, A., SUZUKI, Y., OTA, T., NISHIKAWA, T., YAMASHITA, R., YAMAMOTO, J., SEKINE, M., TSURITANI, K., WAKAGURI, H., ISHII, S., SUGIYAMA, T., SAITO, K., ISONO, Y., IRIE, R., KUSHIDA, N., YONEYAMA, T., OTSUKA, R., KANDA, K., YOKOI, T., KONDO, H., WAGATSUMA, M., MURAKAWA, K., ISHIDA, S., ISHIBASHI, T., TAKAHASHI-FUJII, A., TANASE, T., NAGAI, K., KIKUCHI, H., NAKAI, K., ISOGAI, T. & SUGANO, S. 2006. Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res,* 16**,** 55-65.

KOHTZ, J. D., JAMISON, S.F., WILL, C.L., ZUO, P., LUHRMANN, R., GARCIA-BLANCO, M.A., ET AL 1994. protein-protein interactions and 5' splice-site recognition in mammalian mRNA precursors. *Nature,* 368**,** 119-124.

KOJIMA, J., ARAYA, J., HARA, H., ITO, S., TAKASAKA, N., KOBAYASHI, K., FUJII, S., TSURUSHIGE, C., NUMATA, T., ISHIKAWA, T., SHIMIZU, K., KAWAISHI, M., SAITO, K., KAMIYA, N., HIRANO, J., ODAKA, M., MORIKAWA, T., HANO, H., ARAI, S., MIYAZAKI, T., KANEKO, Y., NAKAYAMA, K. & KUWANO, K. 2013. Apoptosis inhibitor of macrophage (AIM) expression in alveolar macrophages in COPD. *Respir Res,* 14**,** 30.

KOSAKI, A., NELSON, J. & WEBSTER, N. J. 1998. Identification of intron and exon sequences involved in alternative splicing of insulin receptor pre-mRNA. *J Biol Chem,* 273**,** 10331-7.

KOSTI, I., RADIVOJAC, P. & MANDEL-GUTFREUND, Y. 2012. An integrated regulatory network reveals pervasive cross-regulation among transcription and splicing factors. *PLoS Comput Biol,* 8**,** e1002603.

KOX, M., POMPE, J. C., PETERS, E., VANEKER, M., VAN DER LAAK, J. W., VAN DER HOEVEN, J. G., SCHEFFER, G. J., HOEDEMAEKERS, C. W. & PICKKERS, P. 2011. alpha7 nicotinic acetylcholine receptor agonist GTS-21 attenuates ventilator-induced tumour necrosis factor-alpha production and lung injury. *Br J Anaesth,* 107**,** 559-66.

KRACUN, S., HARKNESS, P. C., GIBB, A. J. & MILLAR, N. S. 2008. Influence of the M3-M4 intracellular domain upon nicotinic acetylcholine receptor assembly, targeting and function. *Br J Pharmacol,* 153**,** 1474-84.

KUEPPERS, F., MILLER, R. D., GORDON, H., HEPPER, N. G. & OFFORD, K. 1977. Familial prevalence of chronic obstructive pulmonary disease in a matched pair study. *Am J Med,* 63**,** 336-42.

KUMAR, P. A. C., M. c2012. *Kumar and Clark's Clinical Medicine*, Edinburgh : Saunders Elsevier, c2012. .

LAM, D. C., GIRARD, L., RAMIREZ, R., CHAU, W. S., SUEN, W. S., SHERIDAN, S., TIN, V. P., CHUNG, L. P., WONG, M. P., SHAY, J. W., GAZDAR, A. F., LAM, W. K. & MINNA, J. D. 2007. Expression of nicotinic acetylcholine receptor subunit genes in non-small-cell lung cancer reveals differences between smokers and nonsmokers. *Cancer Res,* 67**,** 4638-47.

LANSDELL, S. J., GEE, V. J., HARKNESS, P. C., DOWARD, A. I., BAKER, E. R., GIBB, A. J. & MILLAR, N. S. 2005. RIC-3 enhances functional expression of multiple nicotinic acetylcholine receptor subtypes in mammalian cells. *Mol Pharmacol,* 68**,** 1431-8.

LARKIN, M. A., BLACKSHIELDS, G., BROWN, N. P., CHENNA, R., MCGETTIGAN, P. A., MCWILLIAM, H., VALENTIN, F., WALLACE, I. M., WILM, A., LOPEZ, R., THOMPSON, J. D., GIBSON, T. J. & HIGGINS, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics,* 23**,** 2947-8.

LAURELL, C. B. & ERIKSSON, S. 1963. ELECTROPHORETIC ALPHA1-GLOBULIN PATTERN OF SERUM IN ALPHA1-ANTITRYPSIN DEFICIENCY. *Scandinavian Journal of Clinical & Laboratory Investigation,* 15**,** 132-&.

LE NOVERE, N. & CHANGEUX, J. P. 1995. Molecular evolution of the nicotinic acetylcholine receptor: an example of multigene family in excitable cells. *J Mol Evol,* 40**,** 155-72.

