GENETIC DIVERSITY AND ADAPTATION TO ENVIRONMENTAL CHALLENGES OF ETHIOPIAN INDIGENOUS CHICKEN

Adriana Rocío Vallejo Trujillo, B.Sc., M.Sc.

A thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

June 2021

Genetics, Ecology and Evolution Group

School of Life Sciences

Declaration

I hereby declare that this thesis has not been previously presented or submitted for examination to this or any other university. The work described in this thesis is entirely of myself except where explicitly stated. References to other people's work have been duly acknowledged.

This thesis may be made available for consultation within the University of Nottingham library and can be photocopied or loaned to other libraries for consultation.

Adriana Rocío Vallejo Trujillo

June 2021

Abstract

Indigenous livestock species are characterized as being locally adapted, displaying low productivity but high levels of genetic diversity. They are usually resistant to diseases and suit smallholder farmers' needs in harsh environments; hence, they represent an important resource for achieving food security in developing countries. Understanding the genetic basis of their environmental adaptations has important implications for the design of breeding improvement programs. Likewise, it may guide conservation initiatives aiming to preserve their unique adaptive diversity. Both are equally important in the current climate change scenarios challenging agricultural production systems and threatening livestock diversity.

Under variable tropical ecologies and scavenging conditions, Ethiopian indigenous chickens exhibit unique adaptations to their environments. However, little is known about the genetic mechanisms of these environmental adaptations with, so far, no in-depth analysis of the agroclimatic adaptive stressors.

Here, an innovative approach - that integrates Ecological Niche Modelling (ENM) and genome analyses (signatures of positive selection and genomic-environmental association analyses) – is applied (i) to dissect Ethiopian chicken agro-ecologies leading to the identification of the key environmental stressors, (ii) to identify and to define Ethiopian chicken ecotypes, and (iii) to identify the genetic responses to key environmental stressors.

After a brief explanation of the research context in Chapter 1 (Setting the scene), the thesis includes four results chapters.

Chapter 2 reports the diversity and population structure of the studied populations. Using *Galgal_6.0* (*GRCg6a*) as the reference genome, 243 whole-genome sequences were analysed from Ethiopian indigenous village chickens belonging to 25 populations from diverse agro-climatic zones representing wide ecological conditions. About 19.6 million SNPs were characterized in the populations of which 29% (n = 5.6 million) were novel. While significant genetic variations are observed within the studied populations, little genetic differentiation was observed among the populations based on *Fst*, PCA, and admixture analyses. Population structure analyses show two major groups and several subgroups in the PCA plot, whereas three ancestral gene pools are revealed by admixture. The two main groups in the structure of the populations may likely reflect two temporally distinct migration paths of chicken into the African continent (Egypt and the East African coast). In comparison, the three ancestral genetic backgrounds can reflect these two migration routes and an additional third genetic background that remains unknown.

Before dissecting the genomic patterns of adaptation, the environmental characterization of the studied populations was performed in Chapter 3. By following an Ecological Niche Modelling approach, we identified six agro-climatic variables from an initial set of 34 as important predictors of the chicken habitats. These key predictors include one temperature variable (correlated to elevation), three precipitation variables (related to water availability), and two soil/land variables (linked to food availability). Twelve chicken ecotypes were proposed by characterizing the population habitats based on these six key environmental stressors.

Genomic signatures of environmental adaptation were then investigated in Chapter 4 and Chapter 5.

In Chapter 4, populations were ranked accordingly to the six key agro-climatic variables identified in Chapter 3. Signatures of positive selection for each extreme population pair (highest and lowest values for each environmental parameter) were investigated using *Fst* and

XP-EHH methods. Strong candidate selected regions identified overlapping genes that have highly relevant functions for adaptation to high-altitude stresses (e.g. hypoxia, thrombosis, and cold temperatures) (*UTP18, SLC43A3, P2RX3, CLP1, YPEL4, RTN4RL2, PGR2/3)*, high temperature (*TOGARAM1*), water scarcity (*MANEA, HTR2C, EPHA7*), and food availability under scavenging conditions (*THSD4, HBE1)*.

In Chapter 5, genome-environmental association analysis was dissected within the 12 Ethiopian chicken ecotypes identified in Chapter 3. Within-ecotypes selection signature analyses were first performed by applying the *Hp* and *iHS* methods. This was followed by a multivariate Redundancy Analysis (RDA) to identify genomic marker outliers overlapping the sweep regions. A total of 616 outlier SNPs were identified. Only variants with environmental correlation ~ 0.3 were retained for further investigation. Candidate genes included some with functions related to thermotolerance (*LDLRAD3, TNIP2, GPX7*), immune response (*PTPRZ1*), and feed metabolisms (*GPCPD1, NHLRC2, TSHD7B*).

Overall, this PhD research exemplifies the use of ENM for dissecting environmental adaptation in indigenous chicken populations. It demonstrates that ENM integrated with genome-environmental association analyses is a powerful tool for studying environmental adaptation in livestock species. The outlined methodology identified several key environmental stressors allowing subsequently to investigate the genomic responses to natural selection. Moreover, the ENM based environmental characterization of the chicken habitats represents an innovative approach for livestock ecotypes identification with potential major applications for conservation and sustainable breeding improvement.

Publications, Conferences and Awards

Publications

Gheyas, A., **Vallejo-Trujillo, A**, Kebede, A., Lozano-Jaramillo, M., Dessie, T., Hanotte, O., and Smith J. (2021). Integrated environmental and genomic analysis reveals the drivers of local adaptation in African indigenous chickens. *Mol Biol Evol*, preprint. doi:10.1093/molbev/msab156. (Parts of Chapters 3 and 4 contributed to this publication).

Banos, G., Lindsay, V., Desta, T.T., Bettridge, J., Sanchez-Molano, E., **Vallejo-Trujillo, A.**, Matika, O., Dessie, T., Wigley, P., Christley, R.M., Kaiser, P., Hanotte, O. and Psifidi, A. (2020) Integrating Genetic and Genomic Analyses of Combined Health Data Across Ecotypes to Improve Disease Resistance in Indigenous African Chickens. *Front Genet*. 11:543890. doi: 10.3389/fgene. 2020.543890.

Hume, D.A., Gutowska-Ding, MW., Garcia-Morales, C., Kebede, A., Bamidele, O., **Vallejo Trujillo, A**., Gheyas, A.A., and Smith, J. (2019). Functional evolution of the colony-stimulating factor 1 receptor (CSF1R) and its ligands in birds. *J Leukoc Biol*, 107: 237 – 250. doi: 10.1002/JBL.6MA0519-172R.

Selected Conferences

50th Population Genetics Group Meeting. January 5 – 7, 2016 – Churchill College, University of Cambridge, UK. (Attended).

36th International Society of Animal Genetics (ISAG) Conference. July 16 – 21, 2017 – University College Dublin, Ireland. (Attended).

II Joint Congress on Evolutionary Biology. August 19 – 22, 2018 – Corum, Montpellier – France – Poster presentation: Microgeographic and climatic adaptation of Ethiopian indigenous chicken: A genome-wide analysis.

School of Life Sciences Postgraduate Research Symposium, June 24 – 15, 2019 - University of Nottingham, UK. Oral presentation: Integrating Ecological Niche Modelling with genomics to dissect tropical adaptation in Ethiopian indigenous chicken.

37th International Society of Animal Genetics (ISAG) Conference. July 7 – 12, 2019. University of Lleida, Spain. Poster and oral presentation: Integrating Ecological Niche Modelling with genomics to dissect tropical adaptation in Ethiopian indigenous chicken.

53rd Population Genetics Group Meeting. January 5 - 7, 2020 - University of Leicester, UK. Oral presentation: Unravelling the local environmental adaptation of Ethiopian indigenous chicken by integrating Ecological Niche Modelling with genomic analyses.

Awards

Vice-Chancellor Scholarship for Research Excellence, UoN	2016
COLCIENCIAS Grant (Administrative Department of Science, Technology and Innovation) Colombian Government	2016

Dedication

In the loving memory of my darling Simon Mark Kerry.

Acknowledgments

First and foremost, I want to express my deepest appreciation to my supervisor Professor Olivier Hanotte for his guidance and support during this research journey. His scientific expertise, knowledge, human qualities, and unwavering enthusiasm for genetics have been a source of inspiration. My experience as his student has encouraged me to reach my goals, build more confidence, and develop critical thinking for my research. It is an honour for me to have shared the last four years of discovery by his mentoring.

My appreciation also extends to my external supervisor Dr Almas Gheyas. I am extremely grateful for her assistance and suggestions throughout my research. Her patience and excellent teaching skills allowed me to grow as a scientist. Almas, you make the most complicated things look so simple and clear. Thank you very much. Almas and Olivier, thanks for believing in me and being there for me, especially in the hardest times.

Special thanks to the Administrative Department of Science, Technology, and Innovation (COLCIENCIAS) from the Colombian Government for sponsoring my PhD studies in the UK. Also, I give thanks to the University of Nottingham, who funded my tuition fees in full through the Vice Chancellor's scholarship (International). Thanks to all of the team involved in the African Chicken Genetic Gains (ACGG) programme, the Roslin Institute, and ILRI in Ethiopia. In particular to all those who were involved in sampling, DNA extraction, and sequencing for this project. Special thanks to Dr Adebabay Kebede, Dr Jaqueline Smith, and Dr Maria Lozano-Jaramillo for your contributions to this work. Many thanks to all of the farmers in Ethiopia that contributed with chicken samples. Thank you for the incredible job you do in conserving these genetic resources and contributing to food security. I would also like to acknowledge the support of the School of Life Sciences staff, UoN, during my PhD.

A special mention for Dr Raman Lawal for supporting me during the first years of my studies. Thanks for your friendship and for giving me the confidence and assistance I needed to develop my skills in bioinformatics. A big thanks to my past and current colleagues in the lab, Dr Abdulfatai Tijani, Dr Abulgasim Ahbara, Dr Akil Alshawi, Dr Ahmed Al-Jumali, Dr Ibrahim Akinci, and Mifta Rahman (soon to be Dr Mifta Rahman). You guys are incredible human beings and made this whole experience more enjoyable.

To my two main motivations, my sons Santiago and Diego. I cannot thank you both enough for believing in this. You have been very brave throughout this journey. This success is also yours. Also to my dearest Paul for his loving support and to Lily, and Lucas for joining us in this adventure. I appreciate all of your patience, understanding, and love, thanks. To my parents Gustavo and Florelia, this could not have happened without your constant support, guidance, and example. You have worked so hard to fill my life with opportunities. I hope this success can honour all of your sacrifices. To my brother, sister, and all of my family back in Colombia, you are part of this achievement too. A big thank you to all of my friends in the UK for your unconditional support. Finally, to Simon, I know you are still looking after us and always will.

Table of Contents

List of Figures

List of Tables

Abbreviations

ACGG	African Chicken Genetic Gains
AEZ	Agro-ecological Zones
AIC	Akaike Information Criterion
AUC	Area Under the Receiving Operator Curve
bp	base pair
BQSR	Base Quality Score Recalibration
BWA	Burrows-Wheeler Aligner
CRISPR	Clustered regularly interspaced short palindromic repeats
DAGRIS	Domestic Animal Genetic Resources Information System
dbSNP	Single Nucleotide Polymorphism Database
EIAR	Ethiopian Institute of Agricultural Research
ENM	Ecological Niche Modelling
EVA	European Variation Archive
FAO	Food and Agriculture Organisation
FC	Feature classes
Fst	Fixation index
GATK	Genome Analysis Toolkits
GEA	Genotype-environment association
GFM	Genomic Feature Models
Hp	Pooled Heterozygosity
iHS	Integrated Haplotype Score
ILRI	International Livestock Research Institute
INDELS	Insertions / Deletions
kb	Kilobase pair
KEGG	Kyoto Encyclopaedia of Genes and Genomes
LandUse	Percentage of cultivate land
LD	Linkage Disequilibrium

LRH	Long-range haplotype
MAF	Minor Allele Frequency
MAS	Marker Assisted Selection
MHC	Major Histocompatibility Complex
minTemp	Minimum temperature of the coldest month of a year
NOR	Nuclear Organising Region
PCA	Principal Component Analysis
precSeasonality	Precipitation Seasonality of a year
precDQ	Precipitation in the Driest Quarter of a year
precWQ	Precipitation in the Wettest Quarter of a year
QTL	Quantitative trait loci
RDA	Redundancy Analyses
RM	Regularization multiplier
SD	Standard deviation
SDM	Species Distribution Modelling
SMRT	Single-Molecule Real-Time
SNNPR	Southern Nations, Nationalities, and People’s Region
SNP	Single Nucleotide Polymorphism
SoilOrgC	Soil Organic Carbon content
SV	Structural variants
UN	United Nations
VEP	Variant Effect Predictor
VQSR	Variant Quality Score Recalibration
WFP	World Food Programme
WGS	Whole Genome Sequencing
XP-EHH	Cross Population Extended Haplotype Homozygosity

Chapter 1

General introduction

This thesis presents the research results aiming to develop a novel approach for the identification, characterisation, and study of tropical adaptation in African poultry ecotypes. The outcomes of the research are expected to contribute to the African Chicken Genetic Gains (ACGG) programme (https://www.ilri.org/research/projects/african-chicken-genetic-gains-acgg) under the auspices of the International Livestock Research Institute (ILRI). In particular, we aim to address the current challenges of improving the village poultry productivity of local farming systems and improve the nutritional and financial security of the smallholder poultry keepers in Ethiopia by understanding the environmental constraints of scavenging systems and the genome responses against such challenges.

1.1 Context of the study

Human population growth has led to increased demand for food. Today, two billion people, or 25.9% of the global population, do not have regular access to sufficient food and are experiencing hunger (FAO, 2020a). The situation is particularly acute in Eastern Africa, where smallholder farmers are paying a heavy toll due to climatic fluctuation, including the effects of climatic changes and natural disasters (e.g. desert locust outbreaks) (WFP, 2020).

With around 112 million people, Ethiopia is the 12th most populated country in the world (UN, 2019). Roughly 80% of the Ethiopian population relies on agriculture and livestock for their livelihoods. It is a country where unforeseen climate disasters often leave many people exposed to food insecurity (FAO, 2020b). Livestock production is a crucial element of the agricultural production system in Ethiopia. The farm system is divided into smallholder mixed farming in the highlands and pastoralism in the lowlands (Kassahun & Poulton, 2014).

The poultry population plays a crucial role in the country's economy and is considered a valuable resource for improving the livelihoods of the growing Ethiopian human population (Dessie, 2011a). The Ethiopian poultry population comprises approximately 60 million chickens, 90.8% of which is represented by native breeds, 4.4% by hybrids, and 4.8% by exotic breeds (FAO, 2019). Indigenous village chickens serve as an important source of protein supply for human consumption, even though very little capital is invested in their farming. Egg sales enable poultry keepers to obtain money for their family needs, making poultry production a poverty alleviation tool (Wilson, 2010). Besides, from cultural and religious perspectives, poultry represents a particularly useful source of animal protein. It is exempt from any religious taboo, except during fasting time for the Ethiopian orthodox community, and poultry products are consumed by all ethnic groups in different social strata (Wilson, 2010). Also, poultry production is an important strategy to promote gender equality in rural areas where women often are the custodians of backyard-reared chickens and benefit from the sale of the birds and eggs (Gheye, 2000). Poultry farming in Ethiopia (and in Africa in general) is largely extensive. It is represented predominantly by a backyard scavenging or semi-scavenging production system, consisting of small chicken flocks of typically 7 to 10 mature birds per household (Pagani & Abebe, 2008; Alemu, 2008; Wilson, 2010). Indigenous chickens, reared in such smallholder farming practices, show high phenotypic diversity, including plumage colour and pattern, comb type, weight-body conformation, and eggshell size (Wilson, 2010) but often suffer from low productivity, low hatchability, and high chick mortality through predation (Manyelo *et al.*, 2020). However, indigenous chickens have special adaptive features that make them more resistant and suitable for production purposes in harsh and variable agro-ecological environments than most exotic commercial breeds (Pagani & Abebe, 2008; Psfidi *et al.*, 2016; Muchadeyi & Dzomba, 2017).

The Ethiopian landscape includes about 50% of the African mountains, causing large variation in elevations, ranging from 126 m below sea level in the Danakil depression to 4533 m above sea level in the Semien Mountains (Asefa *et al.*, 2020). The combination of topography, soils, temperature, and rainfall in Ethiopia represents eleven Afro-tropical eco-regions. Therefore, the country has been designated a hotspot of global importance for biodiversity conservation (Olson & Dinerstein, 2002).

Indigenous chickens that show adaptation to their agro-climatic conditions are often referred to as ecotypes. However, wide variations in landscape patterns and large diversity in chicken phenotypes observed in Ethiopia represent a challenge in characterising chicken populations and defining ecotypes associated with specific agro-ecologies. Also, little is known about how the adaptive pressure of different environmental variables has shaped the genetic characteristics of these populations or potential ecotypes. Likewise, little effort has been made to characterise the ecosystems to identify the main environmental selection factors (Muchadeyi & Dzomba, 2017). Previous studies in poultry scrutinise a particular environmental stressor (e.g. elevation or heat stress) (Zhang *et al.*, 2016a; Cedraz *et al.*, 2017) or specific ecotypes (generally defined based on specific geographic regions or bearing a distinctive morphological feature) (Walugembe *et al.*, 2019).

1.2 Rationale and objectives for the study

Insights of how evolutionary forces shape the genetic diversity and adaptation of livestock species will provide useful information for conservation, management, and genetic improvement (Mignon-Grasteau *et al.,* 2005; Boettcher *et al.,* 2015). Many efforts have been made in the genetic characterisation of farm animals (Groeneveld *et al.,* 2010; Eynard, 2018; van Marle-Köster & Visser, 2018). However, indigenous chickens living in diverse agro-climatic conditions have been up to now poorly characterised. This study aims to identify and to characterise the environmental adaptation of Ethiopian indigenous chicken at genome level. A novel approach for defining ecotypes by integrating Ecological Niche Modelling (ENM) with genomic analyses is proposed. It includes a detailed characterisation of the agro-climatic conditions and the detection of signatures of positive selection.

The principal objectives of this study are as follows:

- To analyse the genetic diversity and population structure of different Ethiopian indigenous chicken populations (Chapter 2).
- To perform an environmental characterisation of the agro-ecologies where Ethiopian indigenous chickens survive and identify the major agro-climatic variables playing an important role as potential ecological stressors or adaptation drivers (part of Chapter 3).
- To define chicken "ecotypes" based on environmental characterisation of the agro-ecologies of population habitats (part of Chapter 3).
- To identify the genomic regions under strong positive selection in response to specific environmental stressors across and within the defined ecotypes (results split in Chapters 4 and 5).

Chapter 2

Genetic diversity and population structure of indigenous Ethiopian village chicken

Abstract

Local chicken populations used for backyard smallholder farming in Ethiopia are adapted to the diverse agro-climatic conditions observed in the country. They show a higher tolerance for biotic and abiotic stressors than improved commercial breeds. Understanding the genomic diversity and population structure of such indigenous populations is the first step in exploring the adaptive features that have allowed them to survive in diverse agro-climatic ecologies. To this effect, the present study analyses whole-genome sequence data from 243 indigenous chicken samples from 25 different populations across diverse Ethiopian landscapes. Sequence reads were mapped against the latest chicken reference genome (*GRCg6a)* to identify Single Nucleotide Polymorphism (SNP) variants. Different measures of genetic diversity were explored, including the proportion of novel variants, heterozygosity estimation, inbreeding levels, population structure based on Principal Component Analysis (PCA) and Admixture analysis, population differentiation based on *Fst*, and genome-wide LD structure. A total of 19.6 million SNPs were identified with 10 to 11.6 million SNPs per population and observed mean population heterozygosity ranging between 0.26 to 0.33. Except for two main groups (populations from the eastern part of the country), only a weak sub-structuring of the populations is observed. Even though admixture analysis suggested contributions from three major ancestral populations, low levels of pairwise population differentiations were observed with *Fst* values ranging between 0 to 0.12. Average LD r^2 values range between 0.2 – 0.3 across population groups, consistent with previous results on crossbreed chicken populations. These results indicate a large genetic variability but weak population structure among the Ethiopian indigenous chickens examined.

2.1 Introduction

Ethiopia ranks first in Africa and tenth globally in terms of inventory of livestock populations (Getu, 2014). This includes a chicken population of around 60M, of which most are represented by indigenous breeds (FAO, 2019). Backyard farming predominantly depends on indigenous breeds due to their adaptability to local farming systems (e.g. scavenging conditions) and harsh environments. However, the Ethiopian agricultural landscape shows varied climatic conditions such as hot-arid, hot-humid, cold-arid, and cold-humid agro-ecologies, represented by 17 agro-climatic zones (Hurni, 1998). These wide diversity patterns in phenotypes and environments represent a challenge for classifying Ethiopian indigenous chickens as breeds or ecotypes. The Domestic Animal Genetic Resources Information System (DAGRIS) currently includes information on ten varieties/ecotypes of chickens from Ethiopia (Tepi, Kei, Tilili, Naked neck, Horro, Netch, Chefe, Tikur, Jarso, and Gesima) (DAGRIS, 2007).

Several studies have now examined the phenotypic and genetic diversity of Ethiopian indigenous village chickens. For example, Dessie *et al.* (2011b), and Desta *et al*. (2013), reported the phenotypic diversity of Ethiopian chickens from Horro (Western Ethiopia) and Jarso (Eastern Ethiopia) regions, which represent different agricultural farming systems and ecological landscapes. They showed that chickens display various phenotypes within populations with differences in plumage colour, pattern, size, comb shape, and skin colour (Dessie *et al.*, 2011b; Desta *et al.*, 2013). Mitochondrial DNA analyses have shown the predominant presence of two main haplogroups of sequences (Mwacharo *et al.*, 2011; Al-Jumaili *et al.*, 2020). One haplogroup, likely the more ancient one, originated primarily from the Indian subcontinent and probably reached Africa via Egypt following a terrestrial route. The second haplogroup is likely the legacy of the Indian Ocean maritime trading routes along the East African coast, originally coming from as far Southeast and East Asia (Muchadeyi *et al.,* 2008; Mwacharo *et al.,* 2011; Lyimo *et al.,* 2013; Prendergast *et al.,* 2017; Al-Jumaili *et al.,* 2020). Studies examining the microsatellite and nucleotide genetic diversity of East African chicken have reported at least two (Goraga *et al.,* 2011) and possibly three (Mwacharo *et al.,* 2013a) genetic backgrounds across village chicken. Only a single study (Lawal *et al.* 2018*)* has examined the genetic diversity of Ethiopian chickens using full genome sequencing. This study, however, included only two Ethiopian chicken populations - Horro and Jarso. Approximately 17 million SNPs, corresponding to a density of ~16 SNPs/kb and ~13% novel variants, were identified. Average observed heterozygosity (0.31 – 0.05), the nucleotide diversity ($\pi = 0.003 - 0.004$), and *Fst* (0.11 – 0.14) suggested high genetic diversity but weak differentiation among these chicken populations (Lawal *et al.,* 2018).

Constant improvements in high throughput sequencing technologies reduce the cost of whole-genome sequencing (WGS), thereby allowing the scrutiny of the genetic diversity at a genome-wide level by sequencing many samples. This information can then be used to design breeding programs aiming to enhance productivity and performance (e.g. disease resistance, climatic resilience) while conserving important genetic diversity sustainably (e.g. preserving local ecotypes with unique environmental adaptations) (Bettridge *et al.,* 2018).

WGS analyses of chickens have also been supported by continuous development of the chicken genome reference assembly, the first avian and livestock species whose genome sequence was assembled and has had several improvements (Schmid *et al.,* 2015). The latest reference build, *GRCg6a* (*galGal6),* released in March 2018, represents a significant improvement compared to the previous build, *galGal5*. It was generated based on the single-molecule real-time (SMRT) sequencing technology of Pacific

Biosciences with a sequencing depth of approximately 80x (GRCg6a, 2018). The current reference has a total sequence length of around 1.07 Gigabase, containing 16,878 coding genes, 39,288 transcripts and 7,166 non-coding genes (https://www.ensembl.org/Gallus_gallus/Info/Annotation). In comparison, *galGal5* had a full size of 1.23 Gigabases with 18,346 coding genes, 38,118 transcripts, and 6,492 non-coding genes. The use of long-read sequences not only refined the *galGal6* genome assembly but also led to the improvements to *de novo* assembly algorithms, manual annotation of the assembled contigs, and integration of finished BAC clone sequences (https://www.ncbi.nlm.nih.gov/grc/chicken). These upgrades are reflected in the differences in size and number of annotation categories between *galGal5* and *galGal6*. The *galGal5* assembly includes a total of 23,869 scaffolds with 397 genome gaps and 24,692 contigs. Contrarily, *galGal6* contains only 524 scaffolds with 68 genome gaps and 1,402 contigs, including 34 of the 39 chromosome pairs of the chicken genome (https://www.ncbi.nlm.nih.gov/assembly/GCF_000002315.6). Also, chromosome 16, which contains the Nuclear Organising Region (NOR) and the Major Histocompatibility Complex (MHC), has major improvements in *galGal6* assembly. Its total length in *galGal6* is 2.8 megabases, and the scaffold-N50 length is 1.5 megabases. The corresponding lengths in *galGal5* were only 0.6 megabases and 145 kilobases, respectively (https://www.ncbi.nlm.nih.gov/assembly/GCF_000002315.6#/st). Although there are still some issues regarding the latest *galGal6* chicken genome (e.g. assembly of several microchromosomes still missing) (https://www.ncbi.nlm.nih.gov/grc/chicken/issues), this version is much more informative. It offers a new valuable resource for chicken genomic research.

This chapter reports the genetic diversity of Ethiopian indigenous chickens based on single nucleotide polymorphisms (SNPs) detected from genome sequence data on 243 samples from 25 different populations and using the *galGal6* reference build for mapping and variant calling. It also examines the population structure and differentiation among these populations. Together, these pieces of information provide updated baseline knowledge for the subsequent signature of selection analyses conducted in Chapters 4 and 5, aiming to understand the genomic adaptation of Ethiopian chickens to their environments.

2.2 Materials and Methods

2.2.1 Sampling and sequencing

Sampling was carried out as part of the African Chicken Genetic Gains project (https://africacgg.net/) in collaboration with the Centre for Tropical Livestock Genetics and Health (www.ctlgh.org) and the Ethiopian Institute of Agricultural Research (EIAR) (Kebede, 2018). Two hundred twenty-three adult (over six months old) chickens from 23 villages across 12 districts and six different national regions (Afar, Amhara, Gumuz, Oromia, SNNPR – Southern Nations, Nationalities, and People's Region, and Tigray) were sampled. From any household, a maximum of two birds (one male and one female) were sampled. Blood was collected from the jugular for DNA extraction (see Kebede (2018) for full details).

Twenty additional samples from Horro and Jarso populations from a previous study (Lawal *et al.*, 2018; Bettridge *et al.*, 2018) were also included, bringing the total number of samples to 243 and the number of villages sampled to 25. The sampled populations represent different Ethiopian climatic and agro-ecological zones (Table 2.1 and Figure 2.1). Whole-genome sequencing of individual samples was performed on the HiSeq 200 Illumina platform with an average coverage of ~45X.

Table 2.1 Analysed samples across twenty-five Ethiopian populations

Kebele[a]	N[b]	Longitude	Latitude	Elevation	Mean rainfall (mm)	Mean Temperature (°C)	AEZ-16[c]	Traditional classification[d]	Geographic region
Adane	10	36.894902	11.074785	2280 - 2734	900 - 1000	14 - 18	cool / sub-humid	Moist dega	Amhara
AlfaMidir	10	36.865141	11.084692	3304 – 3451	> 1000	10 - 14	cold / sub-humid	Moist wurch	Amhara
Amesha Shinkuri	10	37.083	10.986	2417 - 2564	900 - 1000	14 - 18	cool / sub-humid	Moist dega	Amhara
Arabo	10	36.906	10.661	1471 - 1624	900 - 1000	18 - 20	cool / semiarid	Moist - weynadega	Amhara
Ashuda	10	36.845155	11.395002	2009 - 2103	900 - 1000	18 - 20	cool / sub-humid	Moist - weynadega	Amhara
Batambie	8	36.892	11.489	2003 - 2127	900 - 1000	18 - 20	cool / sub-humid	Moist - weynadega	Amhara
Bekele Girissa	10	36.223	10.868	1642 - 1668	900 - 1000	18 - 20	cool / sub-humid	Moist - weynadega	Oromya
Dikuli	10	36.165094	10.714984	2022 - 2109	900 - 1000	18 - 20	cool / sub-humid	Moist - weynadega	Amhara
Gafera	10	37.555	12.372	2441 - 2592	900 - 1000	14 - 18	cool / sub-humid	Moist - dega	Amhara
Gesses	9	38.834809	14.364728	1196 - 1546	< 900	18 -24	warm / semiarid	Dry - kolla	Amhara
Gijet	9	38.774904	14.464562	1990 - 2303	900 - 1000	18 - 20	cool / semiarid	Moist - weynadega	Amhara
Hadush Adi	9	39.164709	13.334987	1417 - 1669	< 900	18 -24	cool / semiarid	Dry - kolla	Tigray
Hugub	10	39.174911	13.244464	737 - 979	< 900	18 -24	warm / semiarid	Dry - kolla	Afar
Jarso	14	39.565015	13.365625	1817 - 2136	900 - 1000	18 - 20	cool / semiarid	Moist - weynadega	Oromya
Kido	9	39.834715	11.254644	1281 - 1418	< 900	18 -24	cool / arid	Dry - kolla	Amhara
Kumato	10	39.894959	11.164452	1838 - 1969	900 - 1000	18 - 20	cool / sub-humid	Moist - weynadega	Oromya
LocalHorro	6	39.653	10.321	1921 - 2375	900 - 1000	18 - 20	cool / sub-humid	Moist - weynadega	Oromya
Loya	10	39.778	10.332	1601 - 1847	900 - 1000	18 - 20	cool / sub-humid	Moist - weynadega	Oromya
Meseret	10	40.08476	9.404273	2248 - 2312	900 - 1000	18 - 20	cool / semiarid	Moist weynadega	Tigray
Metkilimat	10	42.164926	9.444748	1709 - 2051	900 - 1000	18 - 20	cool / semiarid	Moist - weynadega	Tigray
Mihquan	10	38.85439	8.154678	1295 - 1495	< 900	18 -24	cool / semiarid	Dry - kolla	Tigray
Negasi Amba	10	38.795003	8.145067	2978 - 3071	900 - 1000	14 - 18	cool / sub-humid	Moist - dega	Amhara
ShubiGemo	10	38.394908	6.514992	1643 - 1651	900 - 1000	18 - 20	cool / sub-humid	Moist - weynadega	Oromya
Surta	9	38.314788	6.485064	2480 - 2575	900 - 1000	14 - 18	cool / sub-humid	Moist - dega	Amhara
TzionTeguaz	10	37.01513	9.934119	1870 - 1968	900 - 1000	18 - 20	cool / semiarid	Moist - weynadega	Amhara

(a) Village or population. Kebele is the smallest administrative unit in Ethiopia.
(b) Number of samples per population.
(c) Agro-ecological regions are defined based on the Global 16 Class classification system (Harvest Choice, 2015).
(d) Source: MoA, 2000.

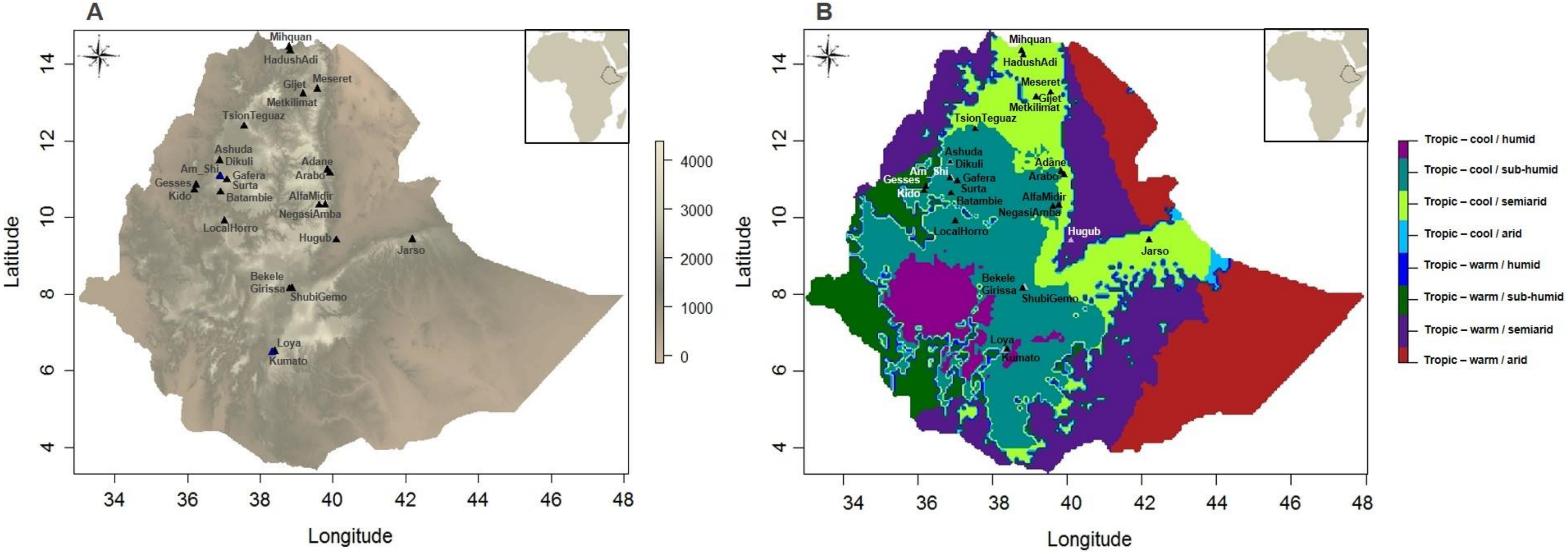


Figure 2.1 Geographic location across Ethiopia of the twenty-five investigated indigenous chicken populations.

(A) Population locations against a physical map based on elevation (meter above sea level, m.a.s.l.) and (B) population locations against Ethiopian main agro-ecological zones (Harvest Choice, 2015).

2.2.2 Sequence alignment, variant calling and annotation

The sequencing reads (FASTQ format) were mapped to the reference genome GRCg6a (galGal6) using the Burrows-Wheeler Aligner (BWA-MEM) version 0.7.15 (Li & Durbin, 2009). The mapped reads (in SAM format) were further processed to remove PCR duplicates using Picard tool v2.9.0 (http://picard.sourceforge.net). Using GATK v3.8.0 (McKenna et al., 2010), a base quality score recalibration (BQSR) was performed. Variant calling was completed using the 'haplotype caller" function in GATK, followed by joint genotyping of all samples. The "joint genotyping" option was applied in order to "rescue" genotype calls at sites where a carrier has a low coverage but other samples withing the call set have a confident variant at that location (https://software.broadinstitute.org/gatk/documentation/article.php?id=4150).