LI, J., HAKATA, Y., TAKEDA, E., LIU, Q., IWATANI, Y., KOZAK, C. A. & MIYAZAWA, M. 2012. Two genetic determinants acquired late in mus evolution regulate the inclusion of exon 5, which alters mouse APOBEC3 translation efficiency. *PLoS Pathog,* 8**,** e1002478.

LI, Q. Y., HUANG, S. G., WAN, H. Y., WU, H. C., ZHOU, T., LI, M. & DENG, W. W. 2007. Effect of smoking cessation on airway inflammation of rats with chronic bronchitis. *Chin Med J (Engl),* 120**,** 1511-6.

LIU, F. & GONG, C. X. 2008. Tau exon 10 alternative splicing and tauopathies. *Mol Neurodegener,* 3**,** 8.

LIU, Q., HUANG, Y., XUE, F., SIMARD, A., DECHON, J., LI, G., ZHANG, J., LUCERO, L., WANG, M., SIERKS, M., HU, G., CHANG, Y., LUKAS, R. J. & WU, J. 2009. A novel nicotinic acetylcholine receptor subtype in basal forebrain cholinergic neurons with high sensitivity to amyloid peptides. *J Neurosci,* 29**,** 918-29.

LOCKE, D. P., ARCHIDIACONO, N., MISCEO, D., CARDONE, M. F., DESCHAMPS, S., ROE, B., ROCCHI, M. & EICHLER, E. E. 2003. Refinement of a chimpanzee pericentric inversion breakpoint to a segmental duplication cluster. *Genome Biol,* 4**,** R50.

LOKKE, A., LANGE, P., SCHARLING, H., FABRICIUS, P. & VESTBO, J. 2006. Developing COPD: a 25 year follow up study of the general population. *Thorax,* 61**,** 935-9.

LOPEZ, A. D., SHIBUYA, K., RAO, C., MATHERS, C. D., HANSELL, A. L., HELD, L. S., SCHMID, V. & BUIST, S. 2006. Chronic obstructive pulmonary disease: current burden and future projections. *Eur Respir J,* 27**,** 397-412.

LUMMIS, S. C., BEENE, D. L., LEE, L. W., LESTER, H. A., BROADHURST, R. W. & DOUGHERTY, D. A. 2005. Cis-trans isomerization at a proline

opens the pore of a neurotransmitter-gated ion channel. *Nature,* 438**,** 248-52.

LUNDBACK, B., LINDBERG, A., LINDSTROM, M., RONMARK, E., JONSSON, A. C., JONSSON, E., LARSSON, L. G., ANDERSSON, S., SANDSTROM, T. & LARSSON, K. 2003. Not 15 but 50% of smokers develop COPD?--Report from the Obstructive Lung Disease in Northern Sweden Studies. *Respir Med,* 97**,** 115-22.

LYNN DOUCETTE-STAMM, L. M., MONTEGGIA, DIANA DONNELLY-ROBERTS, MEI TAI WANG, JOHN LEE, JINGXIANG TIAN, AND TONY GIORDANO 1993. Cloning and sequence of the human alpha7 nicotinic receptor.

MABLEY, J. G., PACHER, P., SOUTHAN, G. J., SALZMAN, A. L. & SZABO, C. 2002. Nicotine reduces the incidence of type I diabetes in mice. *J Pharmacol Exp Ther,* 300**,** 876-81.

MAKOFF, A. J. & FLOMEN, R. H. 2007. Detailed analysis of 15q11-q14 sequence corrects errors and gaps in the public access sequence to fully reveal large segmental duplications at breakpoints for Prader-Willi, Angelman, and inv dup(15) syndromes. *Genome Biol,* 8**,** R114.

MANNINO, D. M., BUIST, A. S., PETTY, T. L., ENRIGHT, P. L. & REDD, S. C. 2003. Lung function and mortality in the United States: data from the First National Health and Nutrition Examination Survey follow up study. *Thorax,* 58**,** 388-93.

MAOUCHE, K., POLETTE, M., JOLLY, T., MEDJBER, K., CLOEZ-TAYARANI, I., CHANGEUX, J. P., BURLET, H., TERRYN, C., CORAUX, C., ZAHM, J. M., BIREMBAUT, P. & TOURNIER, J. M. 2009. {alpha}7 nicotinic acetylcholine receptor regulates airway epithelium differentiation by controlling basal cell proliferation. *Am J Pathol,* 175**,** 1868-82.

MASAHIRO, G., MASASHI, S., SHIGEKI, M., SHIGEKI, M. 2005. high performance system for signal peptide prediction: sosui signal. *Chem Bio Inform J,* 4**,** 142-147.

MASUREL-PAULET, A., ANDRIEUX, J., CALLIER, P., CUISSET, J. M., LE CAIGNEC, C., HOLDER, M., THAUVIN-ROBINET, C., DORAY, B., FLORI, E., ALEX-CORDIER, M. P., BERI, M., BOUTE, O., DELOBEL, B., DIEUX, A., VALLEE, L., JAILLARD, S., ODENT, S., ISIDOR, B., BENETEAU, C., VIGNERON, J., BILAN, F., GILBERT-DUSSARDIER, B., DUBOURG, C., LABALME, A., BIDON, C., GAUTIER, A., PERNES, P., PINOIT, J. M., HUET, F., MUGNERET, F., ARAL, B., JONVEAUX, P., SANLAVILLE, D. & FAIVRE, L. 2010. Delineation of 15q13.3 microdeletions. *Clin Genet,* 78**,** 149-61.

MATLIN, A. J., CLARK, F. & SMITH, C. W. 2005. Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol,* 6**,** 386-98.

MCCLEARN, G. E., SVARTENGREN, M., PEDERSEN, N. L., HELLER, D. A. & PLOMIN, R. 1994. Genetic and environmental influences on pulmonary function in aging Swedish twins. *J Gerontol,* 49**,** 264-8.

MCCLOSKEY, S. C., PATEL, B. D., HINCHLIFFE, S. J., REID, E. D., WAREHAM, N. J. & LOMAS, D. A. 2001. Siblings of patients with severe chronic obstructive pulmonary disease have a significant risk of airflow obstruction. *Am J Respir Crit Care Med,* 164**,** 1419-24.

MILLAR, N. S. 2003. Assembly and subunit diversity of nicotinic acetylcholine receptors. *Biochem Soc Trans,* 31**,** 869-74.

MODREK, B. & LEE, C. 2002. A genomic view of alternative splicing. *Nat Genet,* 30**,** 13-9.

MOSS, S. J. & SMART, T. G. 2001. Constructing inhibitory synapses. *Nat Rev Neurosci,* 2**,** 240-50.

NATHAN, C. 2002. Points of control in inflammation. *Nature,* 420**,** 846-52.

NOTT, A., LE HIR, H. & MOORE, M. J. 2004. Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. *Genes Dev,* 18**,** 210-22.

NURSESLABS. 2012. *Chronic bronchitis/emphysema pathophysiology and schematic diagram* [Online]. Available: http://nurseslabs.com/tag/copd/.

OLOFSSON, P. S., ROSAS-BALLINA, M., LEVINE, Y. A. & TRACEY, K. J. 2012. Rethinking inflammation: neural circuits in the regulation of immunity. *Immunol Rev,* 248**,** 188-204.

ORR-URTREGER, A., GOLDNER, F. M., SAEKI, M., LORENZO, I., GOLDBERG, L., DE BIASI, M., DANI, J. A., PATRICK, J. W. & BEAUDET, A. L. 1997. Mice deficient in the alpha7 neuronal nicotinic acetylcholine receptor lack alpha-bungarotoxin binding sites and hippocampal fast nicotinic currents. *J Neurosci,* 17**,** 9165-71.

PALEARI, L., CATASSI, A., CIARLO, M., CAVALIERI, Z., BRUZZO, C., SERVENT, D., CESARIO, A., CHESSA, L., CILLI, M., PICCARDI, F., GRANONE, P. & RUSSO, P. 2008. Role of alpha7-nicotinic acetylcholine receptor in human non-small cell lung cancer proliferation. *Cell Prolif,* 41**,** 936-59.

PAVLOV, V. A., PARRISH, W. R., ROSAS-BALLINA, M., OCHANI, M., PUERTA, M., OCHANI, K., CHAVAN, S., AL-ABED, Y. & TRACEY, K. J. 2009. Brain acetylcholinesterase activity controls systemic cytokine levels through the cholinergic anti-inflammatory pathway. *Brain Behav Immun,* 23**,** 41-5.

PAVLOV, V. A. & TRACEY, K. J. 2004. Neural regulators of innate immune responses and inflammation. *Cell Mol Life Sci,* 61**,** 2322-31.

PAVLOV, V. A. & TRACEY, K. J. 2006. Controlling inflammation: the cholinergic anti-inflammatory pathway. *Biochem Soc Trans,* 34**,** 1037-40.

PAVLOV, V. A., WANG, H., CZURA, C. J., FRIEDMAN, S. G. & TRACEY, K. J. 2003. The cholinergic anti-inflammatory pathway: a missing link in neuroimmunomodulation. *Mol Med,* 9**,** 125-34.

PEARSON, T. A. & MANOLIO, T. A. 2008. How to interpret a genome-wide association study. *JAMA,* 299**,** 1335-44.

PENG, J. H., LUCERO, L., FRYER, J., HERL, J., LEONARD, S. S. & LUKAS, R. J. 1999. Inducible, heterologous expression of human alpha7-nicotinic acetylcholine receptors in a native nicotinic receptor-null human clonal line. *Brain Res,* 825**,** 172-9.