A variant quality score recalibration (VQSR) was applied to minimize calling artifacts and increase the discovery of real variants. This technique uses machine learning to model the variant profile based on training sets to discover false positives according to GATK best practices recommendations (DePristo *et al.*, 2011; Van der Auwera *et al.*, 2013). Two lists of known variants were used for the VQSR step: (i) a 1 million validated SNPs (Kranis *et al.*, 2013) used as the "Training" or "Truth" set, and (ii) a set of 14 million SNPs from dbSNP (https://www.ncbi.nlm.nih.gov/snp/) as "known" set. A tranch level of 99% was used for the filtration purpose, corresponding to 99% sensitivity relative to the truth set. Only autosomal SNPs that passed the recalibration threshold were selected for downstream analyses. Further filtrations were then applied on the VQSR passed autosomal variants for the downstream genomic analyses. These included selecting only biallelic SNPs from autosomes with a missing genotype rate ≤ 20%, genotype quality ≥ 15, and depth of coverage ≥ 3. Using Ensembl Variant Effect

Predictor (VEP) platform (McLaren *et al.,* 2016), the variant's consequences were identified from Ensembl (release 98).

2.2.3 Genetic diversity and population structure

Several approaches were applied to explore the genetic diversity and population structure of the studied population. To assess the within and between population genetic diversity, VCFtools version 0.1.15 (Danecek *et al.*, 2011) was used to calculate the nucleotide diversity (π) (with the --window-pi 20000 option), inbreeding coefficient (*F*), the pairwise population differentiation (*Fst*), and heterozygosity (--het option). Heterozygosity calculation was based on the proportion of heterozygous markers in each individual and then obtaining mean values for each population. To assess the underlying population structure, Principal Component Analysis (PCA) was performed using the Eigenstrat method in Eigensoft v6.1.4 software (Patterson *et al*., 2006; Price *et al*., 2006). Admixture programme v1.3.0 (Alexander *et al.,* 2009) was executed to identify ancestry structure on the dataset. Cross-validation for different *K* values (i.e. number of potential ancestral gene pools) from 1 to 5 was carried out following Evanno *et al*. (2005) method.

Linkage Disequilibrium (LD) analyses were performed with a reduced set of genome-wide variants rather than the full SNP collection to reduce computational burden. The reduction of the variant set was performed using the "thinning" approach in PLINK v1.9 (Purcell, 2007), keeping 30% (~ 4 M variants) of randomly selected SNPs from the original full set.

LD and haplotype block structures were measured in windows of 20 kb size using Plink software *v.1.07* (Purcell, 2007). For these analyses, genetically close populations (based on geographic proximity) were combined to increase the sample size. In total, six population groups were defined with sample sizes between 10 to 90 individuals. Genome-wide pairwise r^2 metrics for SNPs within 20 kb of distance were plotted with R. The LD decay was calculated and plotted using the PopLDdecay software (Zhang *et al.,* 2019). Using the –block parameter, haplotype blocks were estimated *via* Haploview's interpretation of the block definition suggested by Gabriel *et al.* (2002). This definition denotes blocks as genome regions with a limited haplotype diversity without much internal recombination and separated by segments that show little or no LD.

2.3 Results

2.3.1 Sequencing, variant discovery and annotation

The average number of paired sequence reads per population ranged between 202 to 475 million with a mean sequence genome coverage between 22 to 44 X (Supplementary Table S2.1). Compared with the study of Kebede *et al.* (2018), where *galGal5* was used as a reference genome, we obtained a similar mapping rate (98 – 99 %) (Supplementary Table S2.1). About 19.6 million high-quality SNPs were detected (after VQSR filtration) from whole-genome sequence analyses of the 25 Ethiopian populations. The variants were distributed in the genome with an average density of 19 SNPs per kb or 1 SNP per every 52 bases. After applying the extra filtration parameters (see Materials and Methods), over 14 million SNPs (n=14,091,055) were retained.

The number of SNPs across populations was similar, varying between 10 to 11.6 million, with the lowest number detected in Hugub and the highest number in Kumato (Table 2.2). Between 10% to 11% of the SNP are novel in each population. When all Ethiopian chicken populations are lumped together, 29% of the entire SNPs set (5.6 million out of 19.6) are novel.

As a measure of quality control, Ts/Tv was calculated. We observe a Ts/Tv ratio of around 2.4 - 2.5 across populations and 2.4 when all populations are analyzed as a whole (Table 2.2). DePristo *et al.* (2011) state that Ts/Tv ratio should be about 2.1 for whole-genome sequencing and can be used to validate SNP discovery. If it is lower, the data might include false positives caused by random sequencing errors. Previous studies in chicken have found similar Ts/Tv ratios to ours, e.g. 2.53 (Derks *et al.*, 2018) and 2.31 (Boschiero *et al.*,2018).

Table 2.2 SNPs summary statistics for the studied chicken population.

Population	N	nSNPs	Novel (%)	Ts/Tv	Singletons (%)
Adane	10	11,231,204	11.11	2.48	21.44
Alfa Midir	10	10,539,558	11.38	2.49	22.41
Amesha Shinkuri	10	11,211,292	11.39	2.48	18.59
Arabo	10	10,958,190	11.23	2.48	21.92
Ashuda	10	10,804,200	11.5	2.48	18.04
Batambie	8	10,682,327	10.68	2.48	20.61
Bekele Girissa	10	11,310,581	11.01	2.48	19.34
Dikuli	10	11,265,455	11.15	2.48	19.44
Gafera	10	11,135,900	11.44	2.48	18.77
Gesses	9	10,337,743	11.42	2.48	20.82
Gijet	9	11,046,899	10.79	2.48	19.06
Hadush Adi	9	10,213,712	10.08	2.49	20.37
Hugub	10	10,034,153	10.69	2.48	14.43
Jarso	14	11,155,696	10.63	2.49	17.92
Kido	9	10,543,102	11.09	2.48	18.96
Kumato	10	11,606,289	11.46	2.48	18.19
Local Horro	6	10,170,097	10.18	2.48	26.36
Loya	10	11,078,600	11.12	2.48	19.71
Meseret	10	11,419,999	10.92	2.48	18.29
Metkilimat	10	11,295,871	10.96	2.48	18.63
Mihquan	10	11,428,789	10.83	2.48	17.84
Negasi Amba	10	10,535,898	11.21	2.48	20.63
Shubi Gemo	10	11,217,239	10.82	2.48	17.04
Surta	9	10,984,861	10.42	2.48	20.73
Tsion Teguaz	10	10,954,331	11.16	2.48	19.17
Overall value (All populations)	243	19,571,070	29.02	2.48	0.09

N = number of samples analyzed; nSNPs = total number of SNPs retained after filtration (see Material and Methods); Novel SNPs (%) = non-reported SNPs in dbSNP; Ts/Tv = transition/transversion ratio; Singletons (%) = singleton variant identified

The proportion of singleton SNPs from different populations varied from 14.43% in Hugub to 26.36% in local Horro (Table 2.2). Differences might be a function of the number of studied samples in each population and how genetically diverse they are. These large proportions of singletons reflect the small sample size (6-10 individuals) analysed per population. Therefore, when the entire set of Ethiopian samples are considered, the proportion of singletons reduces drastically to only around 0.09%.

The overall SNP density observed in different chromosomes is presented in Figure 2.2. The lowest SNP density is observed in chromosomes 31 and 16, and the highest in chromosome 26, followed by chromosome 6. A chromosome-wise heatmap of SNP density in 10 kb windows across the genome is presented in Figure 2.3. A range between 100 to 200 SNPs is more commonly found every 10 kb, although there are also many SNP void regions (red areas). Chromosomes 16, 22, 30, 31, 32, and 33 (with many void regions), display a smaller SNP density (< 50 SNPs for 10 kb window) compared to other chromosomes. This may be caused by low coverage due to sequencing difficulties or mapping issues due to high repeat content (Li *et al.*, 2011a).

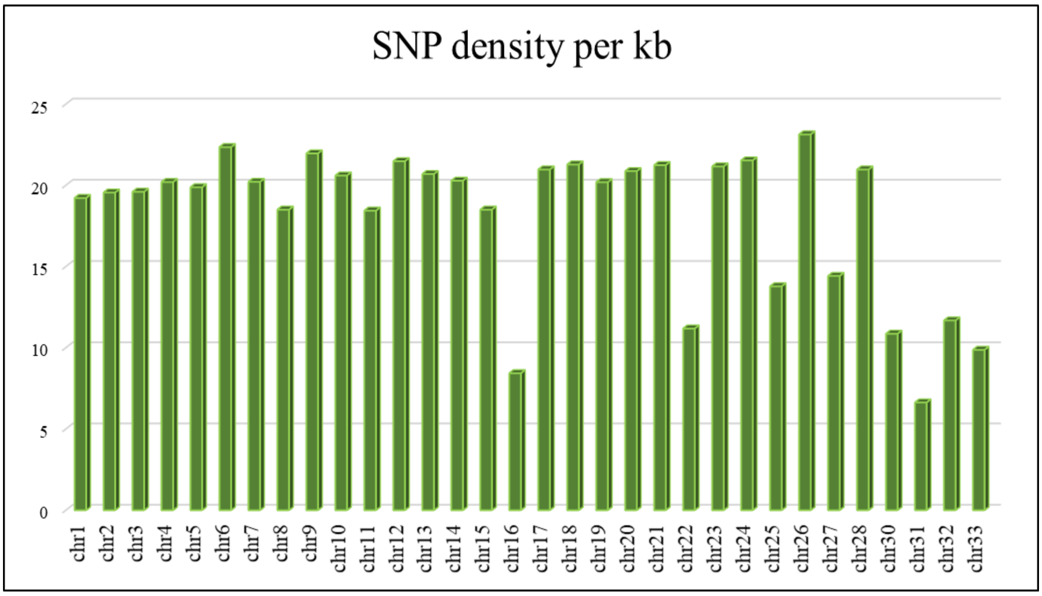


Figure 2.2 Mean SNPs density (SNP count/kb) per chromosome in Ethiopian indigenous chicken.
(Based on 19.6 million SNPs detected from all 25 populations).

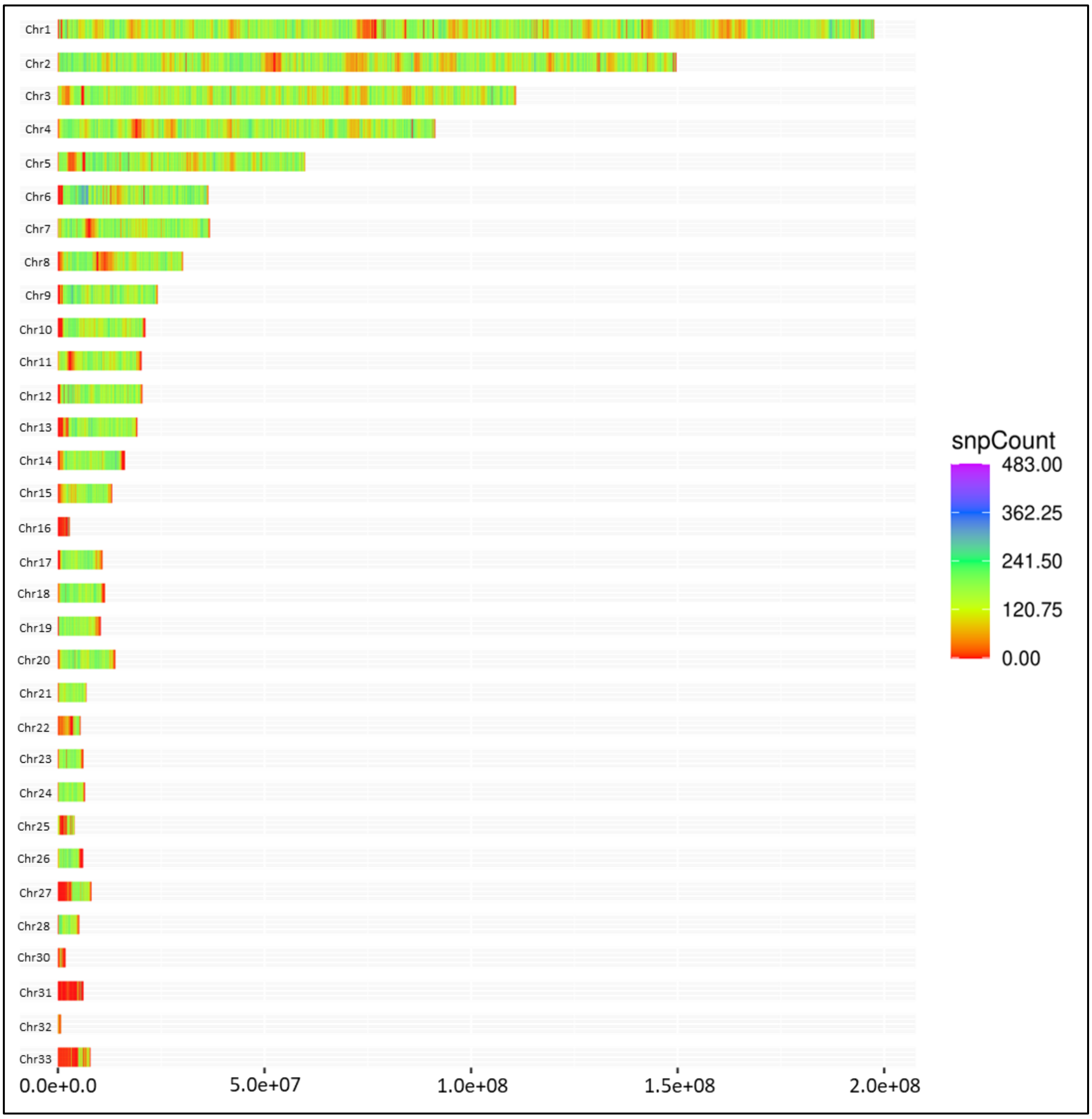


Figure 2.3 Chromosome-wise SNP distribution heatmap across the Ethiopian indigenous chicken genome based on 19.6 million SNPs.

(X-axis denotes the chromosome size in base pairs (bp), and Y-axis the chromosome number). The SNP count was calculated for a 10 kb window size.

Among the 19.6 M SNPs, intergenic (37.8%) and intronic (45.5%) variants are the most common (Figure 2.4). Exonic variants represent 1.5% of the total detected SNPs. Among the exonic SNPs, 73% are synonymous, 26% non-synonymous, 0.3% stop gain and 0.03% stop loss. The majority of the SNPs (between 34% to 0.2%) from all annotation categories belong to low-frequency category of 0 - 0.2. Only a small proportion of the SNPs (between 3.5% to 0.02%) have a high frequency of 0.8 – 1 (i.e.

near fixation). Of these high-frequency SNPs, a total number of 4,771 non-synonymous, 28 stop-gain, and 12 stop-loss SNPs are identified. These categories potentially have a functional impact, and high frequency may suggest they may have some beneficial adaptive roles.

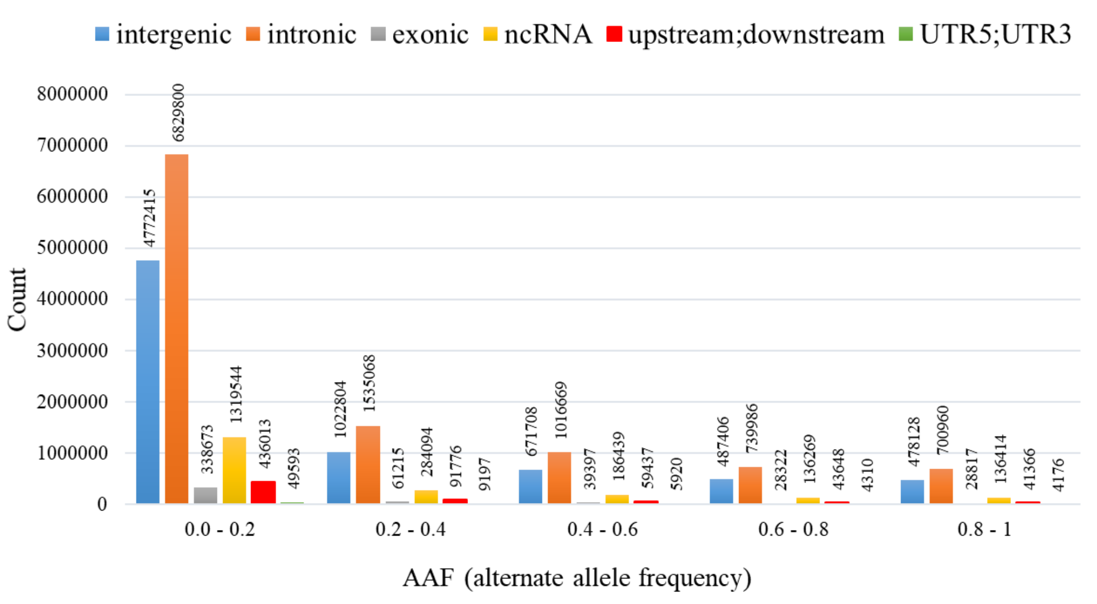


Figure 2.4 Alternate allele frequency distribution by annotation category for the complete set of detected SNPs (~19.6 million) in Ethiopian indigenous chicken populations

2.3.2 Genetic diversity

Genetic diversity within and among populations was investigated in a number of ways. Figure 2.5 shows the population-wise mean observed heterozygosity, which varied from 0.26 to 0.33. The local Horro and Batambie populations have the highest average heterozygosity (0.33), followed by Hugub (0.32). The lowest observed heterozygosities are reported for the Jarso population (0.26), Dikuli (0.28), and Metkilmat (0.28).

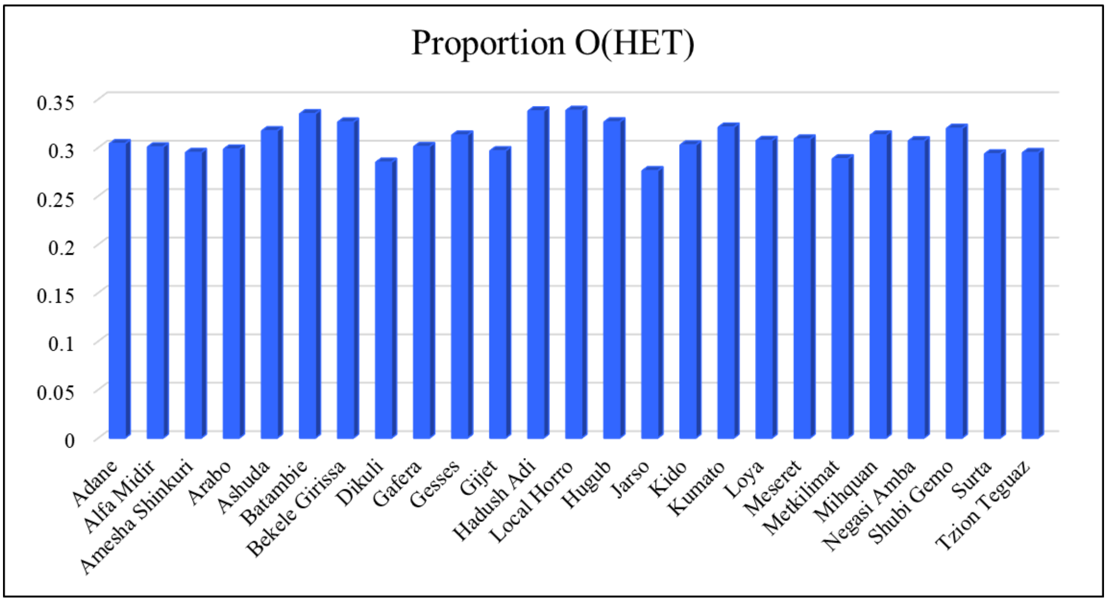


Figure 2.5 Proportion of observed heterozygous variants across the studied Ethiopian indigenous chicken populations.

To further examine the genetic diversity within populations, two other genetic parameters were measured, *viz* the inbreeding coefficient (*F*) and the nucleotide diversity (π) (Table 2.3). Average inbreeding coefficients range from 0.6% (Batambie) to 10.4% (Metkilimat). Majority of the populations show inbreeding below or near 5%, whereas only five populations show values ~10%. The average nucleotide diversity across the investigated populations indicates that the Alfa Midir population displays the lowest average value ($\pi = 0.0032$). In contrast, Bekele Girissa, Kumato, Meseret, Mihquan, Shubi Gemo, and Surta have the highest value ($\pi = 0.0038$). These values are similar to those obtained in previous studies for indigenous chicken (Luo *et al.*, 2020a; Sun *et al.*, 2021).

Table 2. 3 Mean inbreeding co-efficient (F) and nucleotide diversity (π) for each of the 25 studied chicken populations.

Population	**Inbreeding coefficient** ***F***	**Nucleotide diversity** π
Adane	0.0245	0.0036
Alfa Midir	0.0082	0.0032
Amesha Shinkuri	0.0784	0.0037
Arabo	0.0232	0.0034
Ashuda	0.0155	0.0035
Batambie	0.0064	0.0037
Bekele Girissa	0.0059	0.0038
Dikuli	0.1007	0.0037
Gafera	0.0553	0.0036
Gesses	0.0313	0.0033
Gijet	0.1039	0.0037
Hadush Adi	0.0971	0.0037
Hugub	0.0283	0.0034
Jarso	0.0507	0.0033
Kido	0.0744	0.0035
Kumato	0.0098	0.0038
Local Horro	0.0543	0.0036
Loya	-0.0002	0.0037
Meseret	0.0419	0.0038
Metkilimat	0.1044	0.0037
Mihquan	0.0334	0.0038
Negasi Amba	0.0133	0.0033
Shubi Gemo	0.0185	0.0038
Surta	0.1025	0.0038
Tzion Teguaz	0.0716	0.0035

2.3.3 Genetic differentiation and population structure among Ethiopian chickens

These analyses were performed using 14 million autosomal SNPs after applying further filtration on VQSR-passed 19.6 million SNPs (see Materials and Methods).

To examine the population divergence among chicken populations, *Fst* values were estimated. A pairwise-population differentiation (*Fst)* among all 25 populations is presented in figure 2.6. As a general rule, *Fst* < 0.05 between populations would indicate little genetic differentiation, 0.05-0.15 as moderate genetic differentiation, and > 0.15 as large differentiation (Hartl & Clark, 1997; Balloux & Lugon-Moulin, 2002). The figure shows a heat map indicating that the population differentiation is generally quite low, varying between 0 to 0.12. Only Jarso and Hugub populations show a consistent population differentiation of 0.1 against the rest of the populations, whereas values below 0.1 are present for almost all other pairwise comparisons.

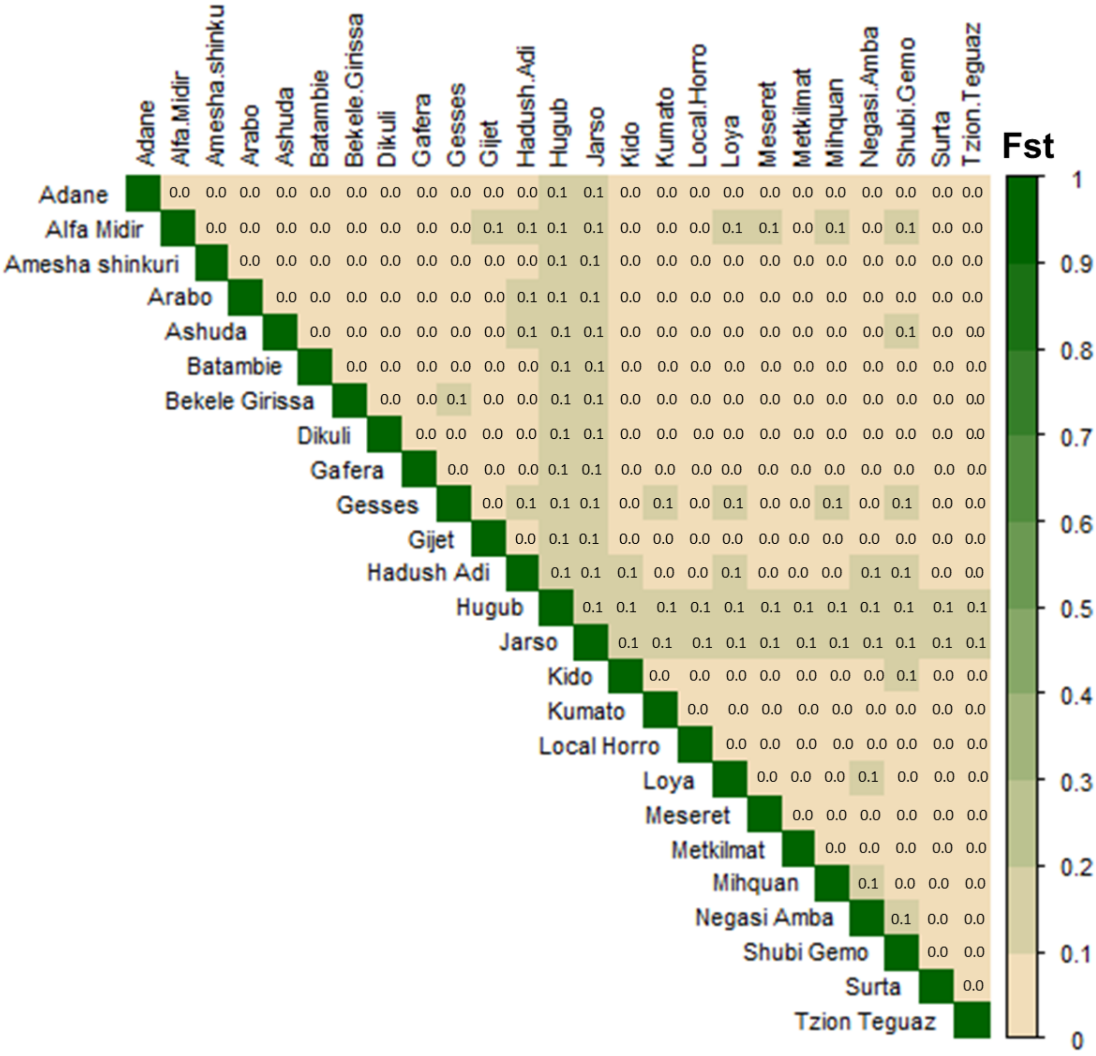


Figure 2.6 Heat map of pairwise weighted Fst values among chicken populations. The right panel explains the *Fst* value (colour) in relation to population differentiation.

The population structure was investigated by principal component analyses (PCA). The first, second, and third principal components accounted for 21.89%, 15.75%, and 10.53% of the proportion of variation among the 25 chicken populations (Figure 2.7a and b).

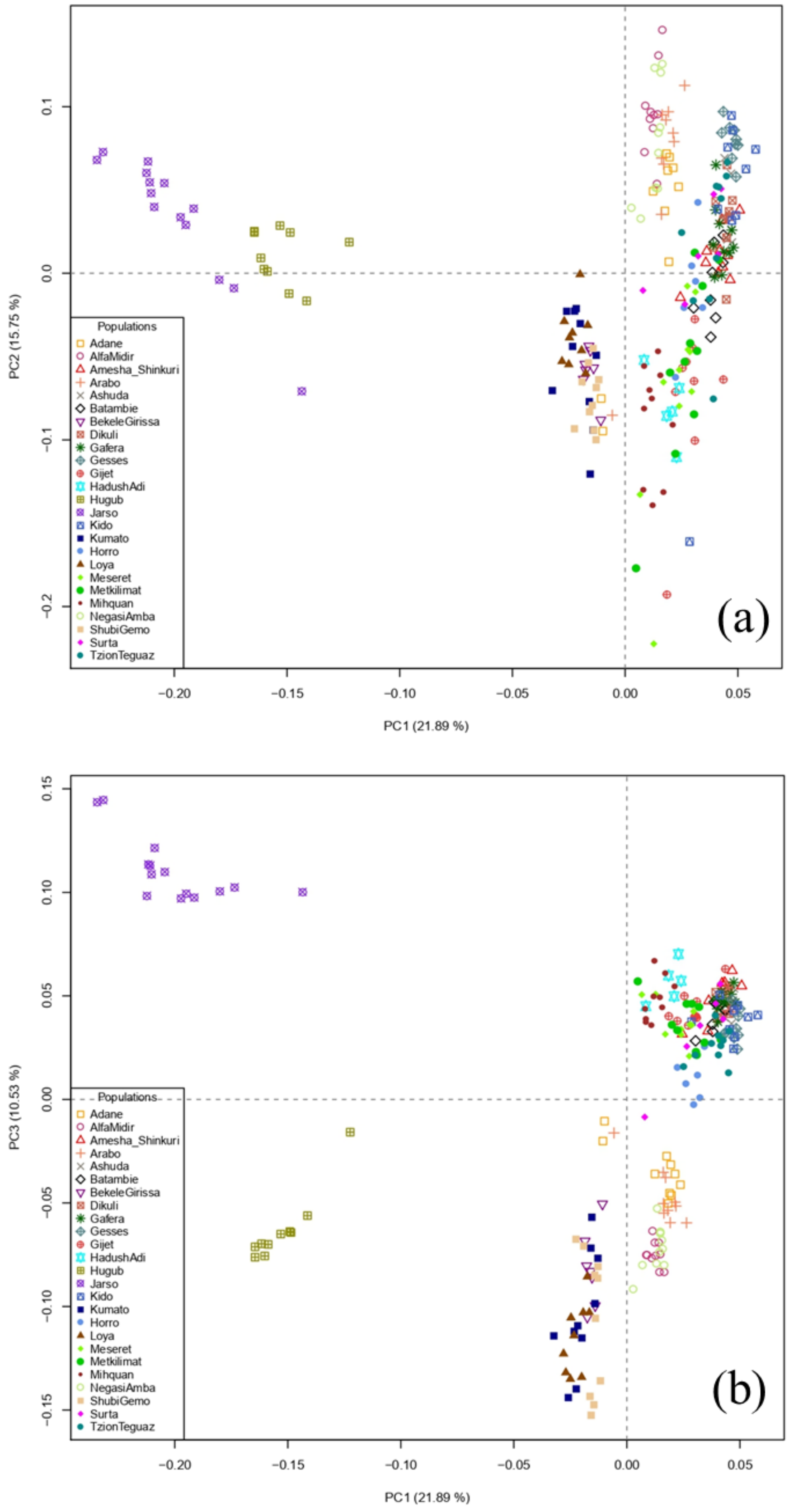


Figure 2.7 Principal component analyses: (a) PCA *vs* PC3 and (b) PC1 *vs* PC3 for Ethiopian chicken populations.

PC1 explains 21.89% of the total variation and clearly differentiates Jarso and Hugub from the remaining populations. PC2, with 15.75% of the variation, separates Metkilimat, Hadush Adi, Mihquan, Gijet, and Meseret populations from Gesses, Kido, Gafera, Horro, TzionTeguaz, Dikuli, Amesha Shinkuri, Batambie, Ashuda, and Surta (Figure 2.7a). However, PC3 (10.53% of variation) gathers most of these populations together in a single group (Figure 2.7b).

2.3.4 Admixture analysis

The genetic ancestry and admixture of the populations were explored through admixture analysis. We investigated 2 to 5 possible *K* values or ancestral gene pools (Figure 2.8). According to the cross-validation test (Evano *et al.*, 2005), the *K* = 3 is the best number of genetic ancestry for the Ethiopian chicken populations investigated (Supplementary Figure S2.1). *K* = 3 clearly separates Jarso and Hugub from the other populations.

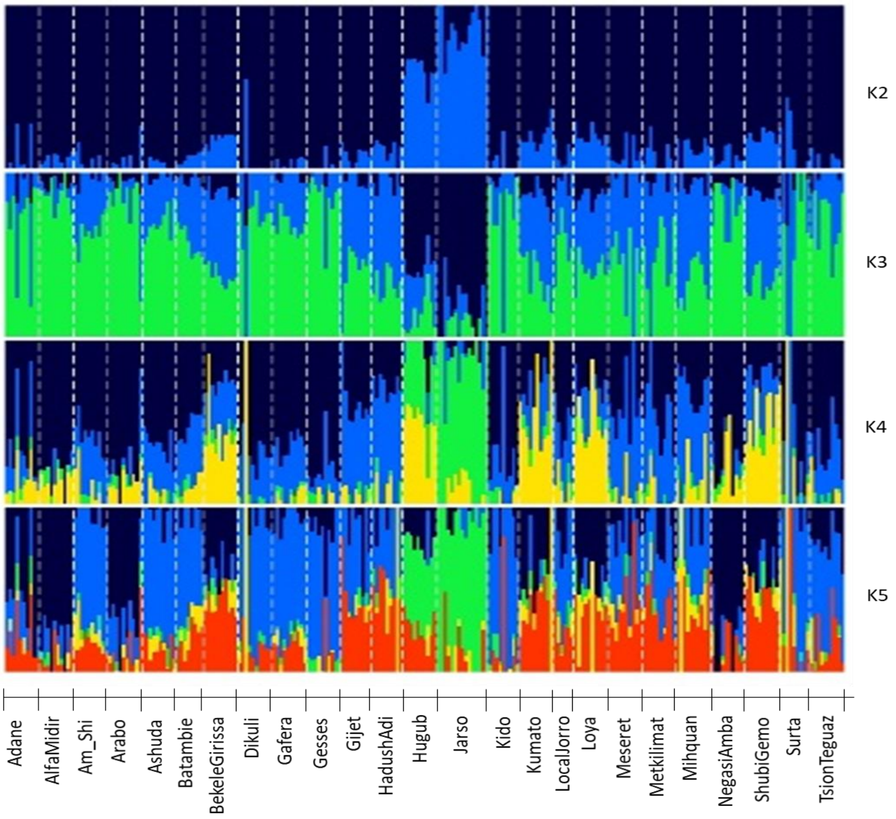


Figure 2.8 Admixture analyses of the 25 Ethiopian chicken populations (K = 2 to 5). Populations are demarcated with dashed lines.

2.3.5 Linkage disequilibrium (LD)

LD can be understood as a non-random assortment of alleles at different loci (Qanbari *et al.*, 2020). It can be influenced by diverse evolutionary forces like migration, selection, genetic drift, mutation, and recombination (Ardlie *et al.*, 2002; Karimi *et al.*, 2014). Therefore, scrutinizing LD levels is relevant to genetic diversity analyses since it is the first step towards understanding how demographic and evolutionary forces have shaped the LD structure of the genome. LD-based haplotype analyses provide an important platform for analysing signatures of selection in the genome. LD analysis is also important to assess how many markers are necessary to perform whole-genome association studies (Ardlie *et al.*, 2002). Moreover, for livestock species, LD is the basis of genomic selection, marker genotype imputation, marker-assisted selection (MAS), and quantitative trait loci (QTL) mapping (Nosrati, 2017).

Previous studies have shown an LD calculation bias due to small sample size and variance, particularly for LD-block calculations (Osabe *et al.*, 2007; Hong & Park, 2012; Hui & Burt, 2020). Therefore, six population groups were defined based on geographic closeness to increase the sample size in calculating LD decay and LD block (Figure 2.9). As an LD measure, we used a squared correlation coefficient between SNP pairs (r^2) (Hill, 1974) instead of D’ (Lewontin, 1964). The (r^2) score is known to be less sensitive to changes in effective population size and allele frequencies (Zwane *et al.*, 2016).

The number LD (r^2) between adjacent SNPs at distances up to 1 Mb was calculated for each population group. In the majority of the cases (73% - 87% of pairwise comparisons), large LD values (r^2 = 0.8 -1) were observed only for SNP pairs within 2 kb distance (Figure 2.10). Using r^2 = 0.8 as a threshold, no LD extended beyond 25 kb (see Supplementary Table S2.2). LD decay plots (Figure 2.11) also show that the r^2

values decrease rapidly after a 1-2 kb distance between variant pairs. With the increase of distances between variants, the r^2 values gradually reduced. Among the population groups, the average r^2 reached 0.2 between 1 kb to 2 kb for groups 3, 4, 5, and 6. In groups 1 and 2, r^2 is close to 0.25 – 0.3 within 1 kb to 2 kb distance.

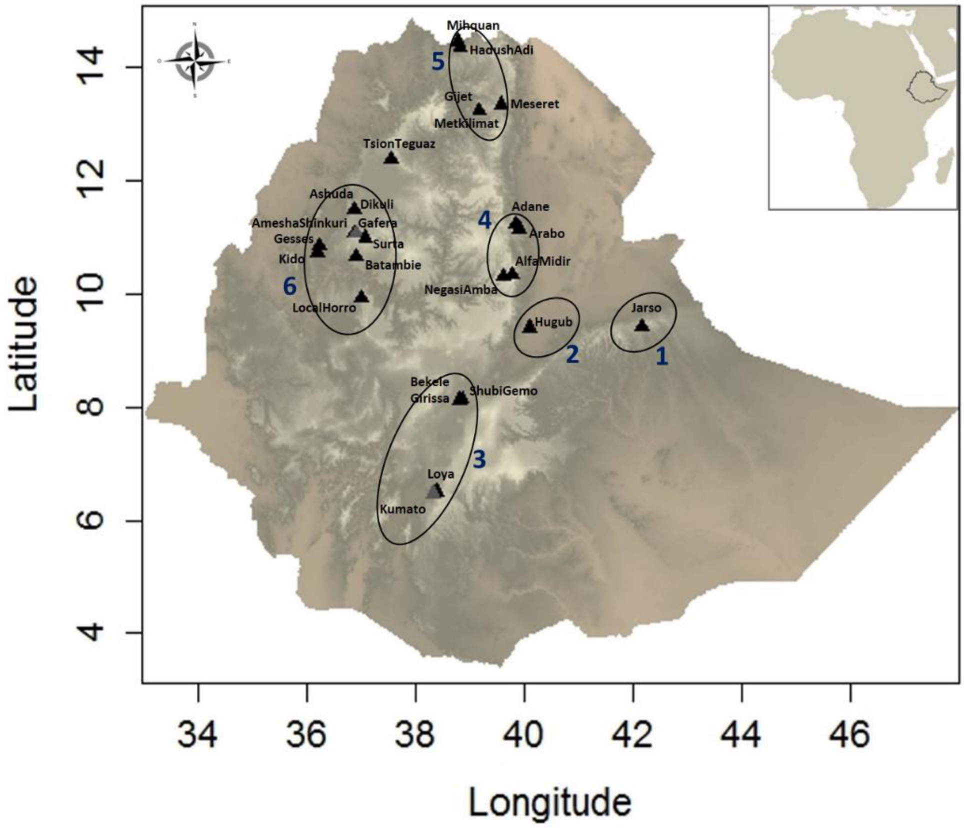
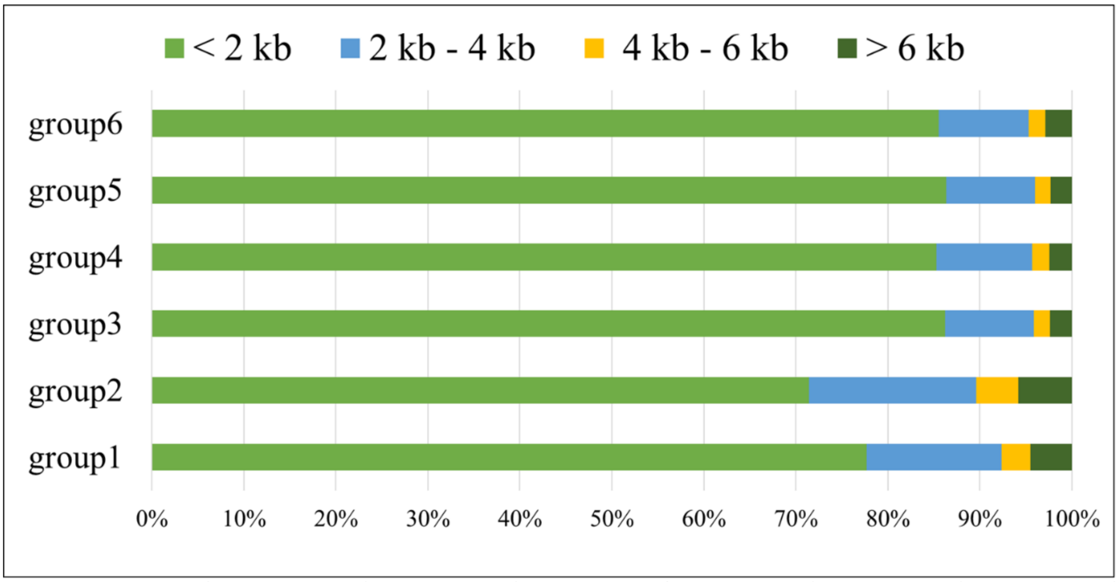


Figure 2.9 Geographic distribution of Ethiopian chicken populations denoting population groups defined for LD calculations.

Overall, the proportion of highly linked SNP (i.e., with r2 = 0.8-1) at different distances is similar between population groups except for group 1 (consisting of only 14 samples from the Jarso population) and group 2 (10 samples of Hugub population only) (Figure 2.10). Differences are most likely due to the small number of samples in groups 1 and 2 compared to the other clusters.



	group 1	group2	group 3	group 4	group 5	group 6
# of populations	1	1	4	4	5	10
# of samples	14	10	38	40	46	90
Total # SNPs	11,155,695	10,034,152	13,441,574	13,078,158	13,472,811	13,735,691
Thin_30 Total # SNPs	3,232,132	2,831,213	3,986,133	3,866,386	3,997,233	3,952,288

Figure 2.10 Proportion of SNP pairs showing high LD (r^2 = 0.8 – 1) in the six population groups across different kb distances.
The table at the bottom summarises the number of samples and SNPs used for LD analyses.

Similar to Figures 2.10 and 2.11, the distribution of haplotypes (based on LD estimation) across the genome (Figure 2.12) shows that most of the haplotype blocks (~ 90%) have a size < 2 kb for all the studied groups, and only a very small proportion (0.02% – 0.03%) have a size > 100 kb. The longest identified haplotype block is ~ 200 kb in length. The number of SNPs in the blocks ranges from ~5 to 1018 (Supplementary Figure S2.2). In all groups, between 80% to 90% of the haplotype blocks contain less than 5 SNPs and between 0.05 % to 0.06% more than 100 SNPs, consistent with the blocks' size as presented in Figure 2.12.

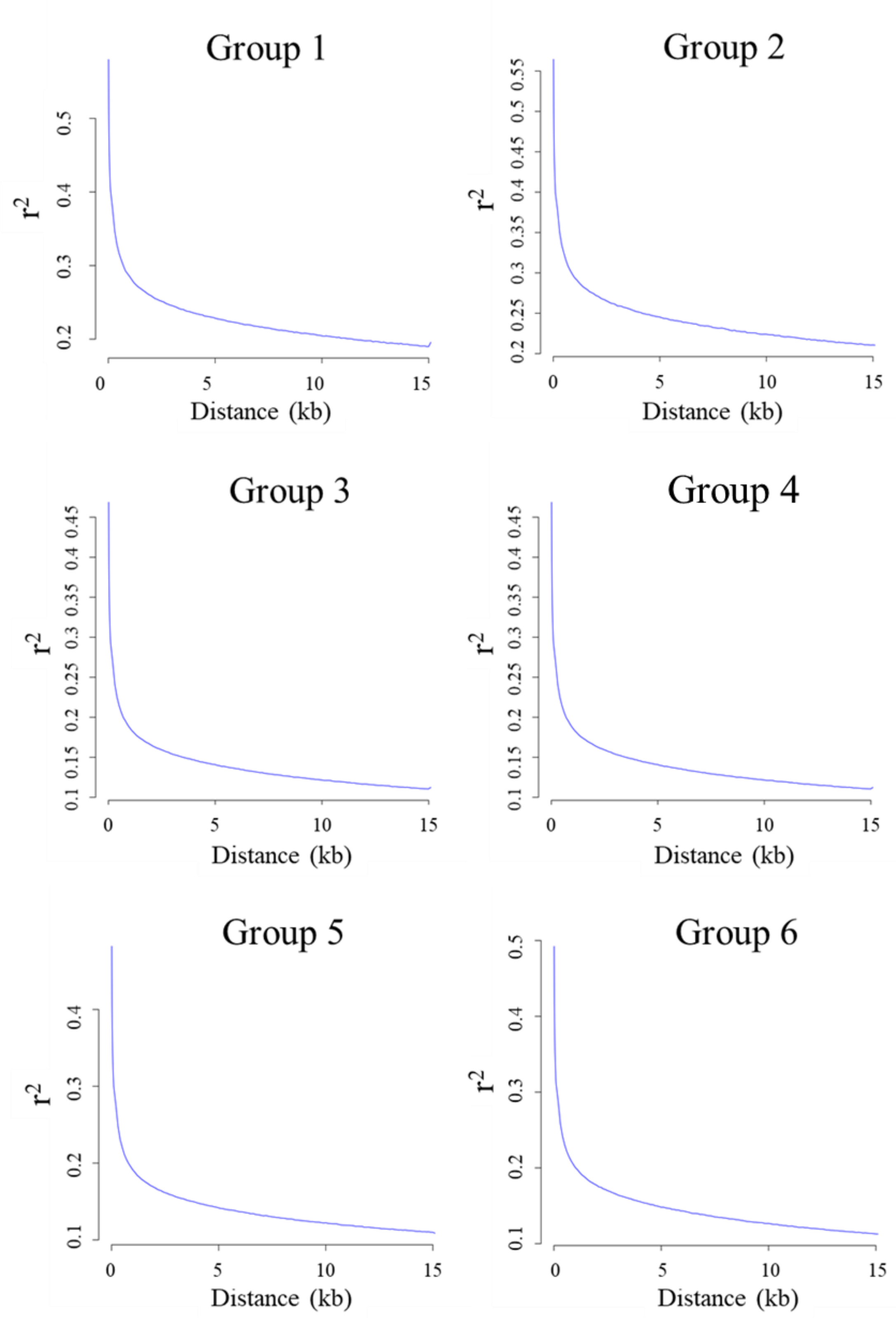


Figure 2.11 LD Decay trend in different Ethiopian chicken population groups. Based on average r^2 from variants within 15 kb across all autosomes.

The largest haplotype block of ~200 kb was detected on chromosome 8 (position 9070505 – 9270269 bp). It also showed many SNPs in groups 4 (660 SNPs), 6 (614 SNPs) and 5 (370 SNPs). This is a pericentromeric region with a small physical distance to the centromeric region in chromosome 8 (position 9300300 – 9800800), influencing the recombination rates in this genomic area, increasing the LD. The centromere is a complex structure that plays a central role in chromosome integrity and cell division (Barra & Fachinetti, 2018). Normally, regions around the centromere have large LD due to suppressed recombination (Bhakta *et al.*, 2015).

The blocks with the highest number of SNPs were further investigated. This block corresponds to a 159.7 kb region in chromosome 14 (position 14015491 – 14175264 bp). This block was identified in group 1 (Jarso), with 1018 SNPs. Interestingly, a haplotype block overlapping the same region was identified for group 3 (543 SNPs, 66.18 kb), group 4 (745 SNPs, 81.18 kb), group 5 (746 SNPs, 68.28 kb), and group 6 (829 SNPs, 93.34 kb). This region overlaps with the gene *JPT2,* which is a protein-coding gene involved in microtubule function (Karpova *et al.*, 2006) and playing an important role as a component of the NAADP receptor complex that is essential for pore channel signalling and control of viral cell entry (Gunaratne *et al.,* 2021).

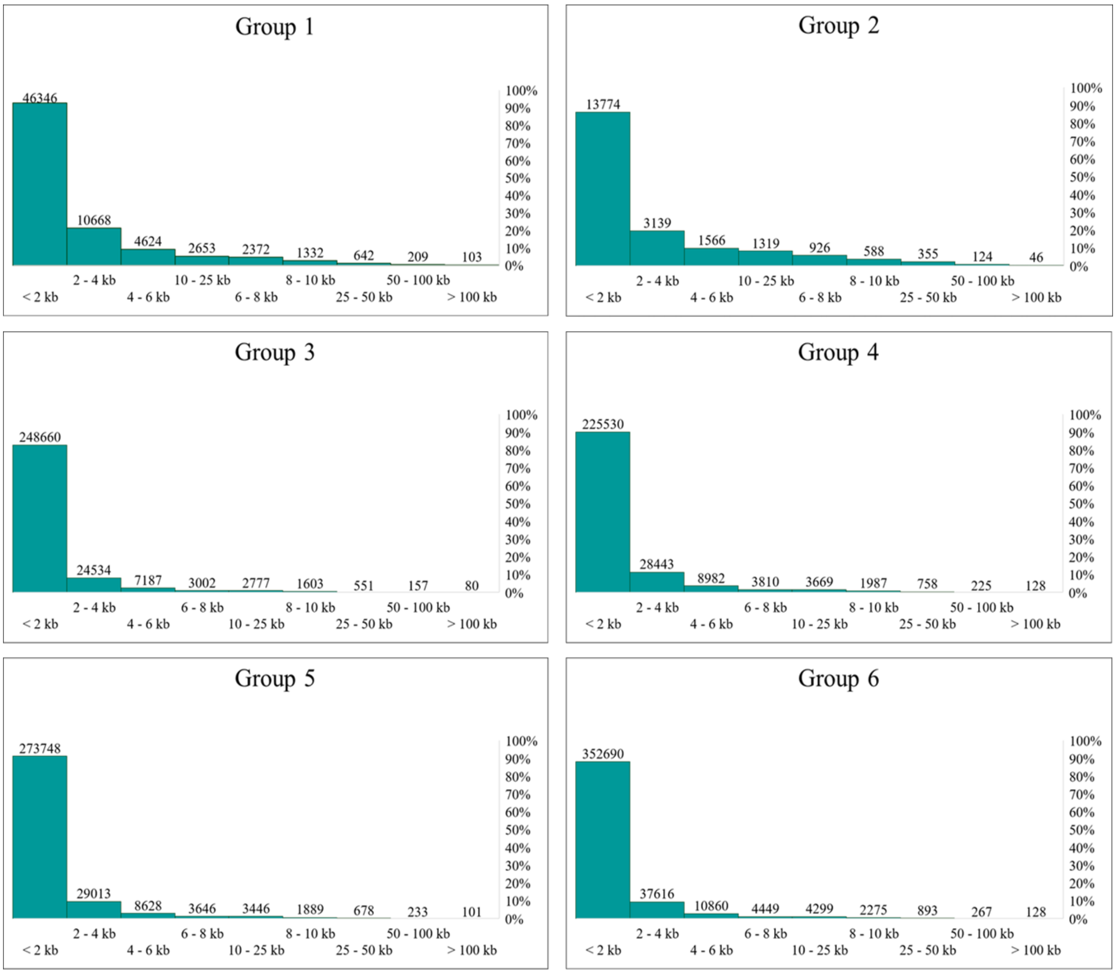


Figure 2.12 Number of LD blocks per range size identified in the six population groups. X-axis delimits range size in kb.

2.4 Discussion

By analysing whole-genome sequence data from 245 Ethiopian indigenous chickens, this chapter reports the discovery of millions of SNPs. It explores the genetic diversity and structure of 25 chicken populations from which these samples originated. Almost a third of the total number of SNPs identified in the present study (29% of 19.6M) are novel variants, which will be submitted to the European Variation Archive (EVA) and, therefore, will increase the global knowledge regarding the diversity of the chicken genome.

The sequence data in the present study were mapped against the latest version of the chicken reference genome (*GRCg6a* or *galGal6).* The SNP density found in this analysis (19 SNPs per kb) is smaller compared to that (21 SNPs per kb) reported for the same populations in a previous study by Kebede (2018) using *galGal5* as a reference. It is possibly a consequence of the improved quality of the latest genome assembly *galGal6.* In contrast, the SNPs density obtained in this study is higher than the density observed in experimental and commercial chicken populations (15 SNPs per kb) (Gheyas *et al.,* 2015). However, both these studies analysed a similar number of samples for SNP detection. This reflects greater genetic variability in the Ethiopian indigenous populations compared to commercial lines. While for most of the chromosomes, the mean SNP density is similar (ranging between 14 and 23 SNPs per kb), a few microchromosomes viz. chr16 and chr30 – 33 show a much lower density (6 – 12 SNPs per kb). Possibly the legacy of repetitive and/or GC-rich sequence regions at these chromosomes as well as differences in recombination rates may have resulted in mapping and/or sequencing issues. (Burt, 2005; Benjamini & Speed, 2012). In Addition, chromosomes 30 – 33 are the smallest microchromosomes of the chicken genome. Some still unmapped contigs might belong to these chromosomes (Solinhac et al., 2010). Considering that the reference genome used in this study (GRCg6a) does not show the full repertoire of variations in the genome due to missing genomic regions, technical errors of assembly, etc., a pangenomic approach could be applied to cover all genomic variants on the chicken genome (Vernikos, 2020). This method, permits the studying a set of all genes that are present in all genomes of an analyzed dataset. This could permit avoid missing data and possible missinterpretation of the results obtained when SNP data is explore.

The genetic diversity architechture of the sex (W/Z) chromosomes are different relative to the autosomes due to differences in the mutation rate and present non-estatic pseudoautosomal regions (Wilson, 2018). Several reasons were taking into account for excluding W/Z chromosomes in this study. First, the assembly of sex crhomosomes may not be perfect. These have low level of heterozygosity (Xu & Zhou, 2020), so many false positive will likely appear. Second, our data set comprises both male and female samples with different sex chromosomes schemes, which would require separate treatment of the sex chromosomes form the rest of the genome. However, further analysis should be carried out on this regions with attention on the mutation rates to avoid understimation of accumulated variation.

Considerable genetic variation within the studied chicken populations was also evidenced through the observed heterozygosity (0.26 – 33). Heterozygosity is an important factor for estimating the genetic variability in domestic animals (Zhang *et al.*, 2018). Thus, the fact that Ethiopian indigenous chickens display reasonable levels of heterozygosity reflects their genetic potential to adapt to diverse and changing natural environments (Groeneveld, 2010) and their importance as genetic resources. Genetic diversity in these indigenous populations also suggests great potential for improvement through breeding programmes.

Two parameters, inbreeding coefficient (F) and nucleotide diversity (π) were also calculated to assess the genetic diversity across the studied populations. Previous studies have evidenced that livestock species have suffered from strong inbreeding due to artificial selection followed by a decline in their genetic diversity (García-Gómez *et al.*, 2012; Kim *et al.*, 2013). Ethiopian village populations are managed and traded at a local scale (Manyelo *et al.*, 2020). Despite this, the level of inbreeding coefficient in most of the studied Ethiopian populations is low (below 5%), and only in few populations the values go ~10%. Inbreeding coefficients less than 5% do not represent substantial harm at the population level, while values higher than 10% may significantly affect individual fitness by accumulating deleterious mutation (Wright *et al.,* 1992; Slate *et al.,* 2004). The Ethiopian indigenous populations studied here also display similar nucleotide diversity values ($\pi \sim 0.003$) to values previously reported for other indigenous chicken populations (Luo *et al.,* 2020a; Sun *et al.,* 2021).

Although substantial genetic variations are observed within the studied populations, not much genetic differentiation was observed among the populations based on *Fst*, PCA, and admixture analyses. Only the Jarso and Hugub populations show some genetic differentiation with the rest of the populations. It is likely a legacy of the ancestral origin and dispersal of chicken across the African continent. Two routes of chicken dispersal have been reported (Mwacharo *et al.,* 2013b), one through the northeast of the mainland with an entrance in Egypt. The second involved maritime introduction along the East African coast that came directly from the Indian subcontinent. These two migrational paths are consistent with the main groups detected in Ethiopian village chickens. Jarso and Hugub populations are located in the East of Ethiopia, closer to the coast. In contrast, the other populations were sampled from the center, the western and the northern parts of the country. Admixture analyses show three genetic backgrounds, which may be reflecting the two chicken migration waves plus a third genetic background still unknown.

This study also explores the LD and haplotype structures of the studied chicken populations as part of the genetic diversity analysis. Previous studies reported average $r^2 = 0.3\text{-}0.4$ between adjacent SNP (separated by less than 5 Mb) across autosomes in pure chicken lines, and crossbred populations displayed $r^2 \approx 0.24$ (Fu *et al.,* 2015). In general, previous studies have shown that the crossbred populations display smaller haplo-block sizes compared to the pure commercial lines (Fu *et al.*, 2015). Larger haplo-blocks in pure commercial lines usually occur under intense artificial selection (Ardlie, 2002). The LD structure observed in Ethiopian indigenous chicken populations is similar to that of crossbred populations ($r^2 = 0.2 – 0.3$). The low LD levels found in the analysed population may suggest high levels of gene flow and possibly high recombination rate in the genome, increasing genetic diversity.

2.5 Conclusions

This chapter documents the genetic diversity and population structure of a large number of Ethiopian indigenous chicken populations, representing diverse agro-climatic conditions. Results clearly illustrate the presence of significant genetic diversity but a low level of population structure or differentiation. The large genetic variability denotes the adaptive capacity of these populations to the diverse agro-ecologies observed in Ethiopia and the potential for breeding improvement. On the other hand, low between-population differentiation indicates that selection signature analyses (as presented in later chapters) will have a lesser burden of false positives from the population structure effect. The results presented in this chapter represent the first stage of exploring the genomic-environmental adaptations to the diverse Ethiopian landscape. Chapters 4 and 5 will consider the information presented here to achieve the signature of selection analyses and identify the possible genomic marks as a response to the environmental pressure.

Chapter 3

Ecological niche modelling for environmental characterisation of indigenous Ethiopian village chickens

Abstract

Livestock populations in different geographic areas are often referred to as ecotypes. In evolutionary ecology, an ecotype is a population that is genotypically adapted to specific environmental conditions. However, livestock ecotypes are often arbitrarily defined based on their geographic origin or morphological features rather than a detailed environmental characterisation of their habitats. Using an Ecological Niche Modelling (ENM) approach (MaxEnt), this chapter presents a novel framework for defining livestock ecotypes through a detailed environmental characterisation of population habitats. The framework is developed and illustrated through the environmental characterisation of 25 indigenous chicken populations from across the diverse Ethiopian landscape and climatic conditions. Important features of the framework include: (i) a pipeline for selecting key environmental variables in population habitats, (ii) optimising model parameters of the ENM to maximise its predictive performance, (iii) predicting the distribution of suitable niche for each population across the Ethiopian landscape, and, finally, (iv) clustering of the studied populations into potential ecotypes based on niche similarity. The study identifies six environmental variables from an initial set of 34 agro-climatic variables as important predictors of the agro-ecology (habitat) of Ethiopian indigenous chicken populations. Characterisation of the habitats, followed by the clustering of populations living within similar habitats, allowed us to propose twelve different Ethiopian ecotypes. These ecotypes represent the population entities, which will be further characterised in Chapter 5 for the signatures of environmental adaptation at genome-wide level. The approach presented here provides a new tool for identifying, managing and conserving locally adapted populations.

3.1 Introduction

Environmental changes such as climate, feed resources, water availability, and disease pressure affect a species' adaptive response (Godber & Wall, 2014). Such response shows extensive variation due to genetic variability, phenotypic plasticity, and evolutionary potential (Manel & Holderegger, 2013). Therefore, understanding how different breeds or populations of livestock species may respond to environmental variation has a major implication for their conservation, sustainable utilisation, and productivity improvement (Dunn & Milne, 2014).

Indigenous populations of livestock species surviving in diverse geographic areas show adaptation to the unique conditions of their local environment. Such locally adapted populations of a species are referred to as ecotypes (Williams, 1966; Köhler-Rollefson, 2001). These ecotypes represent important genetic resources for the sustainable development of the livestock sector in the face of rapid global climate changes. In particular, these ecotypes may hold solutions for achieving climate resilience for sustainable productivity of the livestock sector and, accordingly, may contribute to food security (McDermott *et al.*, 2010; Harvey *et al.*, 2014).

In livestock species, ecotypes have generally been defined based on broad agro-climatic conditions linked to the geographic origin of breeds/populations or on some morphological features, which may or may not be related to environmental selection pressures (Tadelle *et al.*, 2003; Muchadeyi, 2007; Goraga *et al.*, 2011; Lyimo *et al.*, 2013; Keambou *et al.*, 2014; Sanarana *et al.*, 2016). For instance, Tadelle *et al.* (2003) studied five inferred Ethiopian chicken ecotypes defined by their geographic origins. Sanarana *et al.* (2016) investigated five Nguni cattle populations from South Africa (Makhathini, Pedi, Shangaan, and Venda), differentiated by their broad phenotypic differences as ecotypes. Until now, livestock ecotypes have been defined rather ‘loosely’ with no standardised approach in relation to their environment.

Adaptation to the local environment results from the interaction of many biotic and abiotic factors. Considering only a single or few environmental factors (e.g. adaptation to altitude or heat stress) in defining ecotypes fails to acknowledge other elements' contribution in the adaptation process. Since the core concept of an ecotype is an adaptation to a specific ecosystem, the more logical approach in defining an ecotype should be a detailed environmental characterisation of the agro-ecology (habitat). Populations may then be grouped as ecotypes based on habitat similarity. Then, the signature of adaptation may be investigated in the genomes of these ecotypes.

However, an agro-ecosystem or agro-ecology characterisation is inherently complex, considering the expected relationships and interactions between many environmental variables. A modelling-based approach may be used to untangle the relationships between the environmental variables to address this issue.

Ecological Niche Modelling (ENM) has been extensively applied for predicting the distribution of wild species and crop plants. Using environmental data from known species-presence locations (and species-absence sites), such modelling predicts a species’ potential geographical range across a landscape. It also estimates the contribution of environmental variables in the niche of a species. As a result, ENM has been a powerful tool in the conservation efforts of wild species (Thorn *et al.,* 2009). It has also been used to evaluate the risk from invasive species (Jiménez -Valverde *et al.,* 2011), for epidemiological studies (Cardoso-Leite *et al.,* 2014), and to study the effect of climate change on the future distribution of a species (Jeschke & Strayer, 2008). However, its application to livestock species has been rare. Recently ENM has been applied to study the habitat suitability of two exotic chicken breeds across the Ethiopian landscape (Lozano-Jaramillo *et al.,* 2018) to identify suitable habitats for their introduction.

MaxEnt (Phillips *et al.,* 2006) is the most widely used method for ENM, which uses presence-only data. The program applies a maximum entropy-based machine learning algorithm to predict species distribution by contrasting environmental parameters of species-presence locations against the rest of the study region - called the "background" - for which no presence/absence information is available. Although MaxEnt has been extensively applied in wild species, many of these studies have used it as a 'black box' by using mostly its default settings (Morales *et al.,* 2017). However, the default settings can be far from optimum, leading to overcomplex or over-simplistic models (Morales *et al.,* 2017; Yackulic *et al.,* 2013). Therefore, a careful selection of model parameters is required to achieve the best predictive power.

In this chapter, we develop a framework for identifying livestock ecotypes using ENM-based environmental characterisation with the MaxEnt program. Within the framework, we first describe the methods for selecting important environmental variables and optimum model parameters. Using these variables and model parameters, we then predict the niche distribution of the studied populations across a landscape, followed by clustering the populations based on their niche similarity to define distinct ecotypes. Finally, we discuss the potential applications of this modelling approach for livestock farming and management. These critical steps of the modelling and the framework are illustrated in indigenous Ethiopian chicken populations. The approach, however, should be equally applicable to any other livestock species.

3.2 Material and Methods

3.2.1 Populations sampled and environmental parameters

The Ethiopian landscape may be characterised by its altitudinal topography ranging from 126 m below sea level to 4620 m above sea level and climate variation with one or two rainy seasons separated by dry seasons. As a result, the country displays diverse agro-ecological zones and subzones (Mengistu, 2006). The 25 indigenous chicken populations analysed here (already described in Chapter 2) are spread across these several agro-ecological zones (Table 2.1 and Figure 2.1 of Chapter 2). The geographic coordinates (latitude and longitude) of the village sampled for each chicken population were recorded. In addition, for the purpose of the ecological modelling, nine other geographic coordinates in additional grids (separated by 1.2 km^2) were selected for each sampling village using Google Earth Pro 7.3.1.4507 (Google, 2016); this resulted in 250 data points across the 25 populations.

The environmental data – climatic and agro-ecological – were obtained from public databases. These were initially chosen as deemed important in terms of chicken production and biology. The data were accessed from these databases at a spatial resolution of 30 seconds (~1 km^2) (Table 3.1). Climatic data on temperature, precipitation, and water vapour pressure at different seasons were extracted from the WorldClim database (http://www.worldclim.org; v1 and v2) (Hijmans *et al.*, 2005; Fick & Hijmans, 2017). SoilGrids1km v0.5.8 database was used to access seven different major soil groups and soil properties (Hengl *et al.*, 2014). The total water capacity of the soil (mm water per 1-meter soil depth) was obtained from the Spatial Data Access Tool (SDAT; ORNL DAAC 2017) from NASA (Batjes, 2000) with a 0.5-degree grid. The ‘Harmonized World Soil Dataset - Land Use and Land Cover’ with 30 arc-second rasters

(FAO/IIASA/ISRIC/ISS-CAS/JRC, 2009) was used to access data on total cultivated land, forest land, and grass/shrub/woodland variables (Fisher *et al.*, 2008). The crop dominance variable was accessed from the Global Food Security Analysis-Support Data (GFSAD30, 2017; Teluguntla *et al.*, 2014).

For each environmental variable raster layer, the grids' dimension and extension were corrected and homogenised for 1 km^2 based on the Earth-fixed terrestrial reference system and geodetic datum WGS84 using 'rgdal,' 'maptools,' 'rgeos,' and 'raster' R packages.

Table 3.1 Climatic and agro-ecological variables used for the ENM analysis of the Ethiopian chicken populations.

	Variable	**Description**	**Units**	**Database**
Climatic variables	bio1	Annual Mean Temperature	°C	WorldClim - Global climate data http://worldclim.org
	bio2	Mean Diurnal Range	°C (bio2/bio7)	
	bio3	Isothermality	°C	
	bio4	Temperature Seasonality	°C	
	bio5	Max temperature of the warmest month	°C	
	bio6	Min temperature of the coldest month	°C	
	bio7	Temperature annual range	°C (bio5 - bio6)	
	bio8	Mean temperature of the wettest quarter	°C	
	bio9	Mean temperature of the driest quarter	°C	
	bio10	Mean temperature of the warmest quarter	°C	
	bio11	Mean temperature of the coldest quarter	°C	
	bio12	Annual precipitation	mm/m^2	
	bio13	Precipitation of the wettest month	mm/m^2	
	bio14	precipitation of the driest month	mm/m^2	
	bio15	Precipitation seasonality	mm/m^2 (Coefficient of variation)	
	bio16	Precipitation of the wettest quarter	mm/m^2	

	bio17	Precipitation of the driest quarter	mm/m^2	
	bio18	Precipitation of the warmest quarter	mm/m^2	
	bio19	Precipitation of the coldest quarter	mm/m^2	
	WatVapPress01	Water vapor pressure of the wettest month	kPa	
	WatVapPress08	Water vapor pressure of the driest month	kPa	
	Elevation	Meters above sea level	m.a.s.l.	
Soil variables	soil_pH	Soil pH	pH (x10 in H_2O)	Global gridded soil information https://soilgrids.org
	CatEx_Capacity	Cation Exchange capacity	cmolc/Kg at depth 0.00 m	
	BulkD	Bulk Density	Kg/m^3 at depth 0.00 m	
	Organic_Carbon	Soil organic carbon content	g/Kg at depth 0.00 m	
	Clay	Clay content	mass fraction in %	
	Silt	Silt content	mass fraction in %	
	Sand	Sand content	mass fraction in %	
	Water Capacity	Total available water capacity	mm^2/1 mt soil depth	Spatial Data Access Tool (SDAT) - NASA https://webmap.ornl.gov/ogc/wcsdown.jsp?dg_id=
Vegetation variables	Forest	Forest cover	%	Harmonized World Soil Dataset http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/
	Grass	Grass/shrub cover	%	
	Land_Use	Land use for agricultural purposes	%	
	Crop	Crop dominance (major crops)	Category	Global Food Security Analysis-Support DATA https://www.croplands.org/

3.2.2 Exploration of the agro-climatic variables

First, the contribution of the environmental variables on the habitat/niche structure of the studied populations was investigated by Principal Component Analysis (PCA) using the R package 'stats.' Then two complementary approaches were employed to explore the environmental data with the aim to select a shortlisted set of variables for the MaxEnt modelling. The normality of the variables and their correlations were assessed by estimating the Shapiro-Wilk test and Spearman's rank correlation coefficients, respectively in RStudio v1.1.419 (R Core, 2013). Outputs of these tests were plotted using the R package 'ggpubr' (Kasambra, 2017). The R package 'MaxentVariableSelection' was used to identify the most important uncorrelated variables with a relatively large contribution in the habitat structure of Ethiopian indigenous chickens (Jueterbock *et al.*, 2016). Arguments for running MaxentVariableSelection were: maximum correlation threshold of 0.6, minimum contribution threshold of 4% for each variable, and regularization multiplier (RM) values ranging from 0.5 to 15 with an increment step of 0.2. Model performance was assessed with the AUC (**A**rea **U**nder the Receiving Operator **C**urve) and AICc (**A**kaike **I**nformation **C**riterion corrected for small samples) values generated during the analysis.

3.2.3 MaxEnt modelling

To achieve the best predictive power, we initially applied MaxEnt v3.4.1 (Phillips *et al*., 2006) using different combinations of "feature" classes (FCs: linear, quadratic, product, hinge, categorical, and threshold) and regularization multiplier (RM) values (from 0.1 to 6 with steps of 0.5). Then, based on AICc values, the best combination was chosen with the ENMeval package (Muscarella *et al*., 2014) for the final model. All sampling points and selected environmental variables in the previous step were used for preliminary and final models.