PENG, X., KATZ, M., GERZANICH, V., ANAND, R. & LINDSTROM, J. 1994. Human alpha 7 acetylcholine receptor: cloning of the alpha 7 subunit from the SH-SY5Y cell line and determination of pharmacological properties of native receptors and functional alpha 7 homomers expressed in Xenopus oocytes. *Mol Pharmacol,* 45**,** 546-54.

PHELPS, S. F., HAUSER, M. A., COLE, N. M., RAFAEL, J. A., HINKLE, R. T., FAULKNER, J. A. & CHAMBERLAIN, J. S. 1995. Expression of full-length and truncated dystrophin mini-genes in transgenic mdx mice. *Hum Mol Genet,* 4**,** 1251-8.

PIERCE, B. A. (ed.) 2006. *Genetics a conceptional approach,* New York: W.H.Freeman and Company.

PILLAI, S. G., GE, D., ZHU, G., KONG, X., SHIANNA, K. V., NEED, A. C., FENG, S., HERSH, C. P., BAKKE, P., GULSVIK, A., RUPPERT, A., LODRUP CARLSEN, K. C., ROSES, A., ANDERSON, W., RENNARD, S. I., LOMAS, D. A., SILVERMAN, E. K. & GOLDSTEIN, D. B. 2009. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genet,* 5**,** e1000421.

PIVA, F., GIULIETTI, M., BURINI, A. B. & PRINCIPATO, G. 2012. SpliceAid 2: a database of human splicing factors expression data and RNA target motifs. *Hum Mutat,* 33**,** 81-5.

PLACZEK, A. N., GRASSI, F., MEYER, E. M. & PAPKE, R. L. 2005. An alpha7 nicotinic acetylcholine receptor gain-of-function mutant that retains pharmacological fidelity. *Mol Pharmacol,* 68**,** 1863-76.

PLUMMER, H. K. R., DHAR, M. & SCHULLER, H. M. 2005. Expression of the alpha7 nicotinic acetylcholine receptor in human lung cells. *Respir Res,* 6**,** 29.

PONS, S., SALLETTE, J., BOURGEOIS, J. P., TALY, A., CHANGEUX, J. P. & DEVILLERS-THIERY, A. 2004. Critical role of the C-terminal segment in the maturation and export to the cell surface of the homopentameric alpha 7-5HT3A receptor. *Eur J Neurosci,* 20**,** 2022-30.

RAKHILIN, S., DRISDEL, R. C., SAGHER, D., MCGEHEE, D. S., VALLEJO, Y. & GREEN, W. N. 1999. alpha-bungarotoxin receptors contain alpha7 subunits in two different disulfide-bonded conformations. *J Cell Biol,* 146**,** 203-18.

RANGWALA, F., DRISDEL, R. C., RAKHILIN, S., KO, E., ATLURI, P., HARKINS, A. B., FOX, A. P., SALMAN, S. S. & GREEN, W. N. 1997. Neuronal alpha-bungarotoxin receptors differ structurally from other nicotinic acetylcholine receptors. *J Neurosci,* 17**,** 8201-12.

REDLINE, S., TISHLER, P. V., LEWITTER, F. I., TAGER, I. B., MUNOZ, A. & SPEIZER, F. E. 1987. Assessment of genetic and nongenetic influences on pulmonary function. A twin study. *Am Rev Respir Dis,* 135**,** 217-22.

REED, R. & MANIATIS, T. 1986. A role for exon sequences and splice-site proximity in splice-site selection. *Cell,* 46**,** 681-90.

REESE, M. G., EECKMAN, F. H., KULP, D. & HAUSSLER, D. 1997. Improved splice site detection in Genie. *J Comput Biol,* 4**,** 311-23.

REPAPI, E., SAYERS, I., WAIN, L. V., BURTON, P. R., JOHNSON, T., OBEIDAT, M., ZHAO, J. H., RAMASAMY, A., ZHAI, G., VITART, V., HUFFMAN, J. E., IGL, W., ALBRECHT, E., DELOUKAS, P., HENDERSON, J., GRANELL, R., MCARDLE, W. L., RUDNICKA, A. R., BARROSO, I., LOOS, R. J., WAREHAM, N. J., MUSTELIN, L., RANTANEN, T., SURAKKA, I., IMBODEN, M., WICHMANN, H. E., GRKOVIC, I., JANKOVIC, S., ZGAGA, L., HARTIKAINEN, A. L., PELTONEN, L., GYLLENSTEN, U., JOHANSSON, A., ZABOLI, G., CAMPBELL, H., WILD, S. H., WILSON, J. F., GLASER, S., HOMUTH, G., VOLZKE, H., MANGINO, M., SORANZO, N., SPECTOR, T. D., POLASEK, O., RUDAN, I., WRIGHT, A. F., HELIOVAARA, M., RIPATTI, S., POUTA, A., NALUAI, A. T., OLIN, A. C., TOREN, K., COOPER, M. N., JAMES, A. L., PALMER, L. J.,