MaxEnt split presence points (coordinates) into ‘training’ and ‘test’ data. The training data is used to create the model, whereas the test data is used to assess the model accuracy. The best method for partitioning the occurrence or presence data (block, checkerboard, checkerboard2, jackknife, and randomKfold) for training and testing purposes was evaluated with ENMeval.

Using the optimal model parameters and filtered set of variables, ENM was run for individual populations by using a sub-sample of 25% as test data while retaining the remaining 75% as training data. Two outputs - logistic (presence probability) and cumulative (omission rate) - were produced, but the cumulative logarithmic scale was used to plot the habitat suitability maps. Among the different model performance measures generated by MaxEnt, the AUC value was used to test how well a model predicted the habitats’ potential suitability. The AUC values were interpreted as follows: AUC > 0.5 has higher than random predictive power, AUC = 0.5 as a random predictor, and AUC < 0.5 as worse than a random predictor (Jiménez-Valverde, 2012). A jackknife assay was performed for the test and training data to assess the variables’ importance and contribution.

3.2.4 Classifying populations into potential ecotypes

Populations with similar habitat suitability maps from the MaxEnt modelling were grouped to define chicken ecotypes. Different approaches were applied to measure the similarity between suitability maps using the ENMtools Perl software (Warren *et al.*, 2010; Warren & Seifter, 2010). The first assessment was a pairwise Pearson correlation between suitability map files (ASCII raster files). Correlation coefficient values range between +1 and -1, where 1 is a total positive linear correlation, 0 is no correlation, and -1 is a complete negative correlation. The second evaluation was the niche similarity statistic “I,” which ranges from 0 (no overlap) to 1 (identical niche models) (Warren *et al.,* 2010). Suitability maps from MaxEnt logistic output were used for the pairwise comparisons as ENMtools only used this output. Finally, the grouping of the populations based on niche similarity was performed based on hierarchical clustering (R package ‘cluster’). For this, the correlation coefficients and the niche overlap values were converted into “Euclidean” distance using the R package ‘stats.’ Different hierarchical clustering approaches (minimum and maximum linkage, UPGMA, and Ward method) were evaluated to measure the cluster strength based on their agglomerative coefficients. The Ward method produced the highest agglomerative score, and then it was used for the final clustering of the investigated populations. Dendrograms and heatmaps for each similarity dataset were produced using ‘gplots’ and ‘RColorBrewer’ R packages. Niche similarity based on the above approaches and their visual confirmation constituted the approach for grouping the populations into potential ecotypes.

3.3 Results

The principal aim of this study is to present a framework for identifying livestock ecotypes by performing a detailed characterisation of population habitats using ENM. The method presented here provides a pipeline for (i) selecting environmental predictors, (ii) choosing model parameters for executing ENM, (iii) evaluating model performance, (iv) creating suitability maps for individual populations across a landscape, (v) assessing the similarity of the suitability maps from different populations, and (vi) clustering populations based on suitability maps into similar groups to define ecotypes. The overview of the framework and its application are illustrated in Figure 3.1.

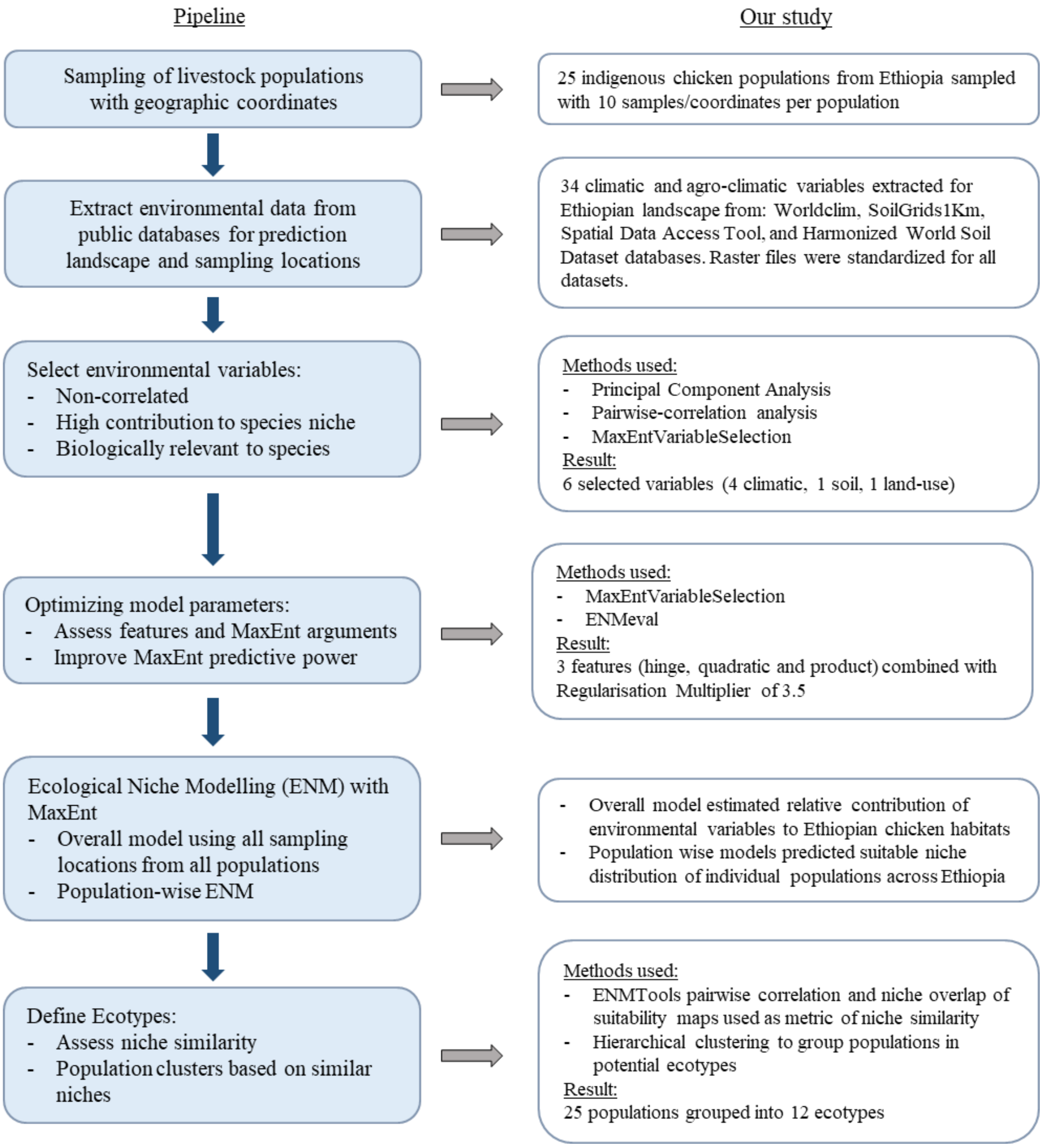


Figure 3.1 A framework for defining livestock ecotypes based on environmental characterisation using the ENM approach.

3.3.1 Selection of agro-climatic predictors

The selection of agro-climatic variables is a critical part of ENM to maximise predictive performance. The choice of environmental variables can limit the prediction of distribution for a species or taxa (Warren *et al.*, 2014; Zeng *et al.*, 2016). Using few variables will decrease the model's predictive power, and too many variables will overburden the model. Besides, using correlated variables can produce biased models that do not accurately represent a species niche (Bradie & Leung, 2017). Therefore, selecting environmental predictors in this study entailed starting with a large set of variables and then shortlisting them by removing the highly correlated ones and those with a little contribution in explaining the model. Initially, 34 environmental variables were chosen from public databases based on their potential relevance to chicken biology and production. These variables were grouped into three clusters: (1) climatic variables (n = 22), with potential impact on the physiological tolerance of chickens; (2) soil variables (n = 8), which may influence food availability; and (3) vegetation and land cover (n = 4) variables, potentially affecting both food availability and exposure to predation (Table 3.1). Variables were then shortlisted based on three different analyses as described below.

3.3.1.1 PCA of populations based on environmental data

PCA was performed for each variable cluster separately to investigate the population habitat structure. PCA plots in Figure 3.2 display the population variance distribution due to the diversity of the variables tested. Arrows indicate the direction of the variables in relation to the PCs plotted. Arrow lengths indicate the extent of contribution of a variable in explaining the habitat structure. Remarkably, PC1 and PC2 explain over 60% of the variants in all three groups, although populations vary widely among the PCAs. Despite differences between PCA plots, several populations persistently appear close to each other in all PCAs, such as Loya and Kumato, Gafera and Amesha Shinkuri, Gesses and Kido, Shubi Gemo and Bekele Girissa. It suggests a similar landscape configuration for these population groups, irrespective of the environmental variables investigated.

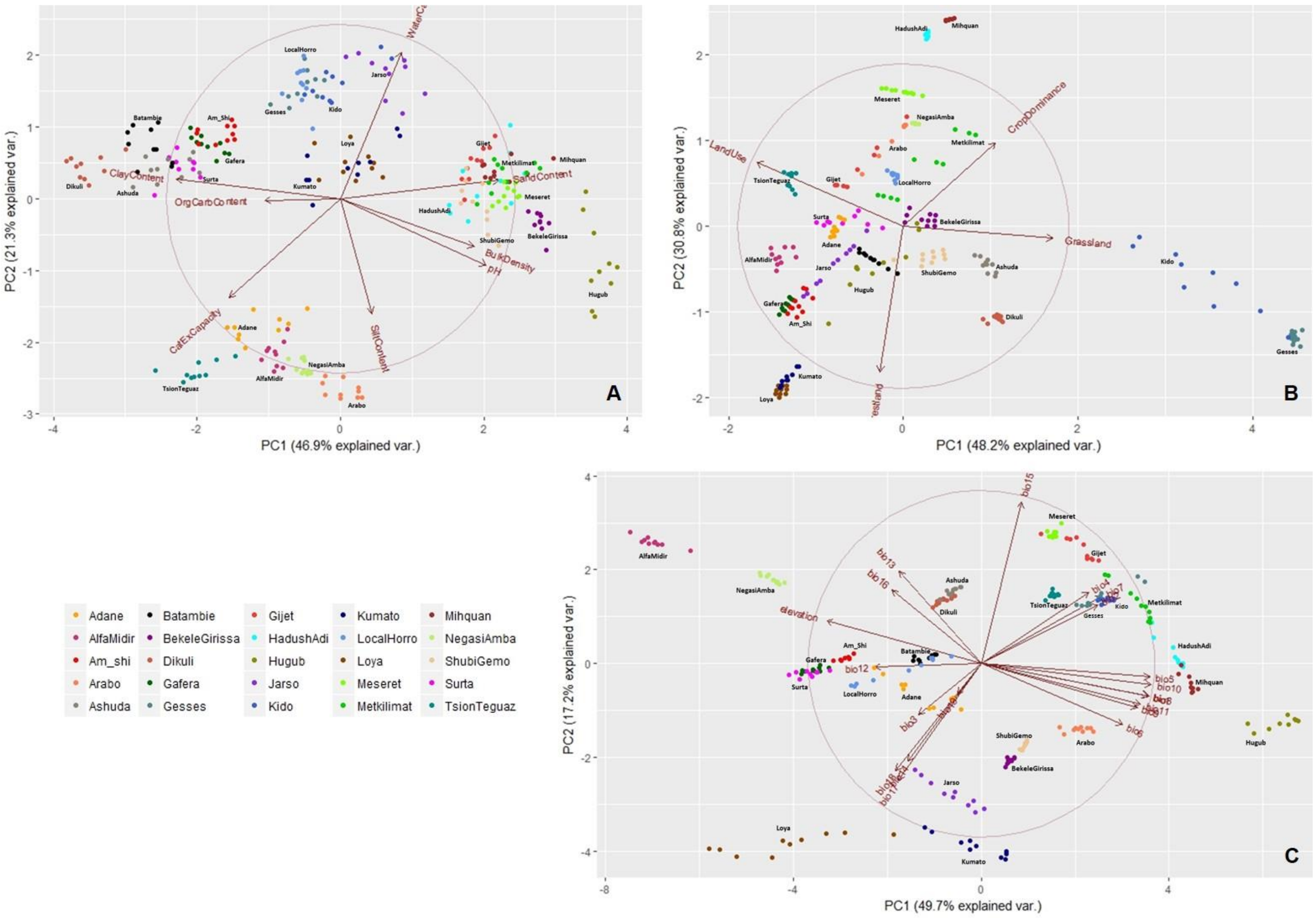


Figure 3.2 A framework for defining livestock ecotypes based on environmental characterisation using the ENM approach. (A) Soil variables, (B) Vegetation and land cover variables, and (C) Climatic variables.

3.3.1.2 Correlations among environmental variables

None of the climatic and agro-ecological variables had a normal distribution in Shapiro-Wilk normality testing ($P > 0.001$) (Supplementary Figure S3.1). Therefore, Spearman's rank-order correlation was used to identify any positive correlation between a pair of variables. Figure 3.3 (A-C) shows the correlations of variable pairs within the three clusters. The strength of correlation was defined as: $r_s \geq 0.8$ as very strong, $r_s \geq 0.6$ to < 0.8 as strong, $r_s \geq 0.4$ to < 0.6 as moderate, $r_s \geq 0.2$ to < 0.4 as weak, and $r_s \geq 0$ to < 0.2 as very weak. For the selection of variables for ENM, we used the threshold of $r_s < 0.6$.

Within the climatic variable group, strong to very strong correlations are observed ($r_s = 0.60\text{-}0.98$) among bio1 (annual mean temperature), bio4 (temperature seasonality), bio5 (maximum temperature of the warmest month), bio6 (minimum temperature of the coldest month), bio8 (mean temperature of the wettest quarter), bio9 (mean temperature of the driest quarter), bio10 (mean temperature of the warmest quarter), bio11 (mean temperature of the coldest quarter), and water vapour pressures (representing humidity).

Among the precipitation variables, strong to very strong correlations ($r_s \geq 0.72$ to 0.97) are observed among bio19 (precipitation in the coldest quarter), bio12 (annual rainfall), bio13 (precipitation of the wettest month), and bio16 (precipitation of the wettest quarter), and among bio14 (precipitation of the driest month), bio17 (precipitation of the driest quarter) and bio18 (precipitation of warmest quarter). Elevation (bio20) shows a very strong negative correlation with many temperature variables (viz. bio1, bio5, bio6, bio8, bio9, bio10, and bio11) and water vapor pressure.

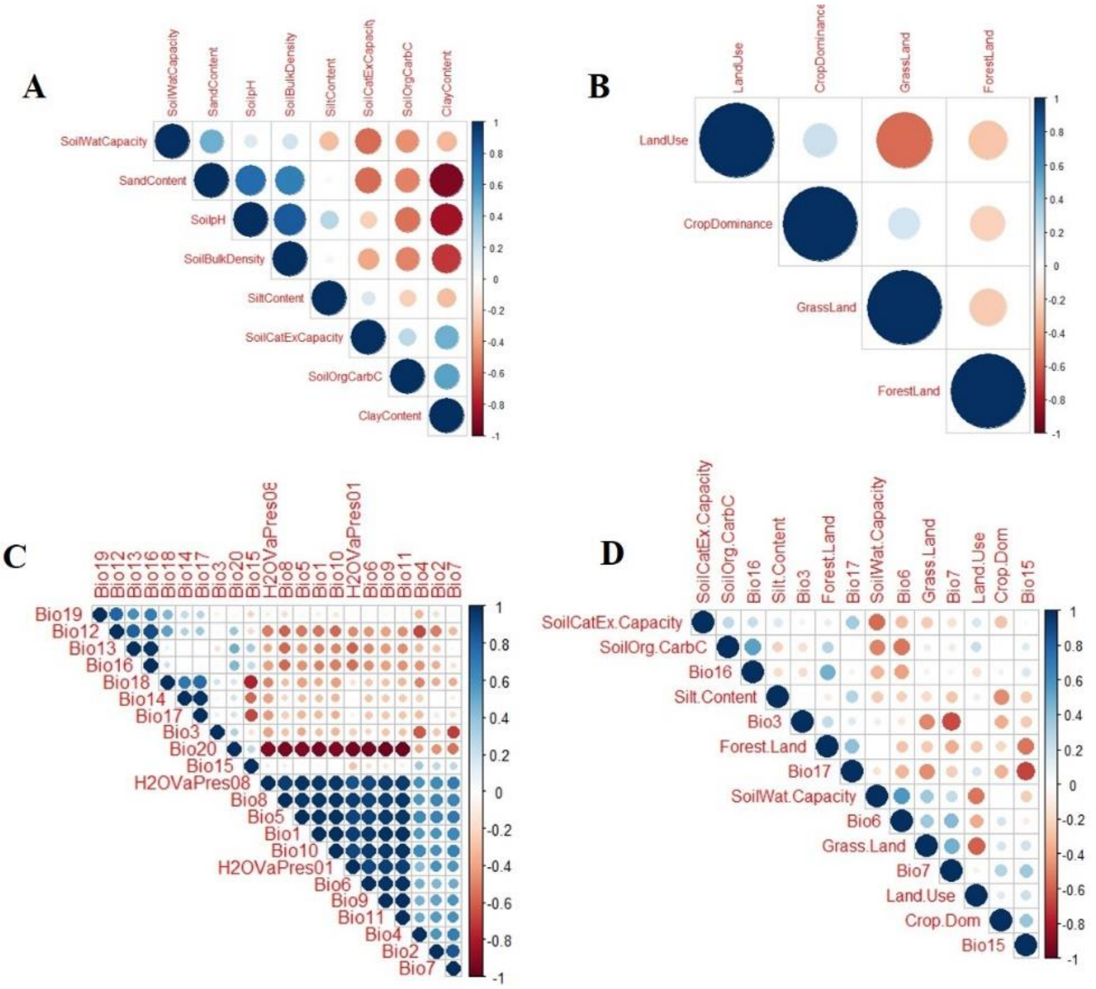


Figure 3.3 Spearman correlation test for three different groups of climatic and environmental parameters evaluated.

(A) Soil variables, (B) Vegetation and land cover variables, (C) Climatic variables, (D) Correlation test between non-correlated variables from different groups. (Order of the variables is based on the strength of correlation).

Among soil variables, soil bulk density and pH show very strong correlations ($r_s > 0.8$), while sand content shows strong correlations with both soil pH and soil bulk density ($r_s = 0.76$ and 0.68, respectively).

After taking one variable from each correlated pair (with $r_s > 0.6$ with $P < 0.001$) along with other non-correlated variables, 14 variables were retained. A final correlation test among these did not find any pair with $r_s > 0.6$ (Figure 3.3D).

3.3.1.3 MaxEntVariableSelection

MaxEntVariableSelection package was used to simultaneously assess the correlation status and contribution level of the variables in chicken agro-ecologies and then to shortlist a set of predictors with low or no correlation among them ($r \leq 0.6$). The shortlisting process identified eight variables with a contribution level ranging between 4 to 15%. However, we further excluded two variables, crop dominance and grass-land, as they had shown high multi-collinearity with LandUse (Gheyas *et al.*, 2021). This left us with a final set of 6 variables for the ENM: minimum temperature of the coldest month (bio6 - minTemp), precipitation seasonality (bio15 - precSeasonality), precipitation of wettest quarter (bio16 - precWQ), precipitation of driest quarter (bio17 - precDQ), soil organic carbon, and, land use (cult_T). Figure 3.4 shows the relative contribution of the six variables.

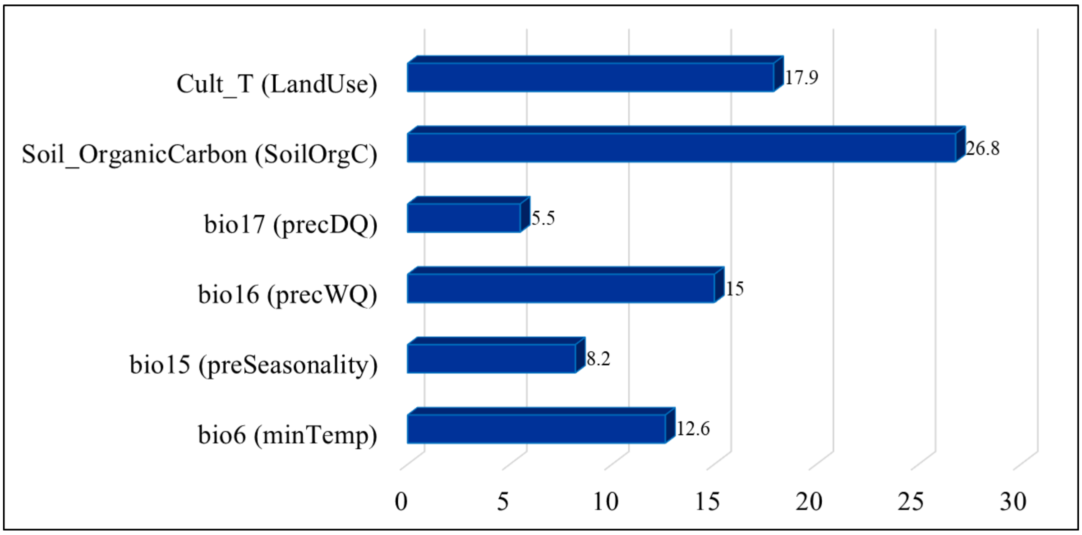


Figure 3.4 Percent contribution of each of the six agro-climatic selected variables to the overall environmental diversity calculated using the MaxEntVariableSelection R package.

3.3.2 Selection of the best model configuration

An optimal combination of FCs (feature classes) and RM (regularization multipliers) is crucial for achieving a balance between the goodness of model fit and model complexity (Muscarella *et al.,* 2017). FCs are an expanded set of transformations of the original environmental variables that allow complex relationships to be modelled (Morales *et al.,* 2017). Several FC options are available in MaxEnt, such as linear (L: the variable itself), quadratic (Q: square of the variable; fits variance), product (P: product of two variables; fits interaction effect), threshold (T: a step function with the different response below and above the threshold), hinge (H: similar to T, but the response above or below the threshold is linear with a positive or negative co-efficient) and categorical (C: fits categorical variables) (Elith *et al.,* 2011). On the other hand, the RM value is used to prevent over-complexity or over-fitting a model by controlling the intensity of the chosen FCs (Morales *et al.,* 2017). A highly complex model may have a good fit, but it may not generalise well (Ellith *et al.,* 2011). The RM value causes a trade-off between model fit and model complexity. Therefore, a series of MaxEnt models combining different FCs and RMs were performed and evaluated in ENMeval using the six selected environmental variables and collectively taking all 250 sampling locations.

In total, 208 RM-FCs combinations were evaluated (Supplementary Table S3.1). AICc value instead of AUC test value was chosen as the evaluation parameter to identify the best RM-FCs combination. AICc has been described as the better option because it is an index that reports the trade-off between fitting (likelihood) and parsimony (number of parameters) of the model (Lobo *et al.,* 2008; Galante *et al.,* 2017). T (threshold) and C (categorical) features did not make any difference in the AICc values at any RM combination; hence these features were dropped in subsequent analyses. Moreover, L (linear) and H (hinge) features were not used together as both represent linear responses, so their combined use is redundant (Elith *et al.*, 2011). Therefore, the feature combinations evaluated were: L, Q, P, H, LQ, LQP, HQC, and HQP. HQP (hinge, quadratic, and product) feature combined with an RM value of 3.5 showed the lowest AICc value (6039.084), indicating that this parameter configuration is expected to provide the best predictive performance for our study (Figure 3.5).

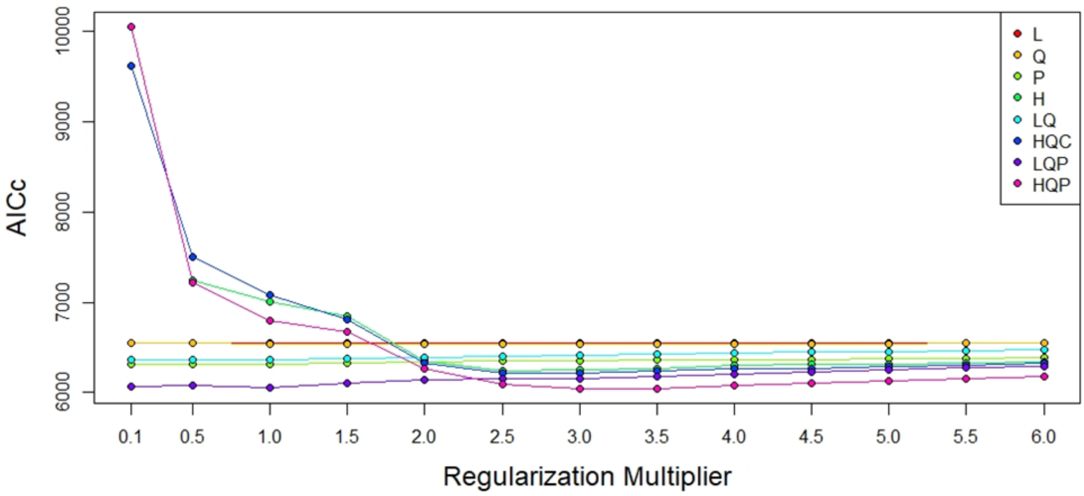


Figure 3.5 AICc values for different combinations of Feature Classes (FCs) and Regularization Multipliers.
The FCs shown are: Linear (L), Quadratic (Q), Product (P), Hinge (H), and Categorical (C).

To assess if these optimized parameters made any difference in prediction over the default settings in Maxent, we compared the suitability maps based on these two parameter sets, taking all sampling locations together (Figure 3.6). Although the two maps showed a high similarity (niche overlap = 0.95), the optimised parameters predicted more suitable regions. These selected parameters were therefore used for all subsequent MaxEnt modelling.

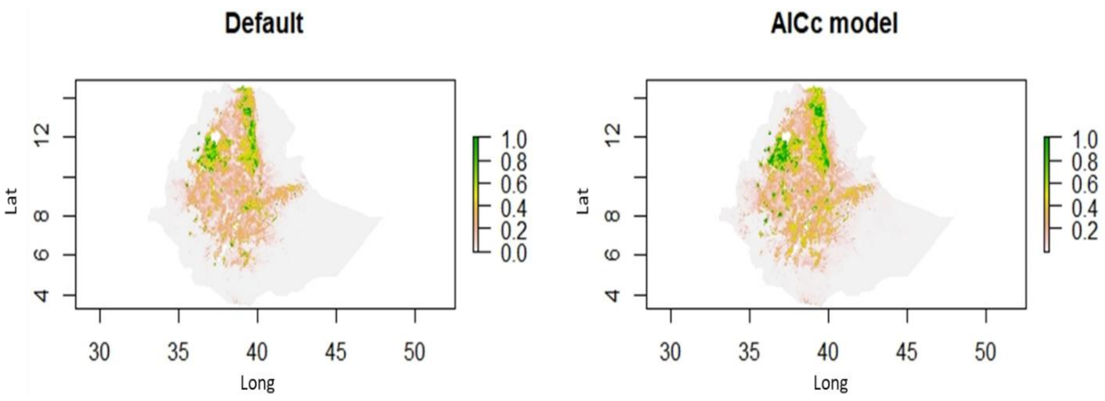


Figure 3.6 Suitability maps for the complete set of Ethiopian chicken populations. Default settings provided by the MaxEnt approach and identified parameters (HQP features and beta-multiplier = 3.5) using ENMeval were used.

3.3.3 Final ENM using selected environmental variables and parameter setting

Using the chosen parameter configuration and environmental variables, a final MaxEnt model with all 25 populations (250 presence points) was executed to assess the model performance further and obtain a refined estimate of the variable contributions on Ethiopian chicken populations as a whole (Figure 3.7).

This model predicted suitable habitats for Ethiopian village chickens with high success rates and low omission rates for training and test data that closely matched the expected omission (Figure 3.7A). The omission is the false-negative rate (i.e. where the model wrongly predicted unsuitable conditions for an occurrence location). Likewise, the model produced high AUC values: 0.911 and 0.891 for training and test data, respectively (Figure 3.7B). A high AUC value indicates that the model classifies presence more accurately than a random position, and AUC > 0.9 should be considered an excellent predictive ability (Pearson, 2010). The jackknife results for individual environmental variables also showed higher than random predictive power with AUC > 0.64 for all variables (Figure 3.7C).

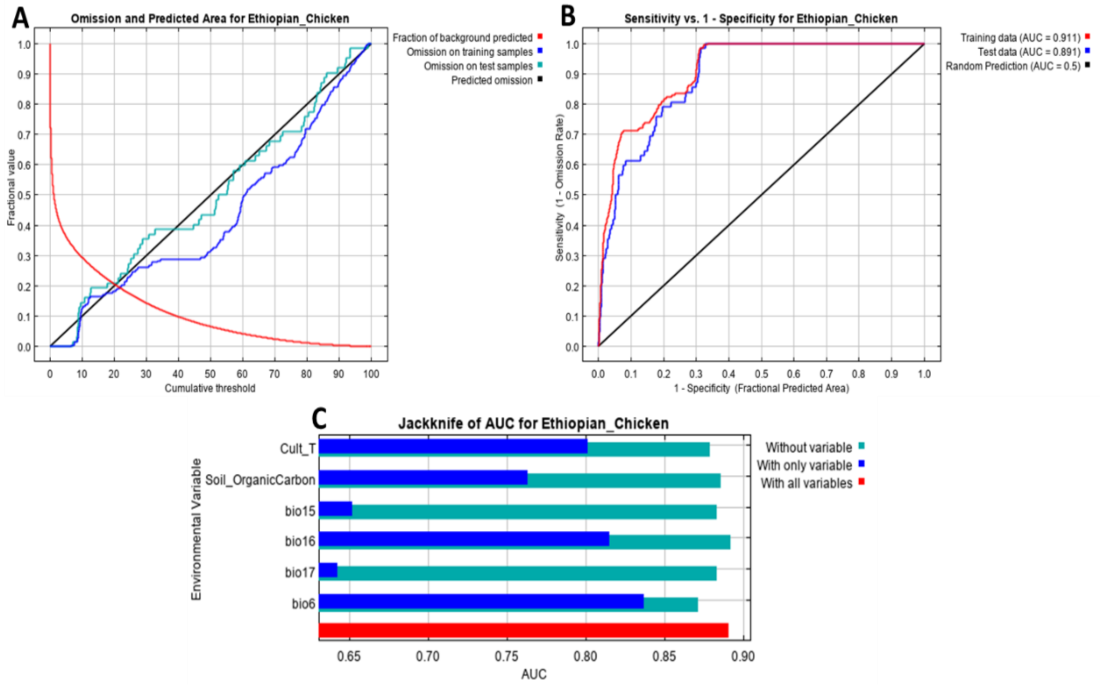


Figure 3.7 MaxEnt model based on six selected variables using the best combination of feature classes (HQP) and the regularization of multiplier value (3.5).
(A) Omission rates and predicted areas at different cumulative thresholds on test and training data, (B) Receiver Operating Curve for training and test data, (C) Jackknife result for AUC (Area Under Receiver Operating Curve) for each selected environmental predictor.

MaxEnt generates several other useful statistics to assess individual environmental predictors, namely - Jackknife of gain function from test and training data (Figure 3.8), response curves per variable (Figure 3.9), and percent contribution based on permutation importance (Figure 3.10).

The jackknife result of test and training gains by individual variables and all variables together is an important metric for assessing variable contribution and model performance. The gain is a penalised likelihood function that MaxEnt maximizes to find a model that best differentiates presence from background locations (Elith *et al.*, 2011). Also, the gain indicates how closely the model is concentrated around the presence samples. It is used to measure the goodness of fit in generalized additive and linear models (Phillips *et al.*, 2006). Jackknife results of training and test gain indicated that bio6 (minimum temperature in the coldest month), bio16 (precipitation in the wettest quarter), cult_T (LandUse), and soil organic carbon content were the variables that had the most useful information when used in isolation of other predictors (Figure 3.8).

On the contrary, the results also showed that cult_T, bio6, and bio15 variants decreased the gains most when excluded from the analysis, indicating that these variables contained most of the information not present in any other predictors. Although precipitation seasonality (bio15) provided only a little gain (Figure 3.8A), omitting it decreased the training gain considerably (see the light blue bar compared with the red bar), indicating that it is important for the accuracy of the model. A similar result is illustrated for the jackknife of test gain (Figure 3.8B).

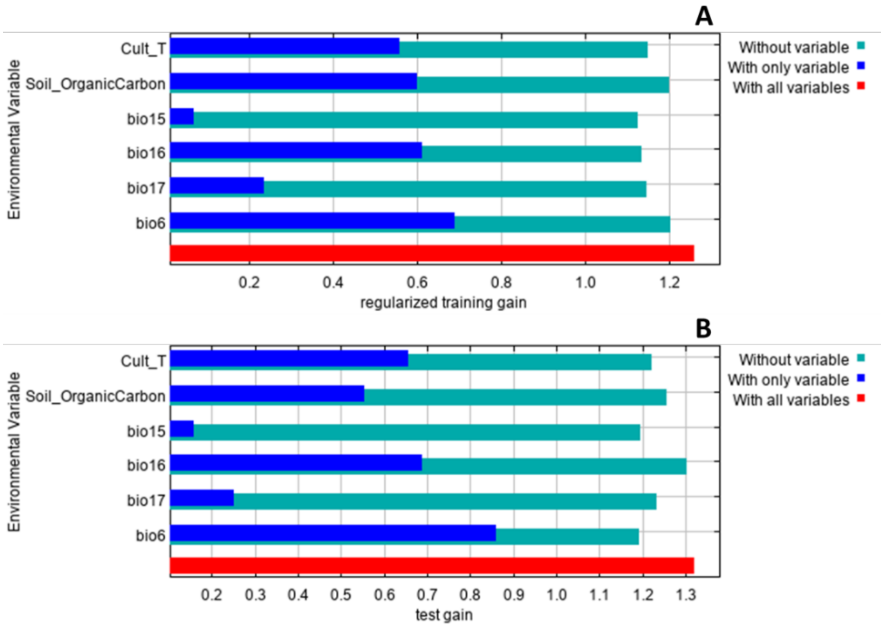


Figure 3.8 Jackknife of regularised (A) training gain and (B) test gain for Ethiopian chicken.

Response curves (Figure 3.9) show how the predicted probability of presence (model predictive power) varies as each environmental variable changes, keeping all the other variables at their average sample value. For example, the response curve for bio6 (minTemp) and LandUse shows that the probability of chickens' presence decreases when the temperature and cultivated land proportion rise. Also, when precipitation seasonality (bio15) and soil organic carbon increase, the likelihood of chicken presence increases. The probability of the model for chicken presence distribution drops gradually when the precWQ (bio16) falls until 600 mm/m^2 and then rises again. In contrast, the likelihood of chicken presence increases when precipitation in precDQ (bio17) rises until 90 mm/m^2 and drops steadily until 200 mm/m^2.