HINGORANI, A. D., WANNAMETHEE, S. G., WHINCUP, P. H., SMITH, G. D., EBRAHIM, S., MCKEEVER, T. M., PAVORD, I. D., MACLEOD, A. K., MORRIS, A. D., PORTEOUS, D. J., COOPER, C., DENNISON, E., SHAHEEN, S., KARRASCH, S., SCHNABEL, E., SCHULZ, H., GRALLERT, H., BOUATIA-NAJI, N., DELPLANQUE, J., FROGUEL, P., BLAKEY, J. D., BRITTON, J. R., MORRIS, R. W., HOLLOWAY, J. W., LAWLOR, D. A., HUI, J., NYBERG, F., JARVELIN, M. R., JACKSON, C., KAHONEN, M., KAPRIO, J., PROBST-HENSCH, N. M., KOCH, B., HAYWARD, C., EVANS, D. M., ELLIOTT, P., STRACHAN, D. P., HALL, I. P. & TOBIN, M. D. 2010. Genome-wide association study identifies five loci associated with lung function. *Nat Genet*, 42**,** 36-44.

RILEY, B., WILLIAMSON, M., COLLIER, D., WILIKIE, H., AND MAKOFF, A. 2002. a 3-Mb map of a large segmental duplication overlapping the alpha-7 nicotinic acetylcholine receptor gene (CHRNA7) at human 15q13-q14. *Genomics,* 79**,** 197-209.

ROBBERSON, B. L., COTE, G. J. & BERGET, S. M. 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol Cell Biol,* 10**,** 84-94.

ROSAS-BALLINA, M., OCHANI, M., PARRISH, W. R., OCHANI, K., HARRIS, Y. T., HUSTON, J. M., CHAVAN, S. & TRACEY, K. J. 2008. Splenic nerve is required for cholinergic antiinflammatory pathway control of TNF in endotoxemia. *Proc Natl Acad Sci U S A,* 105**,** 11008-13.

ROSAS-BALLINA, M., OLOFSSON, P. S., OCHANI, M., VALDES-FERRER, S. I., LEVINE, Y. A., REARDON, C., TUSCHE, M. W., PAVLOV, V. A., ANDERSSON, U., CHAVAN, S., MAK, T. W. & TRACEY, K. J. 2011. Acetylcholine-synthesizing T cells relay neural signals in a vagus nerve circuit. *Science,* 334**,** 98-101.

ROSAS-BALLINA, M. & TRACEY, K. J. 2009. The neurology of the immune system: neural reflexes regulate immunity. *Neuron,* 64**,** 28-32.

SAEED, R. W., VARMA, S., PENG-NEMEROFF, T., SHERRY, B., BALAKHANEH, D., HUSTON, J., TRACEY, K. J., AL-ABED, Y. & METZ, C. N. 2005. Cholinergic stimulation blocks endothelial cell activation and leukocyte recruitment during inflammation. *J Exp Med,* 201**,** 1113-23.

SARAGOZA, P., MODIR, J., GOEL, N., FRENCH, K., LI, L., NOWAK, M. & STITZEL, J. 2003. Identification of an alternatively processed nicotinic receptor alpha7 subunit RNA in mouse brain. *Brain Res Mol Brain Res,* 117**,** 15-26.

SCHEDEL, A., THORNTON, S., SCHLOSS, P., KLUTER, H. & BUGERT, P. 2011. Human platelets express functional alpha7-nicotinic acetylcholine receptors. *Arterioscler Thromb Vasc Biol,* 31**,** 928-34.

SCHMUCKER, D., CLEMENTS, J., SHU, H., WORBY, C.A., XIAO, J., MUDA, M., ET AL 2000. Drosophila Dscam is an Axon Guidance Receptor Exhibiting Extraordinary Molecular Diversity. *Cell,* 101**,** 671-684.

SEGUELA, P., WADICHE, J., DINELEY-MILLER, K., DANI, J. A. & PATRICK, J. W. 1993. Molecular cloning, functional properties, and distribution of rat brain alpha 7: a nicotinic cation channel highly permeable to calcium. *J Neurosci,* 13**,** 596-604.

SEVERANCE, E. & YOLKEN, R. 2008. Novel alpha7 nicotinic receptor isoforms and deficient cholinergic transcription in schizophrenia. *Genes Brain Behav,* 7**,** 37-45.

SEVERANCE, E., ZHANG, H., CRUZ, Y., PAKHLEVANIANTS, S., HADLEY, S., AMIN, J., WECKER, L., REED, C. & CUEVAS, J. 2004. The alpha7 nicotinic acetylcholine receptor subunit exists in two isoforms that contribute to functional ligand-gated ion channels. *Mol Pharmacol,* 66**,** 420-9.

SEVERANCE, E. G., DICKERSON, F. B., STALLINGS, C. R., ORIGONI, A. E., SULLENS, A., MONSON, E. T. & YOLKEN, R. H. 2009. Differentiating nicotine- versus schizophrenia-associated decreases of the alpha7 nicotinic acetylcholine receptor transcript, CHRFAM7A, in peripheral blood lymphocytes. *J Neural Transm,* 116**,** 213-20.