In contrast with the response curves, the marginal response curves represent a different model, a MaxEnt model that only uses the corresponding variables. Thus, the plots reflect the dependence of predicted suitability on the selected variable (Figure 3.9). SoilOrgC and bio6 (minTemp) display a similar response for chicken presence between response curves and marginal response curves. Contrastingly, bio15 - precSeasonality (80 – 100 and 140 - 180 mm/m^2), bio16 - precWQ (precipitation in the wettest quarter = 1200 mm/m^2) , bio17 – precDQ (precipitation in the driest quarter = 90 – 100 mm/m^2) and LandUse (38%) show peaks in the marginal response curve where the chicken presence arises for specific values in each variable. This marginal response needs to be cautiously interpreted. It represents the prediction accuracy of presence response to an individual variable without considering its ecological interactions with other biotic and abiotic factors occurring in a real environment.

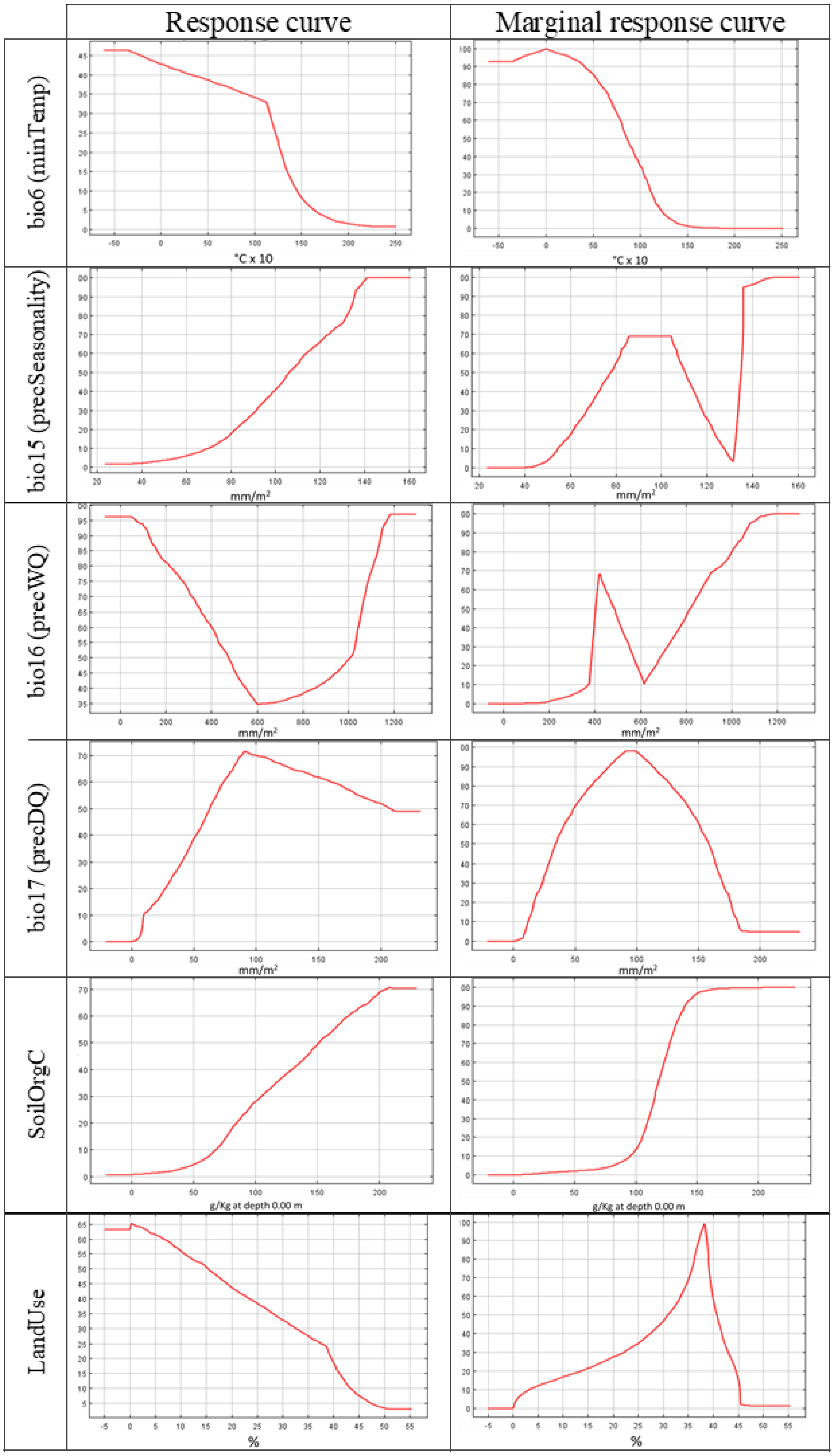


Figure 3.9 Response curves for six selected agro-climatic variables when the ENM is applied across all chicken populations.

The X-axis in each plot shows the environmental values, and Y-axis shows the cumulative output or probability of presence.

Permutation importance can be understood as each variable's contribution, determined by randomly permuting that variable's values among the training points for the presence and background data (Philips, 2017). All selected environmental variables produced high contribution values in terms of permutation importance (> 10%) (Figure 3.10). The three highest contribution values are Soil Organic Carbon (31.2%), Cult_T (18.1%), bio6 (minimum temperature in the coldest month), and bio16 (precipitation in the wettest quarter) (17.4%).

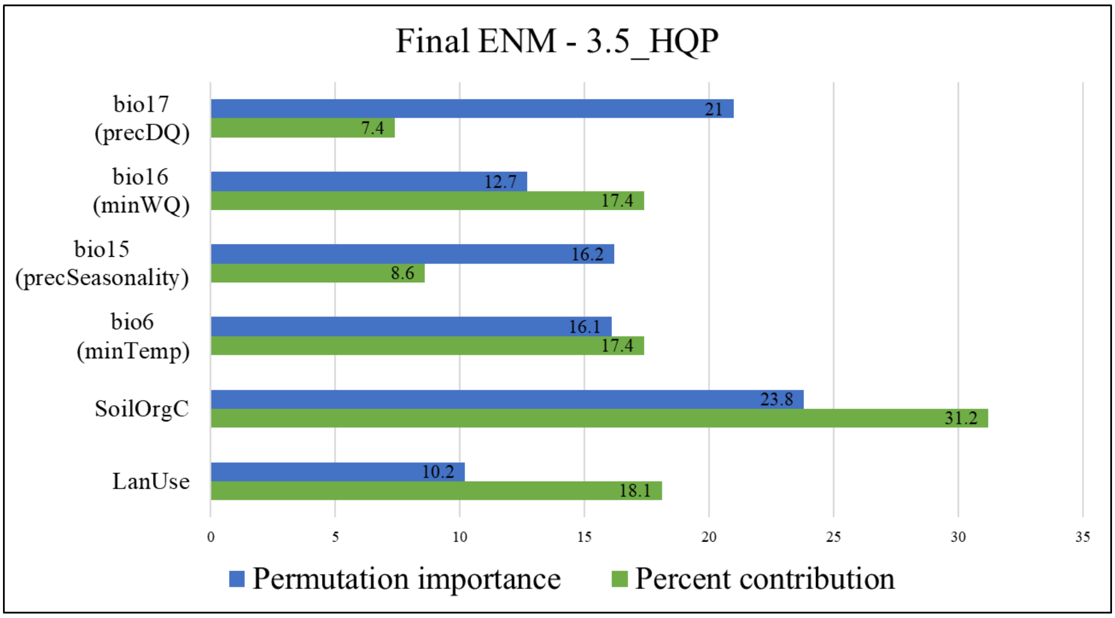


Figure 3.10 Refined estimates of percent contribution (before applying permutation) and permutation importance (% contribution after applying permutation) of six selected environmental predictors from the MaxEnt model.
The model was executed with selected model parameters (feature classes: HQP and regularization parameter: 3.5) and using all (250) presence data points from all populations.

3.3.4 Defining ecotypes

For defining ecotypes, MaxEnt models were first executed for individual populations using the same variables and parameter setting as above. These models predicted a habitat/environmental suitability map of individual populations across the Ethiopian landscape and estimated the relative contribution of environmental variables to each population niche. AUC values of all models were high (> 0.99) for both training and test data (Supplementary Table S3.2). Similarly, the AUC values for the individual environmental variables for each population were generally higher than 0.6, with several exceptions, like bio6 and bio15, which in some populations had values less than 0.6 but over 0.5, which still represent moderate potential.

Pearsons correlation and Niche Overlap statistics were generated by pairwise comparison of the suitability maps for clustering the populations into a similar niche for defining potential ecotypes. The two metrics used for comparing the suitability maps provide somewhat different information. Niche Overlap metric is a measure of similarity of distributions between two populations. In contrast, correlation correlates the underlying models (Warren & Seifert, 2010). The correlation metric is generated by comparing logistic suitability scores (0 or 1) of the two maps' corresponding grid cells.

On the other hand, the Niche Overlap statistic only checks for the proportion of overlap of grid cells that were predicted as suitable (i.e. with logistic suitability score = 1) between two maps. Therefore, it does not include grids with a suitability score of 0 in its calculation. In our study, the results of these two metrics were significantly correlated (Spearman correlation coefficient = 0.61; $P < 0.0001$).

Hierarchical clustering of the populations based on the above two similarity metrics showed different overall topography depending on which statistic was used (Figure 3.11 and 3.12). However, 18 out of 25 populations showed consistent grouping in both analyses. Consistent clusters are (1) Amesha Shinkuri, Gafera, Surta, and Batambie; (2) Ashuda and Dikuli; (3) Bekele Girissa and Shubi Gemo; (4) Gesses and Kido; (5) Kumato and Loya (6) Hugub and Jarso, (7) Hadush Adi and Mihquan. Therefore, we took the most plausible grouping based on at least one test for the rest of the seven populations. A visual inspection of the maps was performed when the tests generated ambiguous results (Meseret, Alfa Midir, Horro, Adane, and Negasi Amba). As a result, three populations – Arabo, Metkilmat, and TzionTeguaz-could not be grouped with other populations and considered separate ecotypes.

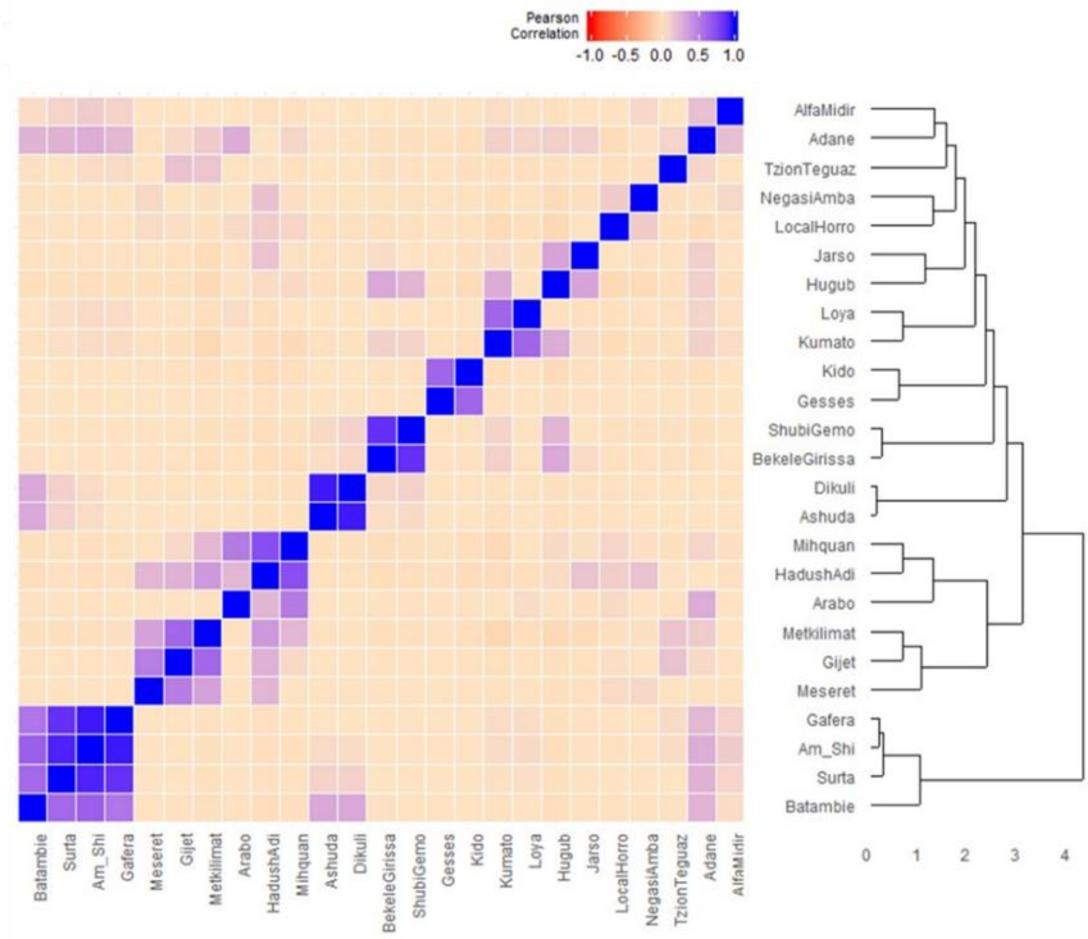


Figure 3.11 Dendrogram and heatmap of pairwise Pearson correlation between suitability maps of individual chicken populations.

Finally, 12 potential ecotypes were identified for the 25 populations examined here (Figure 3.13), reflecting the large heterogeneity of the Ethiopian agro-ecological landscape. (Suitability statistics were generated by pairwise comparison of the ENM suitability maps followed by hierarchical clustering and visual inspection to define the ecotypes).

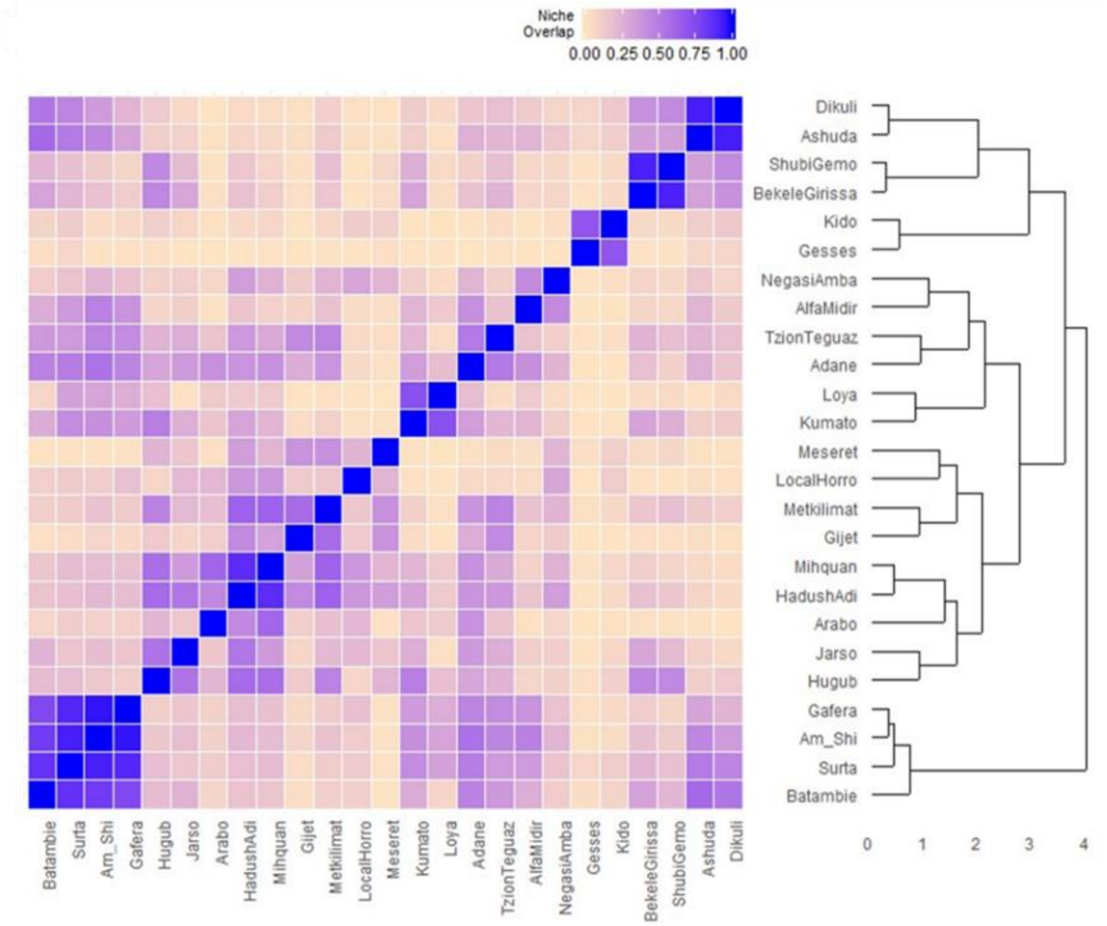


Figure 3.12 Dendrogram and heatmap of Niche Overlap statistic (I) between suitability maps of individual chicken populations.

The most important agro-climatic variables for each ecotype were identified (Table 3.2). These variables are expected to place strong selective pressure on driving adaptation within the ecotype. Interestingly, bio16 (precipitation in the wettest quarter), bio17 (precipitation in the driest quarter), and cult_T (land use) appear to have major contributions in many ecotypes. In contrast, bio15 (precipitation seasonality) and soil organic carbon generate heterogeneity within several ecotypes.

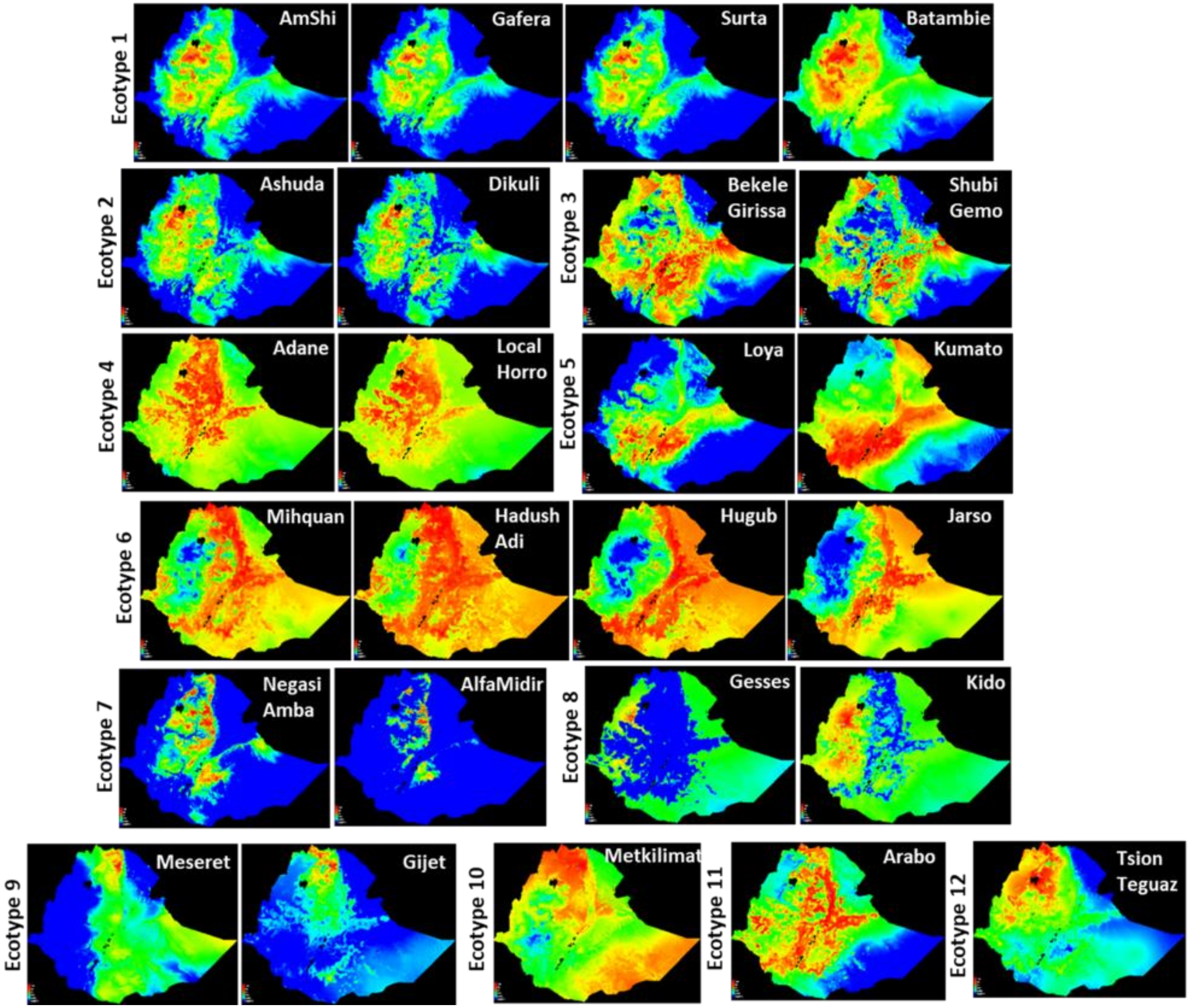


Figure 3.13 The set of 12 potential ecotypes represented by their suitability maps.

Table 3.2 Major contributing environmental variables in each ecotype.
(bio6 (minTemp), bio15 (PrecSeasonality), bio15 (precWQ), bio16 (precDQ), cult_T (LandUse))

Potential ecotype	**Populations**	**Major contribute variables among ecotypes**[1]	**A possible source of differentiation within ecotypes**[2]
E1	AmeshaShinkuri, Gafera, Surta, Batambie	bio6, bio16	bio15
E2	Ashuda, Dikuli	bio16, bio17, cult_T	bio15
E3	BekeleGirissa, ShubiGemo	bio16, bio17, cult_T	Soil organic carbon
E4	Adane, Local Horro	Cult_T, Soil organic carbon	bio6, bio16, bio17
E5	Loya, Kumato	Cult_T, bio15, bio17	bio6, bio16
E6	Mihquan, Hadush Adi, Hugug, Jarso	bio17, cult_T	bio6
E7	Negasi Amba, AlfaMidir	bio6, bio17	bio15, Soil organic carbon
E8	Gesses, Kido	bio16, bio17	bio15, Soil organic carbon
E9	Meseret, Gijet	bio15, bio16, Cult_T	Soil organic carbon
E10	Metkilimat	bio16	
E11	Arabo	bio6, bo16, cult_T	
E12	TzionTeguaz	bio17	

[1] Major contributing variables are those that showed individual AUC > 0.9 for all populations in the same ecotype.
[2] The variables showing the source of differentiation were the ones that showed very diverse individual AUC values for populations within ecotypes.

3.4 Discussion

This study presents an innovative framework of modelling landscape heterogeneity to identify ecotypes of indigenous livestock populations and identify key environmental selection pressures that may have driven local adaptation. Exemplified by Ethiopian village chicken, this novel approach offers a powerful new tool for identifying, characterising, utilising, and conserving livestock diversity.

Choosing Ethiopian indigenous chicken as the study material for ENM had several rationales and benefits. First of all, Ethiopia exhibits great variation in agro-climatic landscapes, predominantly due to extreme altitudinal variation, making it an excellent study landscape for ENM. Moreover, Ethiopian village chickens show diverse morphological and phenotypic characteristics within population (Tadelle *et al.*, 2003), making their classification as ecotypes difficult with their associated phenotypic standards. Therefore, such environmental-modelling-based approach provides a great opportunity for classifying these populations. Besides, these village chicken populations survive in local scavenging or semi-scavenging systems of smallholder backyard farming. Compared to highly productive commercial breeds, they show a higher degree of tolerance to biotic and abiotic stressors in their environment. Therefore, defining ecotypes based on environmental modelling can provide a novel platform for genomic analyses to elucidate local adaptations. The high levels of biological diversity in the Ethiopian chicken populations, along with the diverse environmental conditions of their habitats, make these birds an ideal model to demonstrate the applicability of ENM in livestock species.

An ecotype denotes a genetically adapted population to specific environmental conditions. Following the introduction of domesticated species into new habitats, animals adapt to new environments, leading to new ecotypes (Köhler-Rollefson, 2001). With the distribution of most livestock species across diverse agro-climatic and ecological regions due to human-mediated introduction (Lawler, 2016), livestock ecotypes may also be described as a non-static adaptive variation over many natural landscapes with no discernable boundaries (Lowry, 2012). In practice, researchers and farmers, therefore, face difficulties in identifying and establishing frontiers for ecotypes within the geographic range of a livestock species.

Although previous studies have used different criteria for defining livestock ecotypes (Keambou *et al.,* 2014; Getu & Tadesse, 2014), none addressed the geographic boundaries of ecotypes. Indeed, a detailed characterisation of environment, phenotypes, and genotypes are required to understand these three components' interaction. A comprehensive environmental characterisation of a species habitats may help unravel the clinal variations in genetic and phenotypic traits. With this base, our study developed a framework that livestock researchers and practitioners can apply for defining potential ecotypes by directly modelling important environmental predictors that may drive adaptive evolution.

Our starting point was identifying the landscape variable structures relevant to the formation of Ethiopian village chicken ecotypes. The choice of agro-climatic variables for modelling a landscape configuration faces many potential challenges due to non-uniformity and the variety of ecosystems used by a species (François & Waits, 2016), such as chicken, with its wide distribution range among very diverse landscapes (Lawler, 2016). To avoid arbitrary choices as much as possible, we evaluated 34 environmental parameters to characterise the Ethiopian agro-ecological landscape. No strong correlations were observed during the variable selection between temperature and precipitation variables, which are usually positively correlated (Cong & Brady, 2012). This is because Ethiopia's rainfall patterns do not vary drastically with elevation; instead, it changes with geographic location. In contrast, temperature varies directly with elevation (Jury & Funck, 2013). We then followed a rigorous procedure to select environmental predictors by removing highly correlated variables while retaining only high contributing ones.

The six shortlisted variables included soil features, land-use, and climatic parameters, which resulted in a good prediction of suitable habitats of Ethiopian chicken populations. The same or similar variables were also used in a previous study by (Pitt *et al.,* 2016) for predicting the suitable environmental conditions of early domestic fowl at a global geographic scale based on niche occupied by extant wild Red Junglefowl (*Gallus gallus*). Their study, however, also included bio4 (temperature seasonality) and bio5 (maximum temperature of the warmest month) in their model. These variables were not included in our study either because they showed a strong correlation with one or more variables or had little contribution in explaining the overall model.

We also decided to optimise the best model parameters (RM and FCs) to achieve the best predictive power for our Ethiopian data. We demonstrated that ENM based on selected parameters had a discernible impact on predicting habitat distribution compared to the default setting (Figure 3.6). Model parameters were optimised based on combined analysis of all populations and subsequently applied to individual population analyses. No attempt was made to select the best parameters for a particular population as the data points per population were too small for a meaningful optimisation. Besides, using the same parameters for all populations provided a common ground for comparing suitability maps for ecotype definition.

For ecological niche modelling or species distribution modelling (SDM), it is generally a requirement that the presence data are not spatially autocorrelated and are from a random or representative sampling of environmental space occupied by the focal species and (Elith *et al.,* 2006; Yackulic *et al.,* 2013). The presence of spatially autocorrelated data points would generate a biased inference in the SDM, as these would lay more weight on certain types of habitats as suitable. In our study, however, the sampling points within populations were not uncorrelated. This is because the geographic range of each population was very small, as samples from individual villages were considered a population, and for each population, we included 10 data points. However, we believe that such a sampling bias is not an issue for the context of the present study, but rather it was necessary. This is because the goal of our study was not to predict the distribution of a species but to classify the investigated populations into similar niche categories for defining ecotypes. Therefore, to capture each population's environmental condition via modelling, we had to take several sampling points from each population (even though they were autocorrelated).

3.5 Conclusions

The novel approach presented here demonstrates how environmental modelling leads to livestock ecotype characterisation and definition, providing a baseline for environment-genotype association studies (Chapter 5) and better management practices. Genomic analysis of these ecotypes will allow dissecting the genetic basis of local adaptation. Such knowledge will be invaluable for breeding improvement initiatives through marker-assisted selection or introgression, genomic selection, and genome editing. Incorporating "adaptive capacity" in highly productive breeds will be of great value to ensure sustainable livestock production. Moreover, habitat suitability mapping can be an effective tool for deciding which areas have the best environmental condition to introduce a new livestock breed for farming or determine the most suitable population for agriculture in a given agro-climatic scenario. The framework presented here should be applicable to any livestock species and would be able to incorporate any relevant environmental data (e.g. epidemiological data), promising a much wider implication.

Chapter 4

Whole-genome sequences reveal the drivers of environmental adaptation in Ethiopian village chicken

Abstract

Worldwide, climate change affects agriculture systems, producers' welfare, and food availability for the growing human population. Consequently, different mitigation strategies need to be put in place. Pinpointing environmental stressors at the local level and identifying genomic regions that may strengthen the resilience of individuals against these ecological stress factors provide important entry points for devising mitigating measures. Using the six major agro-climatic variables identified by ENM (in Chapter 3), this chapter investigates the genomic signatures of adaptation to these environmental stressors in Ethiopian village chicken populations. One temperature variable (strongly correlated with elevation), three precipitation variables (related to water availability), and two soil/land variables (linked to food availability for scavenger chickens) were analysed. Between-group genomic analyses (*XP-EHH* and *Fst*) were performed on populations from extreme environments (High *vs* Low) for each environmental variable separately. The analyses identified candidate genomic regions and genes under positive selection for environmental challenges related to high altitude and extreme temperature, water scarcity, and food availability. The results presented here illustrate the importance of identifying key agro-climatic drivers prior to genomic analysis for environmental adaptation, while the detected candidate genes offer valuable targets for the improvement of adaptive performance through breeding programmes.

4.1 Introduction

Nowadays, livestock increasingly undergoes environmental stresses under changing climate conditions (Niyas *et al.*, 2015). Climate change has turned out to be a global phenomenon that threatens the survival of many ecosystems, increasing the temperature and leading to extreme precipitation and fluctuation of weather conditions (Stocker *et al.,* 2013). The agriculture sector is one of the most affected sectors of extreme climates; firstly, by challenging the physiological coping mechanism of livestock species (e.g. temperature stress) (Afsal *et al.*, 2018) and, secondly, by impacting the welfare and food security of smallholder farmers and rural populations (Makuvaru *et al.,* 2018; Dil Bahadur & Akhter, 2018).

Demand for livestock products is rising due to the growing human population (Rashamol & Sejian, 2018). Thus, a sustainable livestock production system is imperative to respond to market demands. One of the possible avenues is developing and utilising more climate-resilient breeds (Karimi *et al.*, 2018).

Compared to exotic breeds, indigenous livestock populations display a better adaptation to their local environmental conditions (Rashamol & Sejian, 2018). These populations represent a valuable source of adaptive genetic diversity (Velado-Alonso, 2020). More than any other livestock species, domestic chicken show a wide range of environmental resilience. They are found in the most extreme environments, from the hottest (e.g. an oasis in the desert) to the coldest (e.g. Tibetan plateau) (Lawler, 2016).

Local breeds of chicken in Africa are adapted to a broad tropical environmental diversity. Since the early human-mediated chicken migration occurred worldwide, different environmental challenges have modulated their adaptation response (Lawler, 2016). For instance, chickens can naturally control their core body temperature homeostatically, but regulation works within a narrow range (Abioja & Abioma, 2021). Today, however, chickens are found ubiquitously worldwide and across very diverse temperature zones, ranging from hot and arid tropical climates (24 – 36 °C) to very cold temperate regions (below 0 °C). Long exposures to such environmental stressors are the driving force for natural selection to occur to achieve a lasting adaptation. Therefore, investigating local adaptations can provide valuable insight into their genetic basis. Such knowledge would be invaluable in combating rapid climate change impacts on livestock populations and creating resilient breeds through genetic interventions. This example, plus other environmental stressors *vs* adaptive responses in chickens, points out that the indigenous village populations – exposed to various environmental stressors - are an interesting focal resource to scrutinise local adaptation.

The Ethiopian agricultural landscape displays a combination of rainfall patterns and temperature at the root of a large diversity of agro-climatic conditions (Fazzini *et al.*, 2015), which vary from hot-humid and hot-arid to cold-humid and cold-arid (Mangistu, 2006). It provides us with a unique opportunity for untangling the genetic drivers of environmental adaptation.

Even though African indigenous chickens are adapted to their harsh environmental conditions, the genetic mechanisms behind these adaptations are still unclear (Muchadeyi & Dzomba, 2017). Until now, studies related to environment-genome adaptation in chicken have focused on specific environmental stressors, such as elevation or temperature stress (e.g. Zhang *et al.*, 2016a; Cedraz *et al.*, 2017). In this chapter, we used the six agro-climatic variables defined in Chapter 3 to identify genomic regions under positive natural selection as an adaptive response to these environmental stressors in the Ethiopian indigenous village chicken populations. We used a comparative genomic approach by 1) ranking the populations for each of the six environmental parameters, 2) identifying extreme groups of populations, and 3) comparing and contrasting the genome differentiation of these two extreme groups. Genomic regions affected by positive selection show long-range linkage disequilibrium (LD) and high population frequency. Hence, we applied two cross-population methods to detect signatures of selection: the fixation index (*Fst)* and the cross-population extended haplotype homozygosity index (*XP-EHH). Fst* statistic (Weir & Cockerham, 1984) measures population differentiation, resulting in differences in population genetic structure following distinct demographic or selection histories. Large *Fst* values between the populations at specific genome regions will represent candidate regions under selection (Biswas & Akey, 2006). The *XP-EHH* is based on the calculation of the long-range haplotype (*LRH*) length around each marker (Voight *et al.,* 2006). It identifies selection signals in cross-population comparison. The *XP-EHH* index was explicitly designed for the signature of positive selection analysis. Based on a linkage disequilibrium calculation approach, it is assumed to give lower false positives than *Fst* (Ma *et al.,* 2015). It also indicates which population experienced selection (Diao *et al.,* 2019). These two methods were chosen based on two main reasons. First, we wanted to have one allele frequency-based approach and on LD-based approach, and second, both methods capture different selection signals and can serve as validation for each other. Therefore, in this study, genomic regions detected by both methods simultaneously were treated as the strongest candidates of selection signatures.

4.2 Material and Methods

4.2.1 Environmental data and definition of extreme groups of populations

The details about the selection of agro-climatic variables can be found in section 3.1 (Chapter 3). Supplementary Table S4.1 shows the data from the six selected variables in each population. The environmental variables are denoted as: the minimum temperature of the coldest month of a year (minTemp), precipitation seasonality (PrecSeasonality) representing the variation in precipitation across a year, precipitation in the wettest quarter of the year (precWQ), precipitation in the driest quarter of a year (precDQ), soil organic carbon content (SoilOrgC) and proportion of cultivated land (LandUse).

After ranking the 25 populations for each environmental parameter from low to high, two populations from each end of the gradations were selected. These low and high groups of populations for the six agro-climatic variables were used to perform the selection signature analyses.

4.2.2 Genomic data

Sequencing data information for 223 out of the 243 samples are available in Supplementary Table S2.1 (Chapter 2). Twenty additional samples from Horro and Jarso populations were accessed from a previous study (Lawal *et al.*, 2018). These 243 samples belong to 25 different Ethiopian chicken populations (Table 2.1 and Figure 2.1 in Chapter 2; Table 3.1 in Chapter 3).

4.2.3 Selection signature analyses

Two methods of detecting signatures of positive selection – *XP-EHH* and *Fst* - were applied over the extreme low and high groups for each environmental variable. The filtered autosomal SNP set described in Chapter 2 was used for these analyses. In both cases, analyses were performed in overlapping 20-kb windows with a 10-kb sliding step. Windows with only SNPs ≥ 10 were retained for further analyses. For both detection methods, empirical *P*-values were calculated by first sorting the scores in descending order (i.e. by ranking the windows) and then dividing their ranks by the total number of windows.

4.2.4 *Fst* analysis

Fst analyses (Weir & Cockerham, 1984) were performed using VCFtools version 0.1.11 (Danacek *et al.*, 2011). Following formula 4.1, *Fst* values were calculated for each SNP as described in Akey et al., (2002), which compare differences in allele frequencies between populations.

$$(4.1) \qquad Fst = P_{i\,total} - P_{i\,within}$$

Where $P_{i\,within}$ and P_i are defined following formulas 4.2 and 4.3.

$$(4.2) \qquad P_{i\,within} = \frac{P_{(i)\,population\ 1} + P_{(i)\,population\ 2}}{2}$$

$$(4.3) \qquad P_i = 1 - fA2 - fT2 - fC2 - fG2$$

With fN being the frequency of nucleotide N (A, T, C, or G), and $P_{i\,total}$ is the total P_i for which allele frequencies in both populations are averaged. The weighted *Fst* values were standardised (*ZFst*) to allow using the same threshold score for each environmental parameter as follows (formula 4.4).

$$(4.4) \quad ZFst = \frac{Fst - \mu Fst}{\sigma Fst}$$

Windows with only SNPs ≥ 10, empirical *P*-value ≤ 0.01, and *Fst* (*ZFst*) > 5 were considered potential sweeps.

4.2.5 XP-EHH

The *XP-EHH* analyses were performed using the Hapbin package (Mclean *et al.*, 2015). As the Hapbin software requires the genetic positions (along with physical position) of variants as input, these were calculated using chromosome-specific recombination rates (Groenen *et al.*, 2009). For chr16, we used the recombination rates of Elferink *et al.* (2010). The recombination rates are based on older reference genome and not GRCg6a assembly. To avoid any issues produced by that, we used “mean” recombination rates for specific chromosomes (calculated by dividing cM length of chromosome by physical length in Mb) for estimating the genetic distance of the SNPs. However, Groenen *et al*., (2009) do not have information on Chr 30 – 33. Then, for these chromosomes, we took the average recombination rate of all microchromosomes (5.4 Cm/Mb).

Missing genotypes were removed, and phasing was performed using Beagle v5.1 (Browning & Browning, 2007). *XP-EHH* values were first calculated for individual SNPs with two parameters: cut-off value for Extended Haplotype Homozygosity (*EHH*) is 0.1, and minor allele frequency > 5%. The calculation was performed according to Sabeti *et al.* (2007) using the formula (4.5), where (s) represent the given marker to be calculated, and *iES* represent the integrated *EHH* (average length of shared haplotypes) in populations 1 and 2, which are being contrasted.

$$(4.5) \qquad unXP\text{-}EHH\ (s) = \ln\left(\frac{iES_{pop1}(s)}{iES_{pop2}(s)}\right)$$

For a given marker (s), *XP-EHH_std* is defined as the standardized *unXP-EHH* in formula (4.6), where mean (*unXP-EHH*) and sd(*unXP-EHH*) represent the mean and standard deviation of *unXP-EHH* computed over SNPs with similar allele frequency.

$$(4.6) \qquad XP\text{-}EHH\_std(s) = \frac{unXP\text{-}EHH(s) - mean(unXP\text{-}EHH)}{sd(unXP\text{-}EHH)}$$

Mean values were calculated within windows for (i) standardised *XP-EHH* (*XP-EHH_std*) and (ii) the absolute value of standardised *XP-EHH* (|*XP-EHH_std*|*)*. Threshold values were defined as follows: windows with only SNPs ≥ 10, empirical *P*-value ≤ 0.01, and absolute standardised *XP-EHH (*|*XP-EHH*|*)* > 3. From these, windows with SNPs showing the same directionality of selection were considered as candidate signatures of selection windows. Positive *XP-EHH* values indicated selection in the low group and negative values selection in the high group.

Moreover, pooled heterozygosity (*Hp*) (Rubin *et al.,* 2010) value for candidate selected windows was calculated (see details in Material and Methods section of Chapter 5) for the low and high groups separately to confirm the directionality of selection. In particular, *Hp* results were referred to when *XP-EHH* results were ambiguous regarding the directionality of selection. Windows were removed from the subsequent analysis when the directionality could still not be resolved.

4.2.6 Functional annotation

Identified selection signature regions (SSRs) were intersected with known genes from Ensembl (release 98) using Bedtools v2.26 (Quinlan & Kindlon, 2017). Candidate genes were checked for their overlap with known chicken QTLs to find a possible and suggestive phenotypic association (Hu *et al.*, 2014). Only significant QTLs with size < 1 Mb were considered.

4.3 Results

4.3.1 Definition of Low and High population groups

The first step towards selection signature analyses in relation to the six ENM-identified environmental variables was to decide which populations to compare. Figure 4.1 shows the ranking of the 25 populations from low to high for each of the six environmental variables.

For each environmental variable, two populations from extreme ends of the variable distribution were chosen. The minimum temperature of the coldest month of the year (minTemp) is highly correlated with the maximum temperature of the warmest month (bio5 - maxTemp) and the elevation (Figure 3.3C in Chapter 3). Accordingly, AlfaMidir and NegasiAmba are the two extreme populations in the Low group, but for the High group, we selected Mihquan and Hugub. Mihquan, instead of Kido, was selected as the maximum temperature of the warmest month (bio5 - maxTemp) for Mihquan is higher than that of Kido, while other correlated variables (i.e. minTemp and elevation) show similar values.

For the precipitation seasonality (precSeasonality), the Low group includes Loya and Kumato, and the High group Meseret and Gijet (Figure 4.1B). Hugub and Jarso are the two extreme Low group populations for the precipitation of the wettest quarter (precWQ) parameter, with Gafera and Gesses defining the High group (Figure 4.1C). Precipitation of the driest quarter (precDQ) has Gijet and Kido in the Low group and Kumato and Loya in the High group (Figure 4.1D). Soil Organic Carbon (SoilOrgC) has Loya and Kumato in the Low group and AlfaMidir and Adane in the High group (Figure 4.1E). Finally, LandUse has Gesses and Kido in the Low and Meseret and AlfaMidir in the High group (Figure 4.1F).

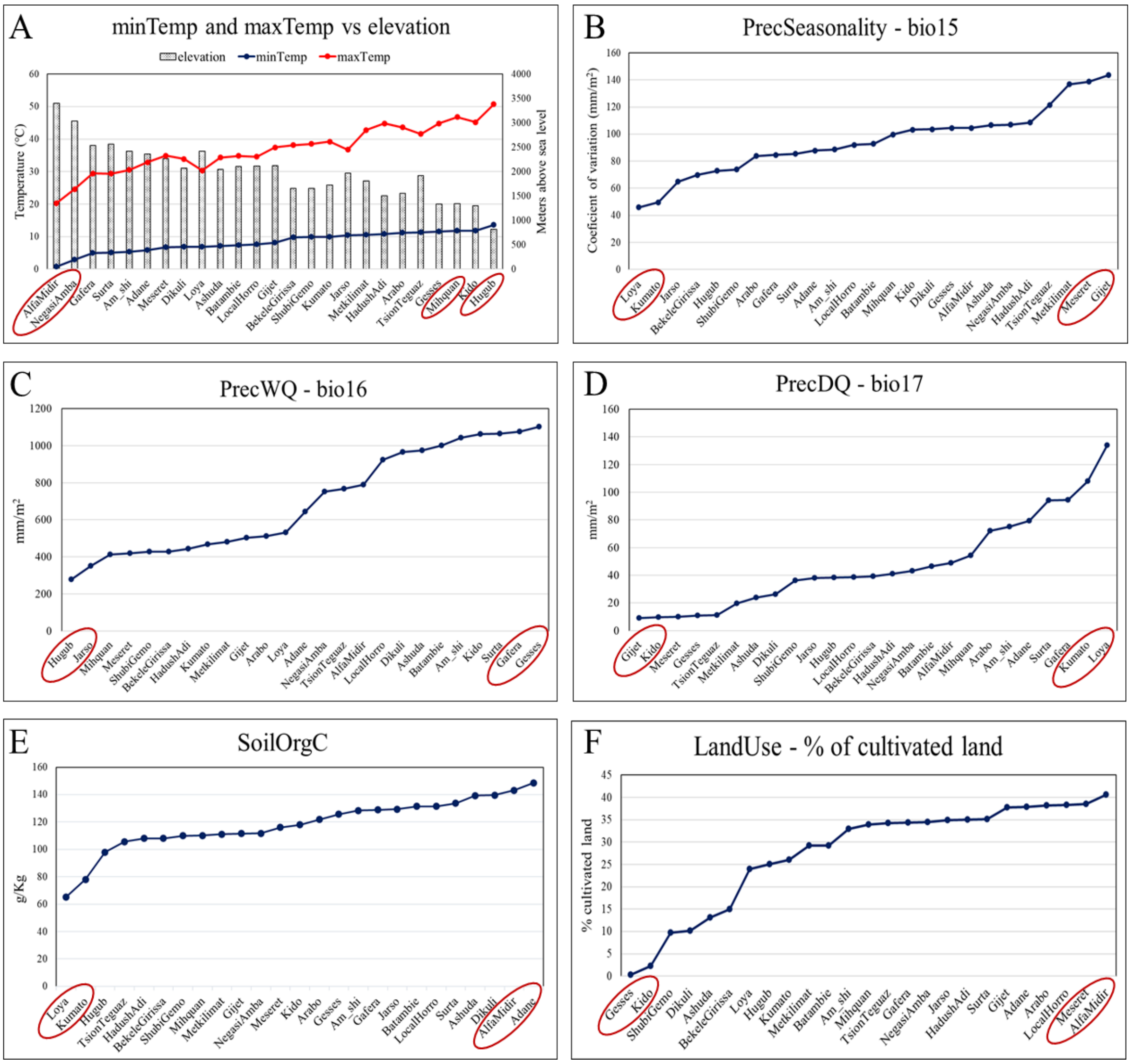


Figure 4.1 Populations organised based on environmental gradients for the six key environmental parameters identified by the ENM approach (Chapter 3).

4.3.2 Identification of candidate windows under selection for environmental adaptation

Candidate windows of selection signatures were identified based on the thresholds defined in the Material and Methods section. The full results are presented in Supplementary Tables S.4.2a to S4.2f. Supplementary Figures S4.1 and S4.2 show the observed distributions of *ZFst* and |*XP-EHH*| variables for each environmental analysis. Supplementary Figure S4.3 shows only a weak positive correlation between *XP-EHH* and *Fst* values (r_S= 0.22 - 0.34, P < 2.2e-16). Therefore, the following sub-sections summarise first the *ZFst* and |*XPEHH*| results and then describe the strongest candidate regions (including genes and variants) related to each environmental parameter. Included in the strongest candidates are genes that were detected by both the *Fst* and *XP-EHH* approaches.

4.3.2.1 Overview of *Fst* analysis

Fst values were calculated for 93,772 - 93,844 windows in different environmental analyses (Table 4.1). The genome-wide mean values of *Fst* between the Low and High groups in each environmental analysis were very low (min: 0.021 ± 0.03 for LandUse, and max: 0.072 ± 0.06 for precWQ), indicating low overall genetic differentiation between the groups. Between 0.1% and 0.3% of the analysed windows passed the set threshold to be considered as candidate signatures of selection. The lowest number of candidate windows were detected in precWQ analysis, while the highest number was detected in precDQ analysis (Table 4.1). After merging overlapping or adjacent candidate windows, 334 sweep regions were identified across the environmental variable groups. The number of regions ranged from 2 (for minTemp High group) and 52 (for SoilOrgC High group).

Although we have limited number of samples (e.g. 20 samples per extreme population groups), the selection signature analysis uses dense genomic data based on SNPs. Also, we have not used individual SNPs for comparison but combined SNP information from within windows. Small sample size is a problem when you are comparing individual SNPs, but previous studies have shown that when dense data from WGS is available, as small as 10-15 samples are good for performing selection signature analysis (Ma *et* al., 2015). In our case we used 20 kb windows and ensured at least 10 SNPs are present per window, but most windows had larger number of SNPs. On the other hand, by using

20 kb window size and dense SNP data, we have got reasonably good resolution compared to any QTL analysis (which are very big, often mega base long).

The mean *Fst* values in the candidate windows varied from 0.22 in LandUse to 0.44 in precWQ. These values are between 6 to 10 times higher than the genome-wide *Fst* average (Table 4.1). The highest *ZFst* scores from different analyses are: minTemp (*ZFst* = 8.16, chr14:3930000-3950000) (Figure 4.2a), precSeasonality (*ZFst* = 12.16, chr2:95770000-95790000) (Figure 4.2b, precWQ (*ZFst*= 7.91, chr4:39340000-39360000) (Figure 4.2c), precDQ (*ZFst* = 10.42, chr3:73130000-73150000) (Figure 4.2d), SoilOrgC (*ZFst*= 11.70, chr33:6470000-6490000) (Figure 4.2e), and LandUse (*ZFst*= 11.75, chr2: 27950000-27970000) (Figure 4.2f).

Table 4.1 Genome-wide Fst statistics for extreme population groups (Low and High) for the six key environmental variables.

Environmental variable	**Total number of windows *Fst***	**Mean ± SD (Total *Fst* windows)**	**Candidate windows (*ZFst*)[a]**	**Mean ± SD (Candidate windows)**	**Group[b]**	**Number of (*ZFst*) sweep regions[c]**
minTemp	93811	*Fst*: 0.04 ± 0.04	90	*Fst*: 0.32 ± 0.03	Low	32
					High	2
precSeasonality	93844	*Fst*: 0.02 ± 0.03	218	*Fst*: 0.25 ± 0.05	Low	26
					High	33
precWQ	93829	*Fst*: 0.07 ± 0.06	88	*Fst*: 0.44 ± 0.03	Low	17
					High	12
precDQ	93772	*Fst*: 0.03 ± 0.06	280	*Fst*: 0.26 ± 0.03	Low	49
					High	22
SoilOrgC	93825	*Fst*: 0.03 ± 0.04	165	*Fst*: 0.27 ± 0.04	Low	13
					High	52
LandUse	93830	*Fst*: 0.02 ± 0.03	254	*Fst*: 0.22 ± 0.04	Low	51
					High	25

(a) Number of candidate (ZFst) windows retained with threshold: SNP ≥ 10, *P*-value ≤ 0.01, and (ZFst) ≥ 5.0.
(b) Low and High population groups are defined by the environmental variable trend across the studied Ethiopian chicken populations.
(c) Sweep regions were determined after merging overlapping or adjacent windows.

Of the 334 *Fst* sweep regions, 153 (46%) have size of 20 kb (i.e. length of a window) while 25 regions (7%) have size larger than 100 kb. The largest candidate regions for each variable are: 100 kb for minTemp (chr4:43410000–43510000), 110 kb for precSeasonality (chr3:85990000–86100000; chr4:74400000-74510000; chr5:40460000-40570000), 240 kb for precWQ (chr8:10430000–10670000), 130 kb for precDQ (chr2:95670000–95800000), 110 kb for SoilOrgC (chr4:50860000–50970000), and 140 kb for LandUse (chr2:82630000-82770000) (Supplementary Tables S4.2.a – f).

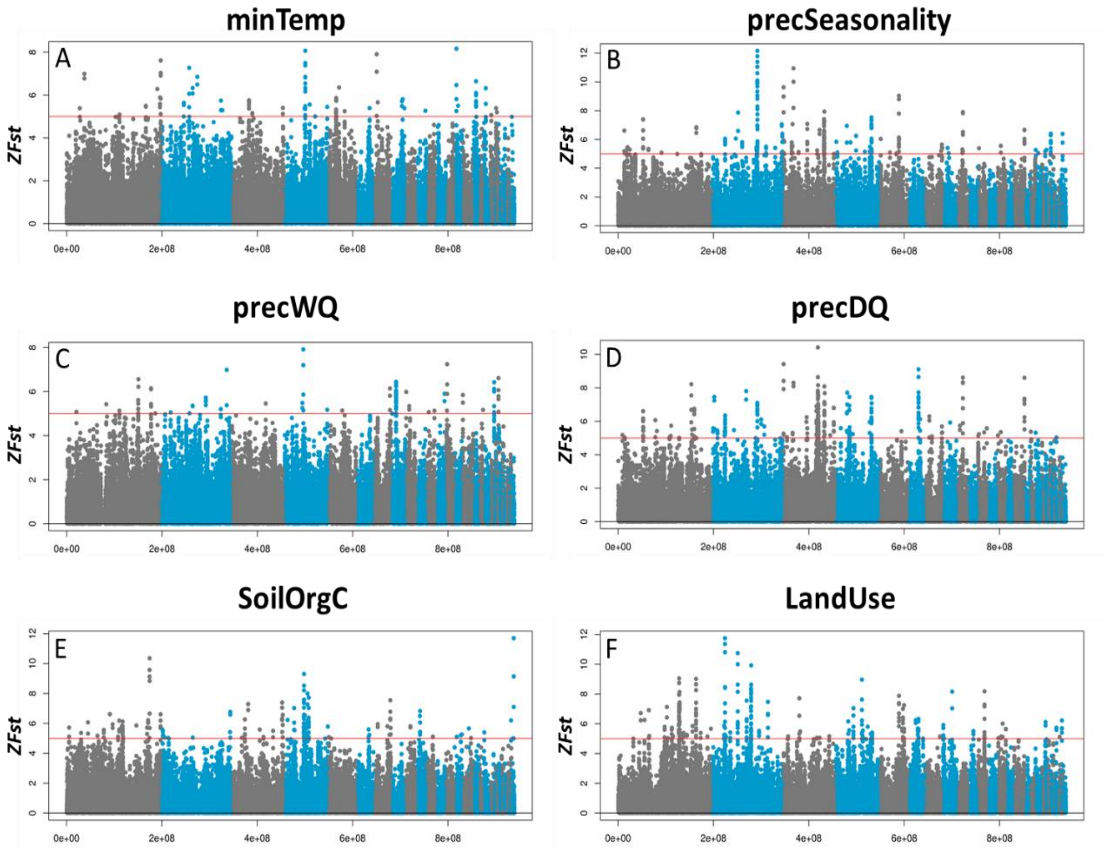


Figure 4.2 *ZFst* candidate signatures of positive selection in Ethiopian village chicken based on six environmental variables.

The red line represents the threshold (*ZFst* = 5). All windows above the threshold were considered as candidates. The X-axis represents the genomic positions within 20 kb windows along autosomes.

4.3.2.2 Overview of *XP-EHH* analyses

Between 92,405 to 92,693 windows were analysed for *XP-EHH* across the six environmental analyses (Table 4.2). The genome-wise mean values of *XP-EHH*, between the Low and High groups in each environmental analysis, were close to 0 (min: -0.06 ± 0.31 for precSeasonality, and max: 0.29 ± 0.37 for minTemp), indicating low overall haplotype-genetic differentiation between groups. Between 0.2% and 0.4% of the windows passed the set threshold to be considered as candidate signatures of selection. The lowest number of candidate windows were detected in precWQ analysis, whereas the highest number was detected in LandUse analysis (Table 4.2). After merging candidate windows, 472 sweep regions were identified across the environmental variable groups. The number of regions ranged from 21 (for SoilOrgC Low group) and 57 (for minTemp Low group).

Table 4.2 Genome-wide (|*XP-EHH*|) statistics for extreme population groups (Low and High) for the six key environmental variables.

Environmental variable	**Total number of windows *XP-EHH***	**Mean ± SD (Total *XP-EHH* windows)**	**Candidate windows (|*XP-EHH*|)[a]**	**Mean ± SD (candidate windows)**	**Group[b]**	**Number of (|*XP-EHH*|) sweep regions[c]**
minTemp	92592	*XP-EHH:* 0.29 ± 0.37	292	*(|XP-EHHstd|):* 1.10 ± 1.11	Low	57
					High	18
precSeasonality	92683	*XP-EHH:* -0.06 ± 0.31	301	*(|XP-EHHstd|):* -0.37 ± 1.11	Low	41
					High	46
precWQ	92405	*XP-EHH:* 0.08 ± 0.42	210	*(|XP-EHHstd|):* 0.47 ± 1.40	Low	52
					High	25
precDQ	92693	*XP-EHH:* 0.05 ± 0.42	285	*(|XP-EHHstd|):* 0.26 ± 1.60	Low	52
					High	33
SoilOrgC	92656	*XP-EHH:* 0.25 ± 0.34	225	*(|XP-EHHstd|):* 1.01 ± 1.04	Low	21
					High	55
LandUse	92635	*XP-EHH:* 0.06 ± 0.47	404	*(|XP-EHHstd|):* -0.08 ± 1.70	Low	47
					High	25

(a) Numbers of candidate (|*XP-EHH*|) windows retained with threshold: SNP ≥ 10, *P*-value ≤ 0.01 and (|*XP-EHH*|) ≥ 3.0.
(b) Low and High population groups are defined by the environmental variable trend across the studied Ethiopian chicken populations.
(c) Sweep regions were determined after merging overlapping or adjacent windows.

The mean |*XP-EHHstd*| values in the candidate windows varied from 1.11 (minTemp and precSeasonality) to 1.7 (LandUse). These values are between 4 to 6 times higher than the genome-wide average (Table 4.2). The highest |*XP-EHHstd*| scores for the different analyses are: minTemp (|*XP-EHHstd*| = 4.50, chr1:197340000-197360000) (Figure 4.3a), precSeasonality (|*XP-EHHstd*| = 4.41, chr7:11550000-11570000) (Figure 4.3b), precWQ (|*XP-EHHstd*| = 4.14, chr4:25510000-25530000) (Figure 4.3c), precDQ (|*XP-EHHstd*| = 5.03, chr1:136580000-136600000) (Figure 4.3d), SoilOrgC (|*XP-EHHstd*| = 5.52, chr1:197490000-197510000) (Figure 4.3e), and LandUse (|*XP-EHHstd*| = 4.84, chr16:620000-640000) (Figure 4.3f).

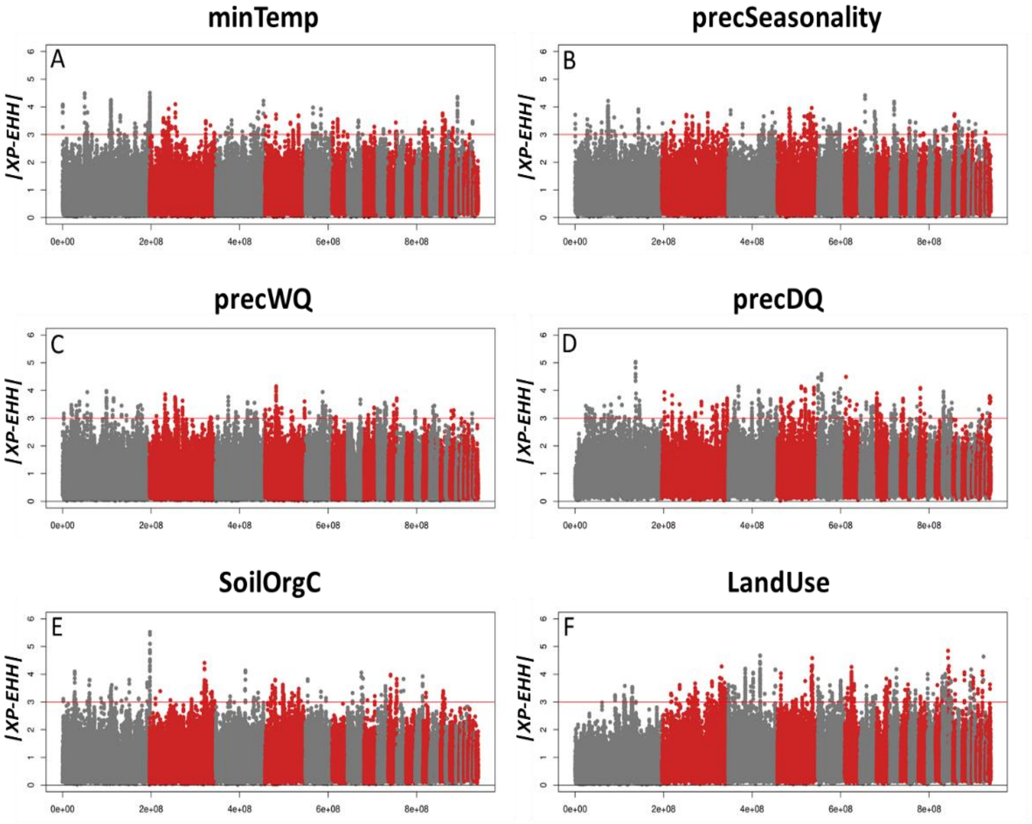


Figure 4.3 (|*XP-EHHstd*|) candidate signatures of positive selection in Ethiopian village chicken based on the six environmental variables.

The red line represents the threshold (|*XP-EHHstd*| = 3). All windows above the threshold were considered as candidates. The X-axis represents the genomic positions within 20 kb windows along autosomes.

One hundred sixty-four regions (35%) are 20 kb in size whereas 38 (8%) are of sizes larger than 100 kb. For each variable, the size of the largest sweeps regions is: 230 kb for minTemp (chr4:63860000-64090000), 290 kb for precSeasonality (ch5:48710000-49000000), 160 kb for precWQ (chr5:39190000-39350000), 130 kb for precDQ (chr3:70960000-71090000), 100 kb for SoilOrgC (ch1:197480000-197580000) and 550 kb for LandUse (chr4:25890000-26440000) (Supplementary Tables S4.2.a – f).

4.3.3 Candidate genes and variants in relation to different environmental adaptation

The number of candidates selected genome regions was determined for each low and high environmental parameter. Regions discovered by both methods were considered as our strongest candidates. These include entire regions or regions partially detected by both approaches (Figure 4.4).

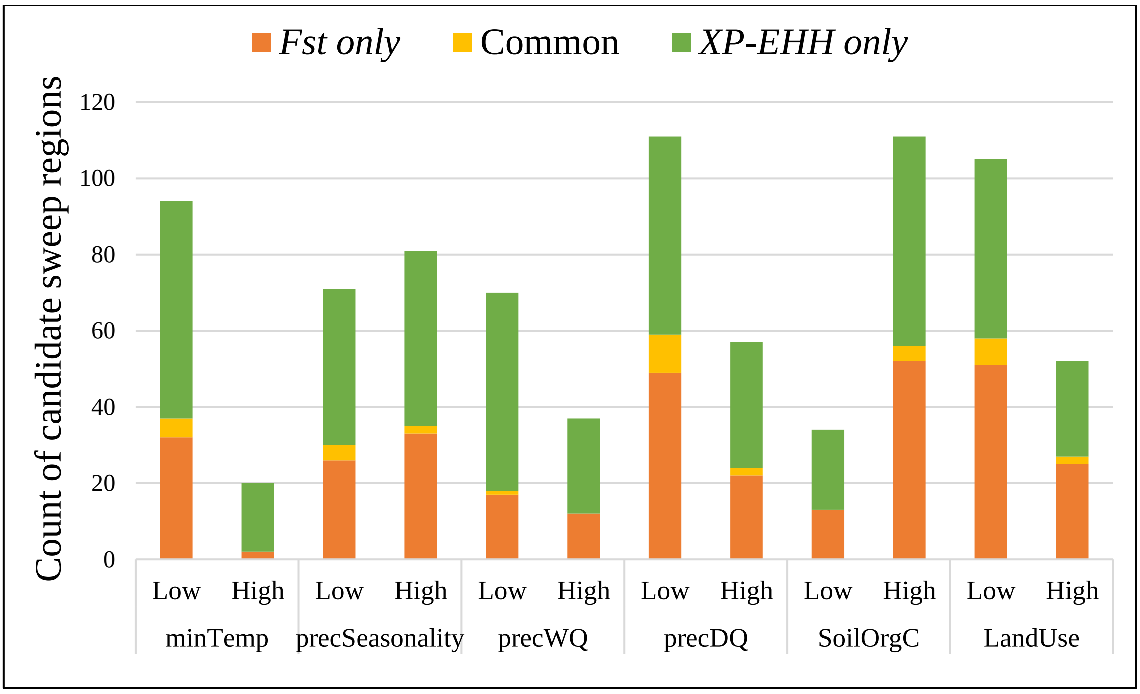


Figure 4.4 Stacked bar plot showing the candidate sweep regions based on Low/High groups and positive selection methods.

Around 76% to 90% of the detected regions overlap with genes. The number of overlapping genes from different analyses is 155 from minTemp, 167 from precSeasonality, 131 from precWQ, 183 from precDQ, 174 from SoilOrgC, and 164 from LandUse (Supplementary Tables 4S.2.a – f). The candidate regions harbour between ~62,000 to ~101,000 SNPs, but only 1.5% to 8% show a large difference in allele frequency (dAAF > 0.5) between the High and Low population groups (Supplementary Tables S4.3.a – f). These SNPs with high allele frequency differences were annotated. Most of these SNPs are intronic (51% out of 597 annotated SNPs). Only a few of these SNPs (between 1 to 35) were annotated into a potentially functional category (non-synonymous, splicing, and ncRNA exonic) (Figure 4.5).

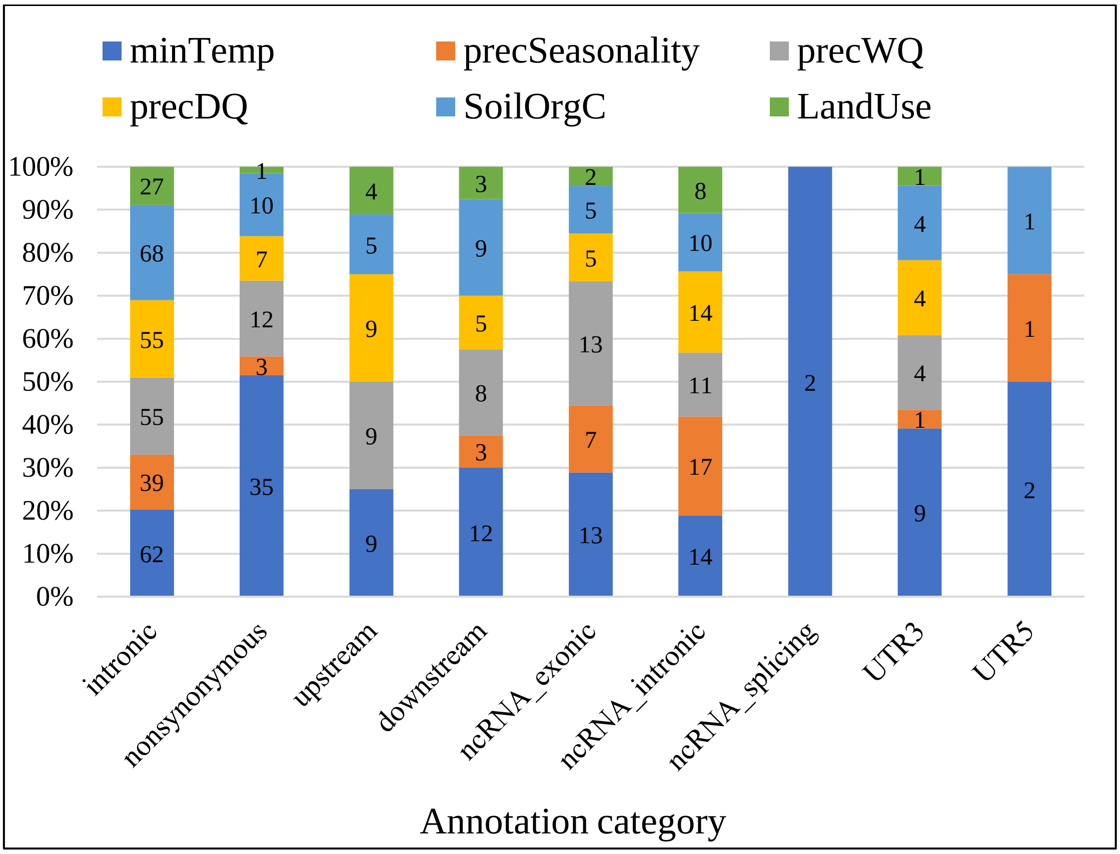


Figure 4.5 Barplots showing the number and their annotation categories of SNPs with a large difference in allele frequency (dAAF > 0.5) between the Low and High groups for different environmental analyses.
Numbers within bars represent the genes count for each category.

Candidate genes were investigated for their overlap with known QTLs (ChickenQTLdb). In total, 174 genes out of 603 overlap with QTLs of size < 1 Mb (Supplementary Table S4.4). Genes overlapping with growth-related QTLs were detected predominantly in all environmental groups except SoilOrgC-Low (Figure 4.6). Overall, QTLs related to growth, egg property, and immune response were more commonly detected.