SHINAWI, M., SCHAAF, C. P., BHATT, S. S., XIA, Z., PATEL, A., CHEUNG, S. W., LANPHER, B., NAGL, S., HERDING, H. S., NEVINNY-STICKEL, C., IMMKEN, L. L., PATEL, G. S., GERMAN, J. R., BEAUDET, A. L. & STANKIEWICZ, P. 2009. A small recurrent deletion within 15q13.3 is associated with a range of neurodevelopmental phenotypes. *Nat Genet,* 41**,** 1269-71.

SILVERMAN, E. K., CHAPMAN, H. A., DRAZEN, J. M., WEISS, S. T., ROSNER, B., CAMPBELL, E. J., O'DONNELL, W. J., REILLY, J. J., GINNS, L., MENTZER, S., WAIN, J. & SPEIZER, F. E. 1998. Genetic epidemiology of severe, early-onset chronic obstructive pulmonary disease. Risk to relatives for airflow obstruction and chronic bronchitis. *Am J Respir Crit Care Med,* 157**,** 1770-8.

SIRONI, M., MENOZZI, G., RIVA, L., CAGLIANI, R., COMI, G. P., BRESOLIN, N., GIORDA, R. & POZZOLI, U. 2004. Silencer elements as possible inhibitors of pseudoexon splicing. *Nucleic Acids Res,* 32**,** 1783-91.

SMITH, C. 2000     alternative pre-mRNA splicing: the logic of combinatorial control     *Trends Biochem. Sci*     25**,** 381-288.

SMITH, C. W. J. A. N.-G., B. 1989. mutually exclusive splicing of alpha tropomysin exons enforced by an unusual lariat branch point location: implications for constitutive splicing. *Cell,* 56**,** 749-758.

SMITH, P. J., ZHANG, C., WANG, J., CHEW, S. L., ZHANG, M. Q. & KRAINER, A. R. 2006. An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Hum Mol Genet,* 15**,** 2490-508.

SMOLONSKA, J., WIJMENGA, C., POSTMA, D. S. & BOEZEN, H. M. 2009. Meta-analyses on suspected chronic obstructive pulmonary disease genes: a summary of 20 years' research. *Am J Respir Crit Care Med,* 180**,** 618-31.

SPIELMANN, M., REICHELT, G., HERTZBERG, C., TRIMBORN, M., MUNDLOS, S., HORN, D. & KLOPOCKI, E. 2011. Homozygous deletion of chromosome 15q13.3 including CHRNA7 causes severe mental retardation, seizures, muscular hypotonia, and the loss of KLF13 and TRPM1 potentially cause macrocytosis and congenital retinal dysfunction in siblings. *Eur J Med Genet,* 54**,** e441-5.

STOLLER, J. K. & ABOUSSOUAN, L. S. 2005. Alpha1-antitrypsin deficiency. *Lancet,* 365**,** 2225-36.

STRACHAN, T. A. R., A. 2010. *HUMAN MOLECULAR GENETICS.*

SU, X., LEE, J. W., MATTHAY, Z. A., MEDNICK, G., UCHIDA, T., FANG, X., GUPTA, N. & MATTHAY, M. A. 2007. Activation of the alpha7 nAChR reduces acid-induced acute lung injury in mice and rats. *Am J Respir Cell Mol Biol,* 37**,** 186-92.

SUN, H. A. C., L.A. 2000. multiple splicing defects in an inronic false exon. *Mol Cell Biol,* 20**,** 6414-6425.

TALY, A., CORRINGER, P. J., GUEDIN, D., LESTAGE, P. & CHANGEUX, J. P. 2009. Nicotinic receptors: allosteric transitions and therapeutic targets in the nervous system. *Nat Rev Drug Discov,* 8**,** 733-50.

THORGEIRSSON, T. E., GELLER, F., SULEM, P., RAFNAR, T., WISTE, A., MAGNUSSON, K. P., MANOLESCU, A., THORLEIFSSON, G., STEFANSSON, H., INGASON, A., STACEY, S. N., BERGTHORSSON, J. T., THORLACIUS, S., GUDMUNDSSON, J., JONSSON, T., JAKOBSDOTTIR, M., SAEMUNDSDOTTIR, J., OLAFSDOTTIR, O., GUDMUNDSSON, L. J., BJORNSDOTTIR, G., KRISTJANSSON, K., SKULADOTTIR, H., ISAKSSON, H. J., GUDBJARTSSON, T., JONES, G. T., MUELLER, T., GOTTSATER, A., FLEX, A., ABEN, K. K., DE VEGT, F., MULDERS, P. F., ISLA, D., VIDAL, M. J., ASIN, L., SAEZ, B., MURILLO, L., BLONDAL, T., KOLBEINSSON, H., STEFANSSON, J. G., HANSDOTTIR, I., RUNARSDOTTIR, V., POLA, R., LINDBLAD, B., VAN RIJ, A. M., DIEPLINGER, B., HALTMAYER, M., MAYORDOMO, J. I., KIEMENEY, L. A., MATTHIASSON, S. E., OSKARSSON, H., TYRFINGSSON, T., GUDBJARTSSON, D. F., GULCHER, J. R., JONSSON, S., THORSTEINSDOTTIR, U., KONG, A. & STEFANSSON, K. 2008. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature,* 452**,** 638-42.