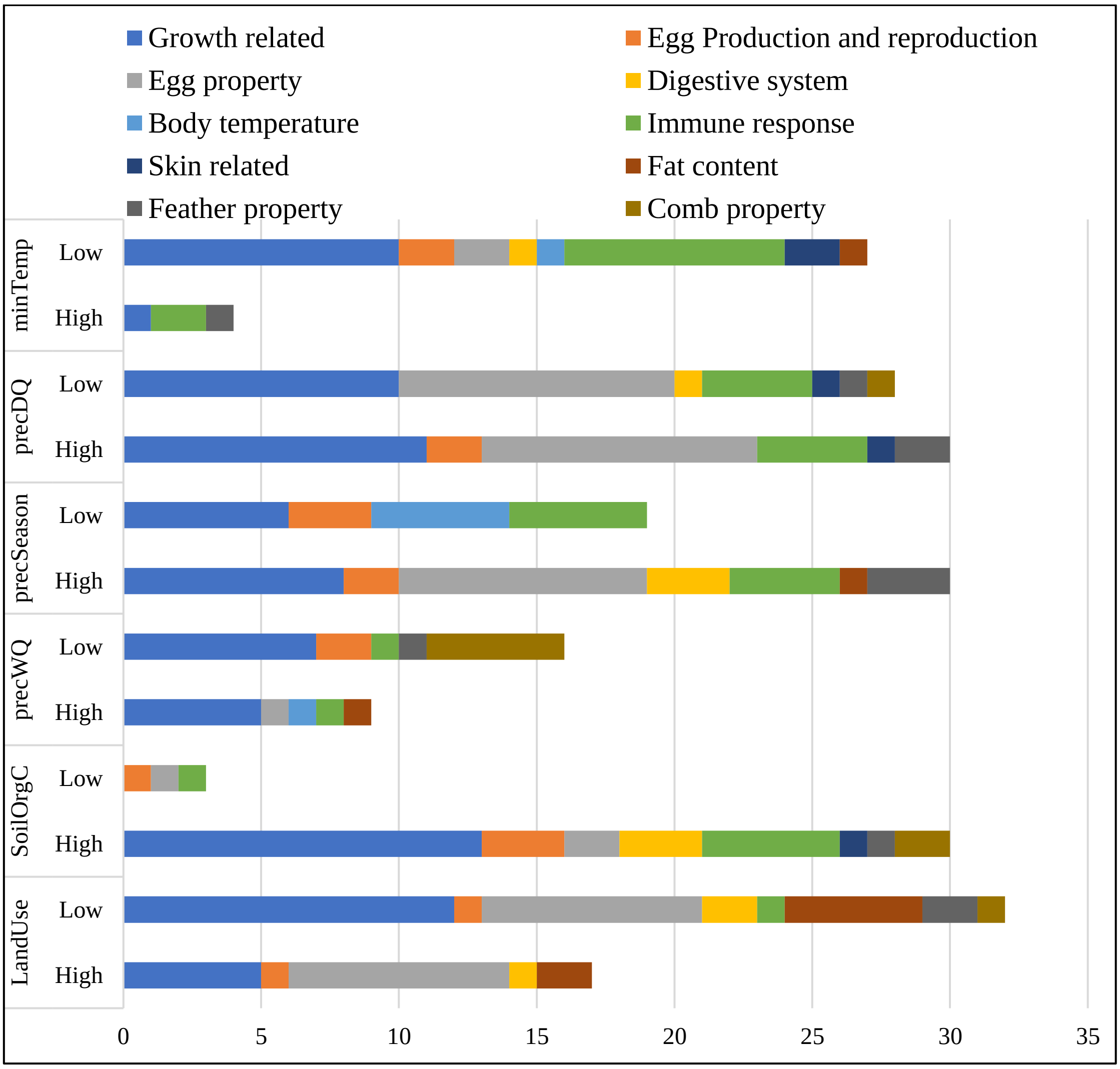


Figure 4.6 Count of candidate genes from selection signature regions overlapping with QTLs of size < 1 Mb.

Candidate genes detected by both (*Fst* and *XP-EHH*) methods were considered our strongest candidates for each environmental variable and were further investigated.

Table 4.3 report these selected genes grouped as (i) adaptation to extreme temperatures and high altitude (minTemp analysis), (ii) adaptation to extreme rainfall patterns (from the analyses of three precipitation variables), (iii) adaptation to food availability, and (iv) adaptation to food abundance related to areas of cultivated land (LandUse) and SoilOrgC content.

4.3.3.1 Adaptation to extreme temperatures and high altitude

As minTemp (minimum temperature of the coldest month of the year) is positively correlated with the maximum temperature of the warmest month (r_s = 0.9) and negatively correlated with elevation ($r_s < -0.91$), the selection signature for these environmental variables mainly reflects adaptation to extreme temperatures and high elevation. Therefore, the Low group populations for minTemp (Alfa Midir and Negasi Amba) live in the coldest environment and the highest altitude. In contrast, the High group population of minTemp live at the maximum temperature and lower altitude (Figure 4.1 and Supplementary Table S4.1). The strongest signals for candidate sweep regions from minTemp analysis were identified in the Low group, suggesting predominant adaptation to low temperature and high altitude (Figure 4.2a, 4.3a, and Supplementary Table S4.2a). Ten candidate genes were commonly detected for *Fst* and *XP-EHH* analyses in the minTemp - Low group. Nine are located in the same candidate region chr5:17250000-17280000 (*CLP1*, *YPEL4*, *ENSGALG00000007381*, *UBEL6*, *TIMM10*, *RTN4RL2*, *SLC43A3*, *PGR2/3*, *P2RX3*), and one additional gene is in a candidate region on chr18:5100000-5120000 (*UTP18*) (Table 4.3).

The genes cluster on chr5 (minTemp - Low group) have functions related to various stress associated with a high-altitude living environment, e.g., hypoxia (Sarkar *et al.*, 2003), thrombosis (Gambhir *et al.*, 2014), and cold temperature (Table 4.3). For instance, *CLP1* is linked to cardiac muscle hypertrophy (Espinoza-Derout *et al.,* 2007), *YPEL4* plays a role in pulmonary diseases (Truong *et al.,* 2018), *P2RX3* and *ENSGALG00000007381* are involved in blood coagulation (Fabregat *et al.*, 2018), and *SLC43A3* plays a possible role in the repair of the lung tissue under oxidative stress (Furukawa *et al.,* 2015). Also, *RTN4RL2, UBE2L6,* and *TIMM10* genes have been found differentially expressed in cells under hypoxic conditions (Lai *et al.,* 2016; Fabregat *et al.*, 2018; Marchesi *et al.*, 2019). The *UTP18* gene, located in chr18, is differentially expressed in the human hypoxic cardiomyocytes cell line (Lee *et al.*, 2018), and it is involved in processing the pre-18S ribosomal RNA (rRNA).

Genes with very high signals, but identified only with one method, were also investigated. *SDK1* (*ZFst* = 8.16) regulates dendritic spine development and synaptic connectivity (de Wit & Ghosh, 2016). A region with a high |*XP-EHHstd*| signal (4.24) overlaps with three genes: *ARFIP2* with a role in autophagy (Judith *et al.,* 2019), *ILK* with a signalling mediator role (Yen *et al.*, 2014), *TRIM3* related to tumour growth and metastasis (Huang *et al.*, 2017) and also identified as a regulator factor of p53 levels (a protein activated by a wide variety of stress signals, including DNA damage, hypoxia, and nutritional deprivation) (Liu *et al.*, 2020).

In the minTemp High group, a total of 25 candidate genes were identified. However, none were simultaneously detected by *Fst* and *XP-EHH* analysis. The highest signal was found for the *TOGARAM1* gene (|*XP-EHHstd*|) = 4.02 (Supplementary Table S4.2a), which plays a role in the assembly of non-motile cilia (Nachury & Mick, 2019). Interestingly, heat shock induces the loss of these organelles (Promodou *et al.*, 2012); thus, this gene might have an adaptive role in easing this effect in high-temperature conditions.

Table 4.3 Candidate genes detected by both Fst and XP-EHH in relation to environmental adaptations.

Environmental adaptation	**Genomic regions and Genes**		**Relevant biological functions for candidate genes**	
	position	**Gene**	**Functions**	**Reference**
Adaptation to high altitude stresses (hypoxia, thrombosis, and cold tolerance) *Selection in minTemp Low group*	Chr5 17230000-17290000	*CLP1*	Linked to cardiovascular function and cardiac muscle hypertrophy	(Espinoza-Derout *et al.*, 2007)
		YPEL4	Role in pulmonary diseases	(Truong *et al.*, 2018)
		ENSGALG 00000007381	Blood coagulation	(Fabregat *et al.*, 2018)
		UBEL6	Involved in protein ubiquitination pathway, and downregulated in cells exposed to hypoxia	(Gambhir *et al.*, 2014) (Fabregat *et al.*, 2018)
		TIMM10	It imports transmembrane proteins into the mitochondrial inner membrane. Downregulated in hypoxic cells	(Lai *et al.*, 2016)
		RTN4RL2	Involved in axon regeneration and protection of motoneurons against apoptosis. Upregulated in the myopathy-affected condition in broiler possibly in response to hypoxia	(Marchesi *et al.*, 2019)
		SLC43A3	Roles in lung tissue repair and growth under oxidative stress	(Furukawa *et al.*, 2015)
		PGR2/3	Immune response	(Parra *et al.*, 2013)
		P2RX3	Blood coagulation	(Fabregat *et al.*, 2018)
	Chr8 5100000-5120000	*UTP18*	RNA binding and involved in pre-18S rRNA processing. It belongs to a gene network targeted by microRNAs differentially expressed in human hypoxic cardiomyocytes cell line	(Gambhir *et al.*, 2014) (Lee *et al.*, 2018)
Water scarcity adaptation *Selection in precDQ Low group*	Chr3 71840000-73980000	*ENSGALG 00000036204*	LncRNA with a possible regulatory role on nearby genes, e.g., the nearest gene is *MMS22L*, with a role in the DNA damage repair process	(O'Donnell *et al.*, 2010)
		ENSGALG 00000025686	*U6* spliceosomal RNA with a possible role in post-transcriptional modification	(Didychuck *et al.*, 2018)
		MANEA	Associated with panic disorder	(Jensen *et al.*, 2014)

		EPHA7	It is involved in many functions, e.g., apoptotic process, axon guidance, brain development, ephrin receptor signalling pathway, and nephric duct morphogenesis	(Clifford *et al.*, 2014)
	Chr3 106430000-106510000	*MSRA*	Involved in cellular protein modification, protein repair, response to oxidative stress	(Singh *et al.*, 2018)
	Chr4 74170000-74300000	*ENSGALG 00000048521* *ENSGALG 00000050078* *ENSGALG 00000046053*	LncRNA genes possible have Cis-regulatory functions on nearby genes. A nearby target is *PPARGC1A*, which is involved in many biological functions	(Puigserver & Spiegelman, 2003)
	Chr6 25040000-25070000	*SLK*	Mediates apoptosis	(Al-Zahrani *et al.*, 2013)
	Chr1 18110000-18130000	*BANP*	Multicellular organism development, transcriptional regulation, and regulation of signal transduction	(Dai *et al.*, 2015)
Possible adaptation to water scarcity *Selection in precSeasonality High group*	Chr4 2730000-2750000	*HTR2C*	It is involved in many functions, e.g., behavioural fear response, regulation of appetite, regulation of corticotropin-releasing hormone secretion, and nervous system processes	(Stamm *et al.*, 2018)
	Chr4 74280000-74310000	*ENSGALG 00000051573*	LncRNA with possible cis-regulatory functions. The nearest protein-coding gene is *ENSGALG00000040208*	(The Uniprot consortium, 2021)
Possible adaptation to excess rainfall and humidity *Selection in precSeasonality Low group*	Chr2 114040000-114080000	*ENSGALG 00000053888*	LncRNA with possible regulatory functions. The nearest gene, *YTHDF3*, has a role in the positive regulation of translation	(Shi *et al.*, 2017)
	Chr7 35360000-35390000	*CACNB4*	Calcium transport, ion transport	(Rima *et al.*, 2017)
Scavenging adaptation to a rich source of insect, worm, and plant-based food *Selection in SoilOrgC High group*	Chr1 197230000-197310000	*HBBA* *HBE* *HBE1* *HBBR*	Genes involved in heme binding, oxygen carrier activity, cellular oxidant detoxification, protein hetero-oligomerisation, response to an organic cyclic compound	(Hardison, 2012)
		ENSGALG 00000052767	novel protein-coding	
	Chr10 6820000-6870000	*THSD4*	Peptidase activity. Possible association with feed efficiency traits	(Yao *et al.*, 2013)
		ENSGALG 00000053176	miRNA with a possible role in RNA silencing and post-transcriptional regulation of gene expression	(The Uniprot consortium, 2021)
	Chr1	*STS*	Lipid and steroid metabolism	(Guo *et al.*, 2020)

Scavenging adaptation to a low surplus of crop and grain-based food *Selection in LandUse Low group*	127750000-127980000	*PUDP*	Nucleotide metabolic process	(von Papen *et al.*, 2013)
	Chr1 129040000-129080000	*ENSGALG 00000052489*	Novel protein coding gene	
	Chr3 34260000-34320000	*SMYD3*	It has a role in transcriptional regulation as a member of an RNA polymerase complex and the cellular response to dexamethasone stimulus	(Bernard *et al.*, 2021)
		KIF26B	Role in cell signalling	(Teng *et al.*, 2018)
	Chr4 56140000-56170000	*NDST4*	It has a strong association with low abdominal fat content in chicken	(Williams *et al.*, 2005)
Scavenging adaptation to crop and grain-rich food *Selection in LandUse High group*	Chr4 75690000-75810000	*NCAPG*	Cell division, mitotic chromosome condensation	(Liu *et al.*, 2017)

4.3.3.2 Adaptation to extreme rainfall patterns

Three out of the six environmental variables are precipitation-related (precSeasonality, precWQ, and precDQ). A decrease in rainfall may have a major impact on the chicken physiology due to water availability, which can alter the animal's general metabolism. On the other hand, rainfall affects the humidity in the atmosphere, which, combined with heat, may exacerbate heat stress. Also, extreme rainfall is often associated with the prevalence of infectious diseases in poultry (Nyoni *et al.*, 2018).

Candidate genes identified by both methods (*Fst* and *XP-EHH*) are found in the precSeasonality Low and High groups and the precDQ Low group. The largest number of commonly detected candidate genes (n = 10) are from the precDQ Low group, which supports water scarcity as the main driver of selection in poultry (Nyomi *et al.*, 2018). Many of these genes are potentially related to regulatory functions (Table 4.3). For instance, *ENSGALG00000036204*, *ENSGALG00000048521*, *ENSGALG00000050078*, and *ENSGALG00000046053*. They are long non-coding RNA genes with possible *cis-*

regulatory functions on nearby genes. Also, *BANP* has known transcriptional regulatory roles (Puigserver & Spiegelman, 2003; Dai *et al.*, 2015). Other genes are involved in various stress responses, e.g., *SLK* mediates apoptosis (Al-Zahrani *et al.*, 2013). *MSRA* is involved in protein modification, repair, and response to oxidative stress (Singh *et al.*, 2018). Another gene, *EPHA7*, is involved in development and morphogenesis (Clifford *et al.*, 2014), while *MANEA is* associated with panic disorder (Jensen *et al.*, 2014). These gene functions illustrate the impact that stress associated with restricted access to water has on the overall metabolism and physiological responses of the chicken.

Genes overlapping with strong signals from a single method (*Fst* or *XP-EHH*) were also explored. *AGTR1* (|*XP-EHHstd*| = 7.9) plays an important role in regulating blood pressure and sodium retention by the kidney (Kim *et al.*, 2015). Two other genes, *TMEEM206* and *AFT3* were detected from a single candidate region with |*XP-EHHstd*| = 10.9. *TMEM206* controls the pH balance in the body (Ullrich *et al.*, 2019), while *AFT3* is a stress-response transcription factor (Zhang *et al.*, 2016b).

Both *Fst* and *XP-EHH* detected four genes for the precSeasonality environmental variable - two in the Low group and two in the High group. The precSeasonality High groups most likely represent an adaptation to water scarcity due to large seasonal variations in rainfall. In this group, *HTR2C* has been linked to many functions, including behavioural fear response and appetite regulation (Stamm *et al.*, 2018), and *ENSGALG00000051573* has possible regulatory roles on gene expression (The UniProt Consortium, 2021).

At the other extreme, the precSeasonality Low group includes populations living in agro-ecologies with low variation in rainfall patterns. It is not evident with what type of stress response these genes are implicated. However, these genes may be associated with the adaptation to an environment with great ambient humidity. Environmental data from analysed populations shows that the relative humidity in the precSeasonality Low group (59 - 109%) is usually higher in the wettest and driest quarters of the year than in the High group (57 - 71%) (Supplementary Figure S4.4). From the precSeasonality Low group, two common genes were detected: *CACNB4,* which has an important role in ion transport (Rima *et al.,* 2017), and *ENSGALG00000053888,* which is a lncRNA-gene with possible *cis*-regulatory function on the nearby genes (Shi *et al.,* 2017).

4.3.3.3 Adaptation to scavenging production systems: using SoilOrgC as a proxy for food availability

The Soil organic carbon content, linked to the organic matter in the soil, is an important biochemical parameter for agriculture (Griffin *et al.,* 2013). A high SoilOrgC reflects the presence of organic compounds and microorganisms like insects and worms that represent an important protein source for the scavenging chicken. So it may be expected that candidate sweep regions identified in the SoilOrgC High group are related to adaptation to a diet rich in organic compounds. Two candidate sweep regions on chr1 and chr10 were detected by both *Fst* and *XP-EHH* analyses (Table 4.3). Interestingly, the chr1 region overlaps three heme-binding genes (*HBBA*, *HBE,* and *HBBR*) related to O_2 carrier activity, and *HBE1* with a role in response to organic cyclic compounds (Hardison, 2012). Studies in broiler chicken with severe myopathic breast muscle have shown overexpression of these genes in the muscle as a possible response to a lack of oxygen in oxidative stress conditions (Pampouille *et al.,* 2019). A novel protein-coding gene (*ENSGALG00000052767*) was also identified for the same region on chr1.

In the second detected region, on chr10, two genes are present: *THSD4* has a peptidase activity and therefore is directly relevant to feeding metabolism (Yao *et al.,* 2013), while the other gene, *ENSGALG00000053176* is a miRNA with a regulatory role in gene expression (The Uniprot consortium, 2021).

A high *ZFst* signal (but not *XP-EHH*) was detected from the SoilOrgC - High group in the region overlapping the *DCLK1* (*ZFst* = 10.36) gene. This gene has a possible role in the calcium-signalling pathway (Lin *et al.,* 2000). It is also associated with ovarian follicle weight (Sun *et al.,* 2015). A high |*XP-EHHstd*| signal (4.03) overlaps *GALNT7,* which is involved in the carbohydrate metabolic process (Bennett *et al.,* 2012).

The SoilOrgC-Low group did not reveal candidate genes identified by both methods. However, a region located on chr33 (position 6470000-6500000) (Figure 4.2e) shows a high *Fst* peak (*ZFst* = 9 – 12). This sweep overlaps with four genes: *DRPA1* with a role in transcriptional regulation (Lüling *et al.*, 2018), *RELA* involved in different functions (including cell growth, immunity, and apoptosis) (Msaki *et al.,* 2011), *UQCC3* implicated in ATP production (Yang *et al.,* 2020), and *KAT5* that acts together with GSK3 as a regulator of autophagy under starvation conditions (Son *et al.,* 2021) (Supplementary Table S4.2e).

4.3.3.4 Adaptation to scavenging conditions in relation to land cultivation pattern

LandUse is another variable strongly linked to food abundance for scavenging chicken, which receives supplementary feeding from the household. A larger percentage of cultivated land may represent a higher amount of supplemental feeding. Between LandUse extreme population groups, a larger proportion of candidate genes was detected in the LandUse Low group (chickens belonging to these groups live in areas with a lower proportion of cultivated land). Four different sweep regions and six genes were identified by *Fst* and *XP-EHH* methods in the LandUse Low group: three genes on chr1, two on chr3, and one on chr4 (Table 4.3). Overlapping the same region on chr1

(127750000-127980000), we find *STS* related to lipid metabolism (Guo *et al.,* 2020) and *PUDP* involved in the nucleotide metabolic process (von Papen *et al.*, 2013). Another gene, - *ENSGALG00000052489 (*a novel protein-coding gene with unknown function), comes from a different region (chr1:129040000-129080000). Within the chr3 region, we find *SMYD3* with a role in transcriptional regulation (Bernard *et al.,* 2021) and *KIF26B* with a role in cell signalling (Teng *et al.*, 2018). The *NDST4* gene overlapping the candidate region on chr4 has been reported to have a strong association with low levels of abdominal fat content in chicken (Williams *et al.*, 2005) (Table 4.3).

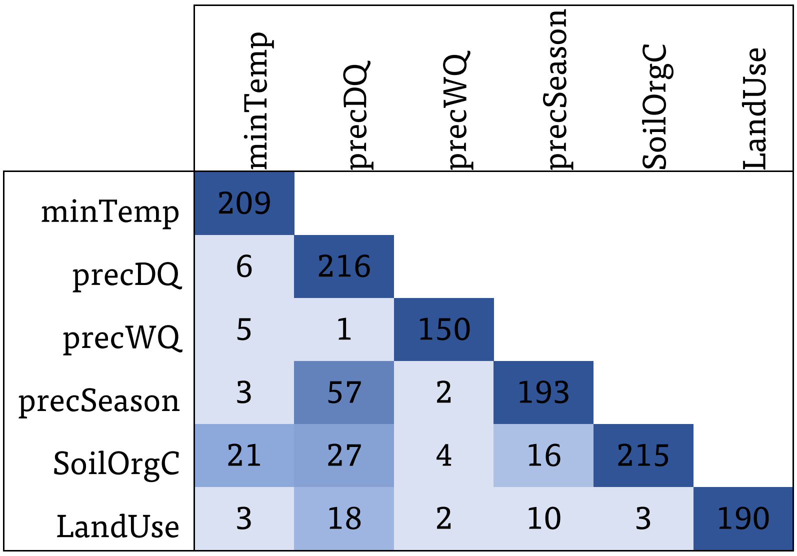
For the LandUse High populations, only a single common candidate gene is detected. This is *NCAPG* involved in cell division and chromosome condensation (Liu *et al.,* 2017).

Large signals for each method (*Fst* and *XP-EHH*) of analysis were also investigated. For the LandUse High group, the *PRDM5* has a *ZFst* of 8.96, and it is involved in the regulation and expression of collagen and other connective tissue components (Galli *et al.,* 2012). A high |*XP-EHHstd*| signal (4.65) was found overlapping *ADIPOR2*, a regulator of glucose and lipid metabolism (Capeau, 2007). For the LandUse-Low groups, a high signal *ZFst* = 10.75 overlaps the gene *ADCY1*, which plays a role in memory, learning, and circadian rhythm (Hwang *et al.*, 2013) (Supplementary Table S4.2f).

4.3.3.5 Shared candidate genes across environmental analyses

Across the six environmental parameters, between 150 and 209 candidate genes were detected (Figure 4.7), giving the total number of 1008 genes from all analyses. Of these, 152 genes (15%) were common to two or more environmental analyses. The range of shared candidate genes between two environmental variables varied between 0.3% and

14% (Figure 4.7). The generally low proportion of overlapping genes shows that the selection of population groups living in the extremes of each environmental variable (Low and High groups) is a useful tool for capturing specific environmental pressures in the genome.



	minTemp	precDQ	precWQ	precSeason	SoilOrgC	LandUse
minTemp	209					
precDQ	6	216				
precWQ	5	1	150			
precSeason	3	57	2	193		
SoilOrgC	21	27	4	16	215	
LandUse	3	18	2	10	3	190

Figure 4.7 Heatmap of shared candidate genes detected across environmental population groups analysed.

4.4 Discussion

Insights of local adaptation are crucial to identify genetic mechanisms that can facilitate livestock environmental resilience. Local adaptation can shape the within-species diversity and create an adaptive divergence in response to extreme agro-climatic variables (Tiffin & Ross-Ibarra, 2014). However, exploring the genetic basis of environmental adaptation is difficult due to the complex interaction between environmental stressors affecting the genome. Thus environmental adaptation is not easy to measure; this is especially true for tropical and subtropical climates with intricate weather patterns (Scholtz *et al.*, 2010).

Different factors have been identified as stressors for livestock animals' survival and productivity, including geographical location and origin, nutrient availability, water availability, photoperiod, management practices, and environmental conditions (Khalifa, 2003; Sejian, 2012). Measuring the adaptation effects to these factors represents a challenge, especially in indigenous populations without specific management systems and living in a broad range of agro-ecologies. Consequently, this study contributes to overcoming this challenge by using an interdisciplinary approach to address this issue. First, this study disentangles and pinpoints the key environmental factors (from a large array of parameters) acting as stressors in particular agro-ecologies and then identifies the candidate genes using complementary genomic approaches.

The results presented here reveal strong signals of adaptation for (i) high altitude with physiological effects as hypoxia and thrombosis, (ii) lack of water during the driest seasons, and (iii) food availability, abundance, and quality.

Perhaps not surprisingly, we did not identify many strong candidate signatures of selection linked to heat tolerance. It is possibly due to the origin and history of African indigenous chickens, whose ancestors were native inhabitants of tropical South and South-East Asia and were already adapted to these regions' hot and humid climates (Pitt *et al.,* 2016).

On the other hand, strong candidates were detected in relation to high altitude and low-temperature adaptation, possibly because when chickens were first introduced from Asia to Africa, the high Ethiopian altitude posed major stress for this newly introduced species, and therefore rapid adaptive evolution was needed to survive in these new environments.

Interestingly, in relation to adaptation to high altitude, no common genes were found between our study and Zhang *et al*. (2016a) and Zhang *et al*. (2016b) studies, which investigated the mechanism of adaptation of Tibetan chickens to their environment.

However, none of their candidates were identified in the present study, although they show similar biological functions. This result may be attributed to different aspects. First, rapid adaptation often works on the "standing variations" in a population (Rees *et al.,* 2020). The different demographic histories of the African and Tibetan chickens may have offered distinct standing variations for natural selection. Thus, even though the adaptive responses were similar in both instances, the associated genes were distinct. Besides, epistatic interactions and pleiotropic effects of genes may favour the selection of one gene over another in different geographic areas (Ostman *et al.,* 2012). These results exemplify the plasticity of the genome in its response to environmental selection pressures.

Among the identified candidate genes, we found several instances of shared genes for two or more environmental analyses, although these were detected in different populations. It may be expected that adaptation to these distinct environmental challenges would involve, to some extent, the same biological pathways. Pleiotropic effects might also explain this phenomenon in which one particular gene can impact the expression of other multiple ones even though they are not directly trait related. Furthermore, geographic origins and biotic interaction with abiotic factors affect responses to the same environmental conditions for different populations or species.

In the long term, environmental adaptation studies on livestock species should aim to identify genomic regions that permit sustainable agro-systems development, producing well-adapted individuals with high production performance. For the Ethiopian indigenous chicken, achieving this goal should involve the sustainable breeding of birds that are already environmentally adapted. Accordingly, the genome regions and genes reported here might be used for animal improvement, not only for backyard scavenging chicken but also for exotic breeds to be introduced in African tropical conditions through markers-assisted selection and breeding improvement strategies.

Our study identified many candidate sweep regions, some with strong signals of selection and overlapping genes were considered as candidates. Obviously because of small sample size, we had to use a window-based approach to identify the sweep regions which did not allow to get very high resolution. This suggest that further validations of these regions would be requires. Such validations can include comparisons with other populations or similar species living in similar environments and also gene expression analyses.

4.5 Conclusion

Our study illustrates the importance of understanding the environmental drivers of selection before investigating genome responses to these environmental stressors. It also exemplifies the use of ENM as a predictive tool to characterise the environment for integration in genomic studies. The genomic data, along with key agro-climatic variables, represent a powerful approach for genome-environmental association analysis aiming to identify candidate signatures of positive selection (selective sweep) and their association with particular agro-climatic variables. Therefore, a powerful tool integrating ecological and genomic approaches can generate more accurate data to improve and sustainable use of Ethiopian indigenous chicken. This approach can be easily applied to any other livestock species.

Chapter 5

Dissecting genome-environment associations in Ethiopian chicken ecotypes

Abstract

Indigenous chickens across Ethiopia are a good example of small populations adapted to diverse and local environmental conditions. Our previous results (Chapter 3) have shown the usefulness of the Ecological Niche Modelling (ENM) approach for the environmental characterisation of poultry population agro-ecologies and for defining ecotypes based on niche similarity. In this chapter, we will now aim to dissect ecotypes' genome-environmental association. Candidate selection signature regions within-ecotype were first identified using *Hp* and *iHS* analysis. A multivariate Redundancy Analysis (RDA) was then used to identify genome-environment associations between genetic variants, overlapping the *Hp* and *iHS* candidate sweep regions and agro-climatic stressors. Sweep regions detected by *Hp* analysis include genomic regions selected following domestication and candidates for immune response-related stress, thermo-tolerance, and food intake stress. *iHS* sweep analysis uncovered genes related to pigmentation, dwarfism, heat stress, water scarcity, and disease resistance. Several of these candidate-selective sweeps were detected in several ecotypes. While the selection signature analyses identify broad candidate regions under selection, the RDA analysis identifies SNPs associated with a specific environmental stressor. A total of 616 outlier SNPs were identified from RDA analysis. These SNPs overlap with genes relevant to different environmental adaptations, such as heat stress and/or water availability (*LDLRAD3*, *TNIP2*, *GPX7*, *PTPRZ1*), as well as muscle growth, body weight and feed metabolism (*GPCPD1*, *NHLRC2 TSHD7B*), and immune response (*MAPK4*).

5.1 Introduction

As described in Chapter 3, Ethiopian indigenous village chicken populations are exposed to diverse ecological conditions. Following an Ecological Niche Modelling (ENM) analysis, key environmental stressors across the Ethiopian landscape were identified, and 12 distinct ecotypes were defined. These proposed ecotypes provide an exciting new platform for exploring environmental adaptation and their genetic basis. This chapter now focuses on dissecting the genetic basis of environmental adaptation in these newly proposed ecotypes.

Various approaches that detect genomic regions under selection have investigated local environmental adaptation (de Simoni *et al.,* 2014). Different factors can influence selection signatures, such as the type of selection, the strength of selection, and the recombination rate (Charlesworth, 2007). Moreover, the selection may have fixed a region or might be an ongoing process so that the regions are yet to reach fixation. Hence, the study in this chapter applies two complementary methods (viz. *Hp* and *iHS)* to capture both types of sweeps – those that have reached fixation (*Hp*) and those yet to reach fixation (*iHS*). The *Hp* method (Rubin *et al.*, 2010) is a site frequency spectrum-based approach used to find sweep regions showing reduced heterozygosity, which have reached either fixation or are close to fixation. In contrast, the *iHS* (Voight *et al.,* 2006) applies an LD-based approach. It measures the amount of extended haplotype homozygosity (*EHH*) at a given SNP location around the ancestral allele relative to the derived allele. This method only detects ongoing selection, i.e., regions yet to reach fixation.

Regarding the time scale, there is not straight forward distinction that *Hp* will detect older sweeps whereas *iHS* will detect newer ones (although that is possible). For instance, *Hp* can also detect strong recent sweep which have gone to fixation very quickly. Similarly *iHS* can also detect slowly applying selection (i.e. weak selection) which is occurring since long time. Thus, considering that chickens in Ethiopia may have been introduced around 820 – 595 BC (Woldekiros & D'andrea, 2016), and they environment is ever changing, these two methods are complementary to understand the selection pressure effect.

The interest in investigating gene-environment association using large datasets of both environment and genetic parameters has led to a new field called landscape genetics, which combines landscape ecology and population genetics/population genomics (Manel *et al.,* 2003). Landscape genetics offers several methods for dissecting genotype-environment association (GEA) (Rellstab *et al.,* 2015), identifying correlations between genetic loci and environmental data as a proxy to local adaptation (Forester *et al.,* 2016).

Most of the early GEA methods are based on univariate statistical analysis testing one locus and one environmental predictor at a time. It includes linear models (Joost *et al.,* 2007; Stucki *et al.,* 2016) and linear mixed models (Coop *et al.,* 2010; Frichot *et al.,* 2013). A disadvantage of these methods is the elevated false-positive rates due to the lack of correction for multiple comparisons (Forester *et al.,* 2016). As large genomic data sets have increased recently, more robust methods that permit analysing simultaneously numerous loci have been developed. Multivariate methods have now been identified as more effective in detecting environmental association with multiple loci (Rellstab *et al.,* 2015).

A broad range of multivariate landscape genetic methods have been developed (Frichot *et al.,* 2013; Bouret *et al.,* 2014; Forester *et al.,* 2016; Stucki *et al.,* 2016). In this study, distance-based redundancy analysis (RDA) was applied (Forester *et al*., 2016). The method is considered a PCA constrained approach that models covarying groups of

SNPs as a response to environmental variables. Based on a locus score, outliers for each constrained axis are identified (Forester *et al.,* 2016). RDA method offers many advantages, such as: (i) it can detect moderate and weakly selected loci, (ii) it offers fast and easy computational application over a large genetic dataset, and (iii) the method is generally robust to weak population structure (Forester *et al.,* 2017). Nevertheless, a strong population structure can result in power decrement in detecting and increasing false positives (Forester *et al.,* 2017). Hence, correction for population structure is needed in such scenarios.

The analysis presented in this chapter has two main objectives. (1) To detect signatures of positive selection within the proposed Ethiopian chicken ecotypes (applying *Hp* and *iHS* methods), and (2) to identify genotype – environmental associations (using RDA) between loci under selection and key agro-climatic predictors of Ethiopian chicken habitats (previously identified by the ENM approach). This study uses knowledge of ecotypes and landscape configuration of Ethiopian village chickens and provides important evidence regarding the genetic mechanisms of local adaptation. The results presented here will be valuable for sustainable farming and conservation programs oriented to smallholder farmers who depend on locally adapted ecosystem-based production.

5.2 Materials and Methods

5.2.1 Genomic data on Ethiopian chicken ecotypes

Sequence data from 243 samples from 25 different chicken populations have been described in detail in Chapter 2. Based on the 12 ecotypes defined in Chapter 3, the genetic variant data were combined for all populations constituting each ecotype.

Though, PCA in chapter 2 showed a weak population structure in the studied samples, this ecotype configuration may contribute to reducing the effect of population structure while downgrading any false positive in selection signature analysis. This is because combining populations to create ecotypes should create heterogeneity in the overall genome except in regions that are under selection, thereby improving the signal strength form adaptive regions. The ecotype details are given in Table 5.1. The number of samples per ecotype varied (10 – 43). In order to keep a similar number of samples per ecotype, we could have subsample larger ecotypes. However, this was not applied, since we could not create major gain for bigger ecotypes in the analysis and also, no improvement for smaller ecotypes.

Table 5.1 Details about proposed ecotypes as defined in Chapter 3.

Ecotype	Populations	No. of samples	No. of SNPs
E1	AmeshaShinkuri, Gafera, Surta, Batambie	37	13,079,450
E2	Ashuda, Dikuli	20	12,242,188
E3	BekeleGirissa, ShubiGemo	20	12,508,181
E4	Adane, LocalHorro	16	12,221,917
E5	Loya, Kumato	20	12,555,563
E6	Mihquan, HadushAdi, Hugub, Jarso	43	13,273,352
E7	NegasiAmba, AlfaMidir	20	11,789,998
E8	Gesses, Kido	18	11,743,436
E9	Meseret, Gijet	19	12,426,154
E10	Metkilimat	10	11,295,871
E11	Arabo	10	10,958,190
E12	TzionTeguaz	10	10,954,331

5.2.2 Selection signature analyses

Putative selective sweeps were detected within individual ecotypes using two complementary methods. First, the pooled heterozygosity (*Hp*) approach was conducted following the procedure described in Rubin *et al.* (2010). The values for the *Hp* were calculated for overlapping 20-kb windows with a 10-kb sliding step using equation (5.1):

$$(5.1) \qquad Hp = \frac{2\Sigma n_{MAJ}\Sigma n_{MIN}}{(\Sigma n_{MAJ} n_{MIN})^2}$$

Where Σn_{MAJ} and Σn_{MIN} are the sums of major and minor allele frequencies for all the SNPs in the 20 kb windows. Afterwards, values calculated for each window were standardised to obtain "Z" scores (*ZHp*). Values were Z-transformed using equation (5.2):

$$(5.2) \qquad Z(Hp) = \frac{Hp - \bar{X}(Hp)}{\sigma\,(Hp)}$$

Where $\bar{X}$ is the mean and σ the standard deviation of the *Hp*. Windows with only SNPs ≥ 10, empirical *P*-value ≤ 0.01, and *ZHp* ≤ - 4.0 were considered outliers and potential sweeps.

The second applied method was the integrated haplotype score (*iHS*). It is computed based on estimating extended haplotype homozygosity (*EHH*) in each bi-allelic SNP, which are ancestral or derived. The sum of *EHH* calculated in both directions from the core SNP is referred to as integrated *EHH*. It is referred to as iHH^{A} for the ancestral allele, iHH^{D} for the derived allele, and *iES* for the SNP site (Gouveia *et al.*, 2014). The *iHS* calculation is based on the extent of decay of linkage disequilibrium surrounding a marker exposed to natural selection, described as a within-population score as the ratio between *iHHA* and *iHHD* (Voight *et al.,* 2006). It was calculated using equation (5.3):

$$(5.3) \qquad uniHS_{(s)} = ln\left(\frac{iHH^{A}_{(s)}}{iHH^{D}_{(s)}}\right)$$

Where $uniHS$ represents the un-standardized log-ratio of ancestral iHH^{A} to derived iHH^{D} of a certain focal marker (s), afterward, a standardization was performed applying equation (5.4):

$$(5.4) \qquad iHS_{(s)} = \frac{uniHS_{(s)} - mean(\,uniHS|_{p_s})}{sd(\,uniHS|_{p_s})}$$

Where $mean(\, uniHS|_{p_s})$ and $sd(\, uniHS|_{p_s})$ represent the mean and standard deviation of $uniHS$ restricted to those variants with a similar derived allele frequency as observed at the focal marker (p_s).

In practice, *iHS* scores were calculated for individual SNPs using the Hapbin package (Maclean *et al.*, 2015), and then means were taken from multiple SNPs within sliding windows (20 Kb size and 10 kb step). All SNPs with missing genotypes were removed from the analysis. This approach requires phasing of the SNP data, which was performed using Beagle v5.1 (Browning & Browning, 2007). The *iHS* values for individual SNPs were calculated using a minor allele frequency of 10% and an Extended Haplotype Homozygosity (*EHH*) value equal to 0.1. Consecutively, mean values were calculated within windows for (i) standardized *iHS* (*iHS_std*) and (ii) absolute value of standardized *iHS* (|*iHS_std*|). Based on |*iHS_std*|, the proportion of SNPs within a window showing |*iHS_std*| ≥ 2.0 was calculated. Empirical P-values were calculated based on the mean of |*iHS_std*| within windows by sorting the scores in descending order (i.e. by ranking the windows) and dividing their ranks by the total number of analysed windows. Criteria to select outlier windows were SNPs ≥ 10, empirical *P*-value ≤ 0.01, window mean of |*iHS_std*| ≥ 2.0, and a minimum of 90% of SNPs with |*iHS_std*| ≥ 2.0 in the window.

Finally, sweep regions showing more extreme signals (either *ZHp* ≤ -5 or *iHS* ≥ 3.5) were investigated more closely for overlapping genes as they potentially represent the most important candidates.

5.2.3 Redundancy analysis (RDA)

RDA analysis was performed based on SNPs overlapping sweep regions detected in the above analyses. As RDA can not deal with large genotype data (scale of millions) (Csala *et al.*, 2017), a reduction of our variant dataset was performed by LD pruning using PLINK v1.9 (Purcell *et al.*, 2007), with the following parameters: --indep-pairwise, window-size = 10 kb, step-size = 10 SNPs, and $r^2 = 0.5$.

RDA was performed using the R package Vegan v2.5-4 (Oksanen, 2015) following (Forester, 2019). LD-pruned SNP genotype data were fitted as response variables (Y). The six environmental variables (Chapter 3) - key predictors of Ethiopian chicken agro-ecologies - were the explanatory variables (x) in the RDA analysis. These environmental variables are the minimum temperature of the coldest month (minTemp), precipitation seasonality (precSeasonality), precipitation of wettest quarter (precWQ), precipitation of driest quarter (precDQ), soil organic carbon (SoilOrgC), and land use (LandUse).

SNP genotypes were converted as a dosage of alternative alleles (i.e. 0, 1, and 2) to be fitted as the response (y) variable. In addition, the RDA model was conditioned on latitude and longitude to correct for spatial autocorrelation and neutral genetic structure due to the geographic proximity of the populations. Moreover, ancestry coefficients (for $K = 3$) from previous admixture analysis (Chapter 2) were also included in the model as covariates to be partially out to correct for possible population structure from different ancestral backgrounds.

The significance of the partial model was estimated with 999 permutations and for each axis with 99 permutations. A two-tail SNP loading distribution was used to determine outliers or candidates for local adaptation with a cut-off value of SD >= 3 (P-value = 0.0005) for each significant axis.

5.2.4 Functional annotation

Candidate regions/variants from selection signature and RDA analyses were investigated with the Ensembl genome browser (release 98) using the Ensembl Variant Effect Predictor (VEP) platform (McLaren *et al.,* 2016) for genes annotation. In addition, molecular functions, biological process, and metabolic pathways were also analysed for the candidate genes using the PANTHER Classification System v.14.0 (Huaiyu *et al.,* 2019) and the Kyoto Encyclopaedia of Genes and Genomes (KEEG) (KOBAS version 3.0, Release 97.0) (Kenehisa & Goto, 2000).

5.3 Results

One of the main goals of defining ENM-informed ecotypes in our study is to use these for dissecting the genetic basis of environmental adaptations. Accordingly, selection signature analyses (*Hp* and *iHS*) were first run within each of the defined ecotypes to identify putative sweep regions under positive selection. Then, the SNPs from the putative sweep regions were used to investigate linear association between genotype and environmental parameters using the RDA approach.

The results are structured in the following subsections. First, the summary results from *Hp* and *iHS* analyses are presented to provide a comparative overview. Then, genes overlapping the strongest candidate regions from different ecotypes are discussed separately for the *Hp* and *iHS* analyses, followed by a section on shared sweeps regions detected by both methods. In the final subsection, results from the RDA analysis are presented.

5.3.1 Summary results from selection signature analyses

For each of the 12 ecotypes, signatures of positive selection were calculated for *Hp* and for *iHS,* analysing between 92,364 and 93,313 windows across the genome (Table 5.2).

Supplementary Figures S5.1a and S5.1b show the observed distributions of *Hp* and *iHS* values for each ecotype analysis.

The genome-wide mean values of *Hp* were similar across all the ecotypes, ranging from 0.29 ± 0.06 for E7 and 0.33 ± 0.05 for E10. However, the number of candidate windows (with *P*-value < 0.01, *ZHp* < -4) varied widely, from only seven candidate windows for E7 to 279 candidate windows for E10. Adjacent or overlapping candidate windows were merged to define the sweep regions. This resulted in 2 to 67 unique sweep regions in individual ecotypes and a total of 320 regions for all combined ecotypes (Table 5.2).

Table 5.2 Genome-wide pool heterozygosity (*Hp*) and integrated haplotype score (*iHS*) for the 12 proposed chicken ecotypes in Ethiopia.

		Pool heterozygosity (*Hp*) statistics				**Integrated haplotype score (*iHS*) statistics**				
Ecotype	**No. of samples**	**No. of windows**	**Mean ± SD**	**No. of candidate window (unique ecotype specific windows)**	**No. of sweep regions**	**No. of windows**	**Mean ± SD**	**No. of candidate window (unique ecotype specific windows)**	**No. of sweep regions**	**No. of Common *Z(Hp)* and |*iHS*| window**
E1	37	93031	0.30 ± 0.04	108 (14)	29	92654	0.813 ± 0.36	176 (67)	41	0
E2	20	93224	0.31 ± 0.05	115 (22)	35	92928	0.797 ± 0.40	234 (161)	53	4
E3	20	93191	0.32 ± 0.04	250 (59)	56	92911	0.804 ± 0.38	112 (56)	32	8
E4	16	93042	0.31 ± 0.05	127 (9)	33	92627	0.809 ± 0.39	87 (54)	29	0
E5	20	93258	0.31 ± 0.04	146 (30)	29	92918	0.816 ± 0.40	389 (178)	74	7
E6	43	92687	0.30 ± 0.04	167 (38)	45	92364	0.811 ± 0.33	48 (35)	23	9
E7	20	93200	0.29 ± 0.06	7 (0)	2	92740	0.788 ± 0.39	181 (130)	44	0
E8	18	93175	0.30 ± 0.05	64 (8)	10	92713	0.808 ± 0.41	243 (144)	56	1
E9	19	93277	0.32 ± 0.04	267 (70)	58	92935	0.818 ± 0.41	312 (263)	80	12
E10	10	93313	0.33 ± 0.05	279 (65)	67	93090	0.815 ± 0.40	81 (44)	22	0
E11	10	93261	0.31 ± 0.06	59 (20)	18	92971	0.752 ± 0.50	559 (435)	180	0
E12	10	93266	0.32 ± 0.05	141 (38)	34	93009	0.791 ± 0.43	154 (102)	46	0

* Candidate (sweep) regions were determined after merging overlapping or adjacent windows.

Similar to *ZHp* analysis, the genome-wide mean values of *iHS* were close across all twelve ecotypes (ranging between 0.75 ± 0.5 for E11 and 0.81 ± 0.4 for E9). Still, the number of candidate sweep windows ($|iHS| \geq 2.0$, P-value ≤ 0.01) showed great variation between ecotypes (Table 5.2). The lowest number of candidate windows were recorded for E10 (n = 36) and E6 (n = 48), and the highest in E11 (n = 559). Merging the overlapping candidate *iHS* windows within ecotypes resulted in 22 to 180 sweep regions (Table 5.2). E6, with the highest number of samples analysed, has the lowest numbers of candidate regions across the chicken genome (n = 48). On the other hand, E11 showed the highest number of candidate regions (n = 180), with the remaining ecotypes with 22 – 80 sweep regions.

The representation of candidate sweep regions across the chromosomes was further explored (Figure 5.1 and 5.2). *Hp* sweeps are distributed across most of the autosomes on the chicken genome except chromosomes 13, 16, 17, 21, 25, and 28, from which no candidate sweeps were detected (Figure 5.1). In contrast, *iHS* sweeps were found across all autosomes (Figure 5.2). In addition, the concentration of sweeps from individual ecotypes varied greatly among chromosomes (Supplementary Table S5.1 and 5S.2).

One chromosome show candidate *ZHp* sweeps from a single ecotype, viz. chr32 for E1, and only chr5 present sweeps from all the 12 ecotypes. Six chromosomes (chr13, 16, 17, 21, 25, and 28) do not have any *ZHp* candidate sweep (Figure 5.1 and Supplementary Table S5.1). Also, five chromosomes show candidate *iHS* sweeps from a single ecotype, viz. chr15 and chr32 have sweeps only from E9, chr22 from E12, chr24 from E5, and chr31 in E2. However, other chromosomes have sweeps from multiple ecotypes. For example, chromosomes 1 and 2 show sweeps from all 12 ecotypes, and chr5 reveals sweeps from 11 different ecotypes (Figure 5.2 and Supplementary Table S5.2).

Among all the candidate *Hp* sweep regions, 47 regions (15%) were detected in at least two ecotypes (Supplementary Table S5.3). E6 has the highest number of shared sweeps with other ecotypes (n = 19), representing 42% of the total sweeps for this ecotype. In contrast, E7 shows only two shared sweep regions.

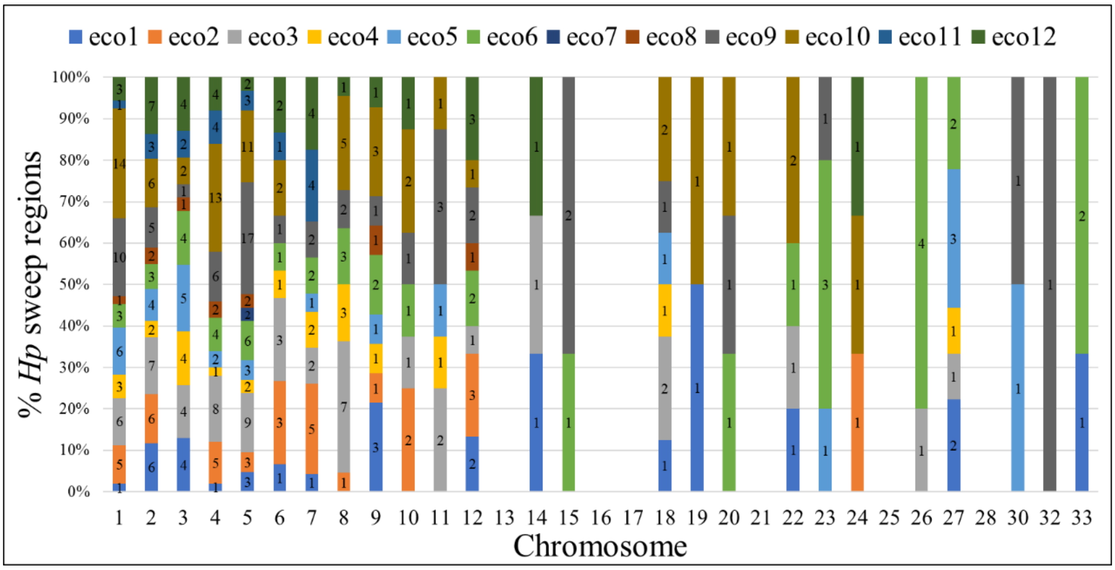


Figure 5.1 Chromosome-wise percentage distribution of *Hp* sweeps within chromosomes by ecotype
(*number in bars represents sweeps count per ecotype).

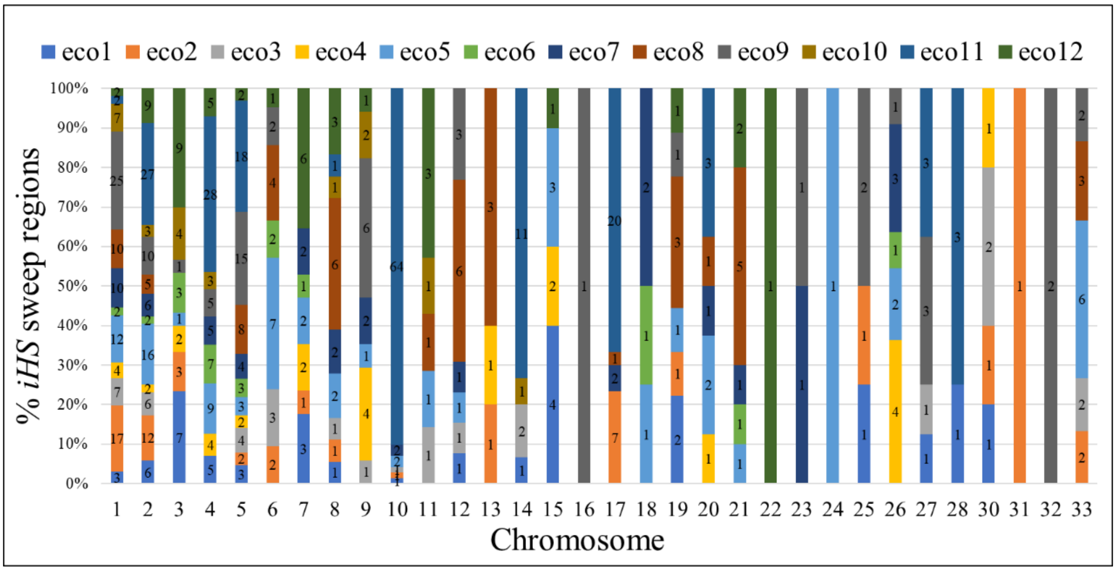


Figure 5.2 Chromosome-wise percentage of *iHS* sweeps within chromosomes by ecotype.
(*number in bars represents sweeps count per ecotype).

A few *Hp* sweep regions were detected across the majority of the ecotypes. For example, one candidate sweep located at chr5: 22,750,000 – 22,800,000 in chromosome 5 (position 22750000 – 2280000) was detected in all ecotypes except E12 (*ZHp* ranges -4.035 to -4.886). (Supplementary Table S5.3). This 50 kb sweep region overlaps with a LncRNA gene (*ENSGALG00000052351*), with a possible regulatory effect on nearby genes. The closest protein-coding gene to this is *FUT8* (24.14 kb forward), a well-conserved region gene across different vertebrate species involved in the protein glycosylation pathway (Coullin *et al.,* 2002).

Candidate sweeps overlapping *TSHR* loci were identified in all the ecotypes. Regions (Chr5: 41000000_41030000 - (30 Kb), overlapping with - *TSHR* locus gene) were detected in E1, E2, E3, E7, E9, and E11, while the region (chr5: 40990000_41040000 - (50 kb), overlapping with - *TSHR* and *GTF2A1* loci) was detected in E4, E5, E6, E10, and E12. These two regions also represented the lowest *ZHp* values for six different ecotypes (E1, E2, E3, E5, E6, and E7), with scores ranging from -4.268 to -6.04 (Supplementary Table S5.3). A number of previous studies have also reported sweep regions harbouring *TSHR* and *GTFA21* genes from different chicken breeds, which are suspected of having played an important role in the early improvement of chickens (Rubin *et al.,* 2010; Gheyas *et al.,* 2015; Lawal *et al.,* 2018).

Twenty *iHS* sweep regions (2.9%) were detected in at least two ecotypes (Supplementary Table S5.4), but unlike the *Hp* results (Supplementary Table S5.3), none of the *iHS* sweeps was identified across all or most of the ecotypes. E8 and E9 had the highest shared sweeps with other ecotypes (3 out of 6). Ecotype 6 did not share any *iHS* sweep with other ecotypes.

The sweep regions show large size variations, ranging from 20 kb -150 kb for *Hp* and 20 kb - 360 kb for *iHS* (Figure 5.3). The largest *Hp* sweeps (150 kb, spanning across 12 windows) were identified in E5 (chr2:147010000-147160000; *ZHp*: -4.39 ± 0.31) and E9 (chr2: 147050000- 147200000; *ZHp*: -4.41 ± 0.27) (Suppl. Tables S5.5.5 & S5.5.9, Figures 5.4e & 5.4i). These 150 kb sweep regions (in E5 and E9) are partly overlapping. The region harbours the *TSNARE1* gene. Sweep regions (of variable size) intersecting with *TSNARE1* have been detected from most ecotypes, except E7, E8, and E11.

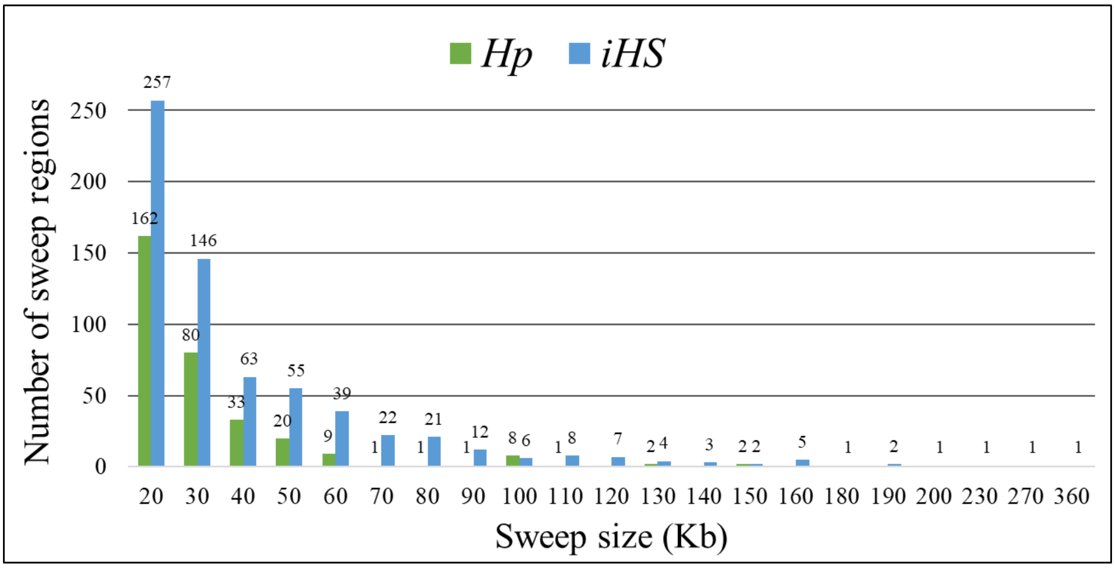


Figure 5.3 Size variation (kb) of detected *Hp* and *iHS* sweep regions.

Nine other *Hp* candidate regions have fragment sizes ≥ 100 kb. Some of these regions detected from different ecotypes were partly overlapping. Beside the region on chr2 presented before, these include a region on chr9 with a 100 kb region in E6 (chr9: 10,360,000-10,460,000) and a 130 kb region in E12 (chr9:10,330,000-10,460,000). Several genes are found within the overlapping region: *ZBTB38* with a role in DNA replication (Miotto *et al.*, 2014), *RASA2* a GTPase activator gene (Li *et al.*, 1996), *RNF7* involved in lipid metabolism (Huang *et al.*, 2011), *GRK7* related with photoresponse (Lamb *et al.*, 2018), and *ATP1B3* (See the sub-tables under the Supplementary Table S5.5 for respective ecotypes).

Forty-two *iHS* regions (6% of the 657 sweep regions detected) were larger than 100Kb. The largest-sized candidate sweep region is 360 kb (chr1:74330000-74690000) in length. It is found in E9 (*iHS* = 2.58 ± 0.12). Two genes, *ANO2* and *VWF*, involved in calcium transport and homeostasis (Stephan *et al.*, 2009; Xu *et al.*, 2012), are found within this region (Supplementary Table S5.6.9 and Figure 5.5i). Two regions have lengths > 200 kb. A sweep of 270 kb identified in E2 (*iHS* = 2.38 ± 0.12, chr2:9240000 – 9510000), which overlaps with the *STRBP* gene involved in spermatogenesis (Meng *et al.*, 2018), and *DENND1A* with a mediated endocytosis role (Gammoh *et al.*, 2015) (Supplementary Table S5.6.2 and Figure 5.5b). Another sweep of 230 kb detected in E7 (*iHS* = 2.55 ± 0.22, chr26:90000 – 320000 bp) includes eleven different genes with multiple functions. The *CLPS* related to lipid metabolism (Pernes *et al.*, 2019), *LHFPL5* involved in hair morphogenesis in mice (Mahendrasingam *et al.*, 2017), *SRPK1*, which regulates alternative splicing (Zhong *et al.*,2009), *SLC26A8* ion transporter (Dirami *et al.*, 2013), *MAPK14* and *MAPK13* involved in cellular proliferation (Lee *et al.*, 2015), *RPF3* with hydrolase activity, *TBC1D22B* with GTPase activity (Fukuda, 2011), *PIM1* related with blood cell proliferation (Magnuson *et al.*, 2010), *CACNA1S* with a roll in ion channel activity and hyperthermia (Carpenter *et al.*, 2009), and *KIF21B* involved with microtubule dynamics (Muhia *et al.*, 2016) (Supplementary Table S5.6.7 and Figure 5.5g).

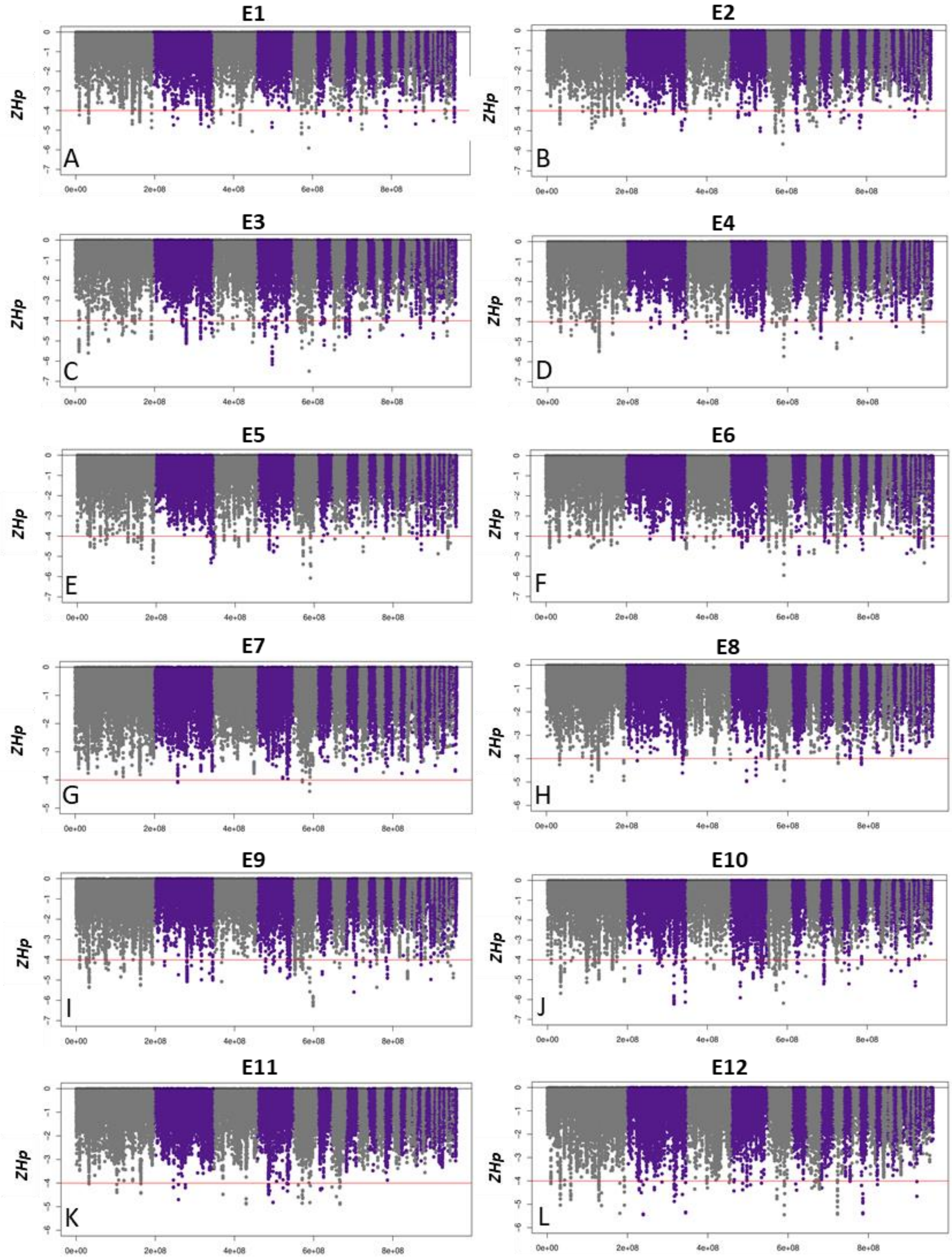


Figure 5.4 Candidate standardized pool heterozygosity (*ZHp*) signatures of positive selection in Ethiopian chicken ecotypes.

The X-axis represents the genomic positions within 20 kb windows along the genome.

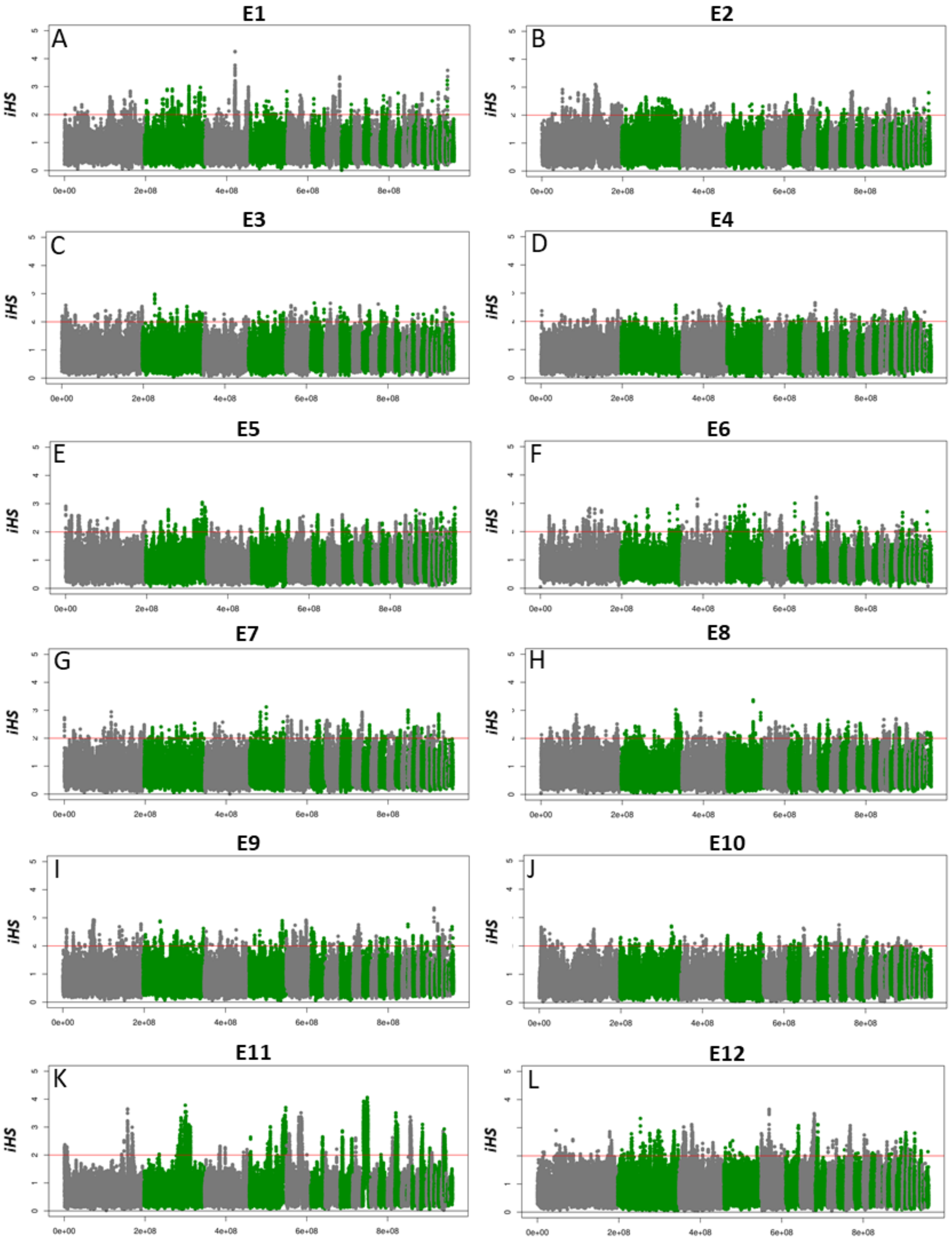


Figure 5.5 Candidate integrated haplotype score (*iHS*) signatures of positive selection in Ethiopian chicken ecotypes using the integrated haplotype score *iHS*.
The X-axis represents the genomic positions within 20 kb windows along the genome.

Even though, the methods used in this chapter 5 and chapter 4 are based on different statistical assumptions, it is interesting to identify which of the detected sweep regions overlap. Such information can be useful as a way of validation of the signals, especially if they are coming from population or ecotypes with similar environmental variables. When the selective regions (SRs) identified by *iHS* and *Hp* analyses are compared with SRs identified in chapter 4 (*Fst* and *XP-EHH* analyses) a total of 162 regions representing all 12 ecotypes overlapped (Supplementary Table S5.7).

5.3.2 Important sweep regions and genes detected in *Hp* analysis

From each ecotype, genes overlapping the extreme signals (sweeps with $ZHp \leq -5$) were further investigated for their function and potential relation to agro-climatic variables (Table 5.3). These include 16 genes distributed across 13 candidate sweeps in nine different ecotypes. For three ecotypes (E7, E8, and E11), no windows passed the threshold of $ZHp < -5$ (Figure 5.6). For these, the candidate region with the lowest significant *ZHp* threshold (-4) was checked for overlapping genes (Table 5.3).

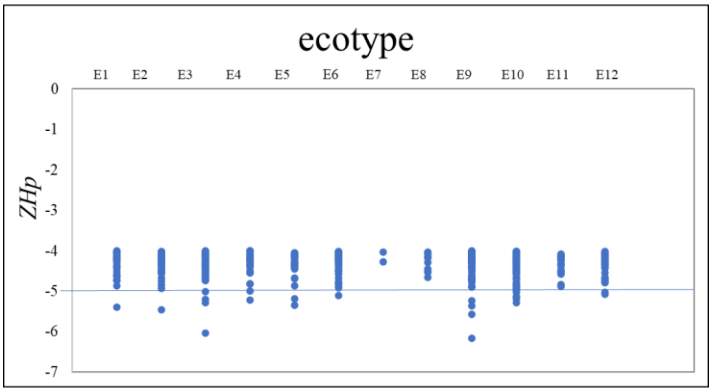


Figure 5.6 *ZHp* values for candidate genome regions in each ecotype.
Genome regions showing sweeps with signals over a threshold $ZHp \leq -5$ (blue line) were further investigated. (See Supplementary Table S5.5 – sub-tables 1 to 12).

Table 5.3 Genes overlapping sweep regions with *ZHp* value < -5 or the strongest signalling *ZHp* sweep in each ecotype.

GENE	Gene function / association	Reference	chr_start_stop	kb	u*ZHp* ± s.d.[a]	Eco[b]
DRD3	It is a dopamine receptor gene with an important role in aggressive chicken behaviour.	Li *et al.*, 2016	1_83730000 _83750000	20	-5.01	E3
EFHC2	Involved in neuron development	Gaudet *et al.*, 2011	1_1124000000 _112430000	30	-4.65	E8
IL1RL2	Interleukin protein. Informative SNPs have been found for immune response and disease resistance.	Kaiser *et al.*, 2008	1_135140000 _135160000	20	-5.05	E12
HHLA1	Protein coding gene. A candidate gene for feralisation in the chicken	Johnson *et al.*, 2015	2_141370000 _141400000	30	-5.19 ± 0.1	E5
OC90	Calcium ion-binding and phospholipase activity. Gene associated with convergent adaptation to the tropical desert climate in birds.	Shilin *et al.*, 2020				
TSNARE1	Neurobehavioural gene. Shows parallel fixation in different commercial traits, suggestive of selection during pre-improvement ages.	Qanbari *et al.*, 