TIKOLE, S. & SANKARARAMAKRISHNAN, R. 2008. Prediction of translation initiation sites in human mRNA sequences with AUG start codon in weak Kozak context: A neural network approach. *Biochem Biophys Res Commun,* 369**,** 1166-8.

TRACEY, K. J. 2002. The inflammatory reflex. *Nature,* 420**,** 853-9.

TSETLIN, V., KUZMIN, D. & KASHEVEROV, I. 2011. Assembly of nicotinic and other Cys-loop receptors. *J Neurochem,* 116**,** 734-41.

TSUNEKI, H., KOBAYASHI, S., TAKAGI, K., KAGAWA, S., TSUNODA, M., MURATA, M., MATSUOKA, T., WADA, T., KURACHI, M., KIMURA, I. & SASAOKA, T. 2007. Novel G423S mutation of human alpha7 nicotinic receptor promotes agonist-induced desensitization by a protein kinase C-dependent mechanism. *Mol Pharmacol,* 71**,** 777-86.

ULLOA, L. 2005. The vagus nerve and the nicotinic anti-inflammatory pathway. *Nat Rev Drug Discov,* 4**,** 673-84.

ULLOA, L. & TRACEY, K. J. 2005. The "cytokine profile": a code for sepsis. *Trends Mol Med,* 11**,** 56-63.

UNWIN, N. 2005. Refined structure of the nicotinic acetylcholine receptor at 4A resolution. *J Mol Biol,* 346**,** 967-89.

UTESHEV, V. V. 2012. alpha7 nicotinic ACh receptors as a ligand-gated source of Ca(2+) ions: the search for a Ca(2+) optimum. *Adv Exp Med Biol,* 740**,** 603-38.

VAN DER ZANDEN, E. P., HILBERS, F. W., VERSEIJDEN, C., VAN DEN WIJNGAARD, R. M., SKYNNER, M., LEE, K., ULLOA, L., BOECKXSTAENS, G. E. & DE JONGE, W. J. 2012. Nicotinic

acetylcholine receptor expression and susceptibility to cholinergic immunomodulation in human monocytes of smoking individuals. *Neuroimmunomodulation,* 19**,** 255-65.

VAN MAANEN, M. A., STOOF, S. P., VAN DER ZANDEN, E. P., DE JONGE, W. J., JANSSEN, R. A., FISCHER, D. F., VANDEGHINSTE, N., BRYS, R., VERVOORDELDONK, M. J. & TAK, P. P. 2009. The alpha7 nicotinic acetylcholine receptor on fibroblast-like synoviocytes and in synovial tissue from rheumatoid arthritis patients: a possible role for a key neurotransmitter in synovial inflammation. *Arthritis Rheum,* 60**,** 1272-81.

VERRALL, S. & HALL, Z. W. 1992. The N-terminal domains of acetylcholine receptor subunits contain recognition signals for the initial steps of receptor assembly. *Cell,* 68**,** 23-31.

VEZAIN, M., SAUGIER-VEBER, P., GOINA, E., TOURAINE, R., MANEL, V., TOUTAIN, A., FEHRENBACH, S., FREBOURG, T., PAGANI, F., TOSI, M. & MARTINS, A. 2010. A rare SMN2 variant in a previously unrecognized composite splicing regulatory element induces exon 7 inclusion and reduces the clinical severity of spinal muscular atrophy. *Hum Mutat,* 31**,** E1110-25.

VILLIGER, Y., SZANTO, I., JACONI, S., BLANCHET, C., BUISSON, B., KRAUSE, K. H., BERTRAND, D. & ROMAND, J. A. 2002. Expression of an alpha7 duplicate nicotinic acetylcholine receptor-related protein in human leukocytes. *J Neuroimmunol,* 126**,** 86-98.

WANG, H., BLOOM, O., ZHANG, M., VISHNUBHAKAT, J. M., OMBRELLINO, M., CHE, J., FRAZIER, A., YANG, H., IVANOVA, S., BOROVIKOVA, L., MANOGUE, K. R., FAIST, E., ABRAHAM, E., ANDERSSON, J., ANDERSSON, U., MOLINA, P. E., ABUMRAD, N. N., SAMA, A. & TRACEY, K. J. 1999. HMG-1 as a late mediator of endotoxin lethality in mice. *Science,* 285**,** 248-51.

WANG, H., LIAO, H., OCHANI, M., JUSTINIANI, M., LIN, X., YANG, L., AL-ABED, Y., METZ, C., MILLER, E. J., TRACEY, K. J. & ULLOA, L. 2004. Cholinergic agonists inhibit HMGB1 release and improve survival in experimental sepsis. *Nat Med,* 10**,** 1216-21.

WANG, H., YU, M., OCHANI, M., AMELLA, C., TANOVIC, M., SUSARLA, S., LI, J., YANG, H., ULLOA, L., AL-ABED, Y., CZURA, C. & TRACEY, K. 2003. Nicotinic acetylcholine receptor alpha7 subunit is an essential regulator of inflammation. *Nature,* 421**,** 384-8.

WANG, Y., PEREIRA, E. F., MAUS, A. D., OSTLIE, N. S., NAVANEETHAM, D., LEI, S., ALBUQUERQUE, E. X. & CONTI-FINE, B. M. 2001. Human bronchial epithelial and endothelial cells express alpha7 nicotinic acetylcholine receptors. *Mol Pharmacol,* 60**,** 1201-9.

WANG, Z. 2004. systematic identification and analysis of exonci splicing silencers. *Cell,* 119**,** 831-845.

WANG, Z., XIAO, X., VAN NOSTRAND, E. & BURGE, C. B. 2006. General and specific functions of exonic splicing silencers in splicing control. *Mol Cell,* 23**,** 61-70.

WEILAND, S., BERTRAND, D. & LEONARD, S. 2000. Neuronal nicotinic acetylcholine receptors: from the gene to the disease. *Behav Brain Res,* 113**,** 43-56.

WILK, J. B., CHEN, T. H., GOTTLIEB, D. J., WALTER, R. E., NAGLE, M. W., BRANDLER, B. J., MYERS, R. H., BORECKI, I. B., SILVERMAN, E. K., WEISS, S. T. & O'CONNOR, G. T. 2009. A genome-wide association study of pulmonary function measures in the Framingham Heart Study. *PLoS Genet,* 5, e1000429.

WILLIAMS, M. E., BURTON, B., URRUTIA, A., SHCHERBATKO, A., CHAVEZ-NORIEGA, L. E., COHEN, C. J. & AIYAR, J. 2005. Ric-3 promotes functional expression of the nicotinic acetylcholine receptor alpha7 subunit in mammalian cells. *J Biol Chem,* 280, 1257-63.

WU, J. & LUKAS, R. J. 2011. Naturally-expressed nicotinic acetylcholine receptor subtypes. *Biochem Pharmacol,* 82, 800-7.

XU, Q., MODREK, B. & LEE, C. 2002. Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic Acids Res,* 30, 3754-66.

ZHANG, C., LI, W. H., KRAINER, A. R. & ZHANG, M. Q. 2008. RNA landscape of evolution for optimal exon and intron discrimination. *Proc Natl Acad Sci U S A,* 105, 5797-802.

ZHANG, J., SUMMAH, H., ZHU, Y. G. & QU, J. M. 2011. Nicotinic acetylcholine receptor variants associated with susceptibility to chronic obstructive pulmonary disease: a meta-analysis. *Respir Res,* 12, 158.

ZHANG, X. H., KANGSAMAKSIN, T., CHAO, M. S., BANERJEE, J. K. & CHASIN, L. A. 2005. Exon inclusion is dependent on predictable exonic splicing enhancers. *Mol Cell Biol,* 25, 7323-32.

ZHANG, X. H. F. A. C., L.A. 2004. computational definition of sequence motifs governing constitutive exon splicing. *genes dev.,* 18, 1241-1250.

ZODY, M. C., GARBER, M., SHARPE, T., YOUNG, S. K., ROWEN, L., O'NEILL, K., WHITTAKER, C. A., KAMAL, M., CHANG, J. L., CUOMO, C. A., DEWAR, K., FITZGERALD, M. G., KODIRA, C. D., MADAN, A., QIN, S., YANG, X., ABBASI, N., ABOUELLEIL, A., ARACHCHI, H. M., BARADARANI, L., BIRDITT, B., BLOOM, S., BLOOM, T., BOROWSKY, M. L., BURKE, J., BUTLER, J., COOK, A., DEARELLANO, K., DECAPRIO, D., DORRIS, L., 3RD, DORS, M., EICHLER, E. E., ENGELS, R., FAHEY, J., FLEETWOOD, P., FRIEDMAN, C., GEARIN, G., HALL, J. L., HENSLEY, G., JOHNSON, E., JONES, C., KAMAT, A., KAUR, A., LOCKE, D. P., MUNSON, G., JAFFE, D. B., LUI, A., MACDONALD, P., MAUCELI, E., NAYLOR, J. W., NESBITT, R., NICOL, R., O'LEARY, S. B., RATCLIFFE, A., ROUNSLEY, S., SHE, X., SNEDDON, K. M., STEWART, S., SOUGNEZ, C., STONE, S. M., TOPHAM, K., VINCENT, D., WANG, S., ZIMMER, A. R., BIRREN, B. W., HOOD, L., LANDER, E. S. & NUSBAUM, C. 2006. Analysis of the DNA sequence and duplication history of human chromosome 15. *Nature,* 440, 671-5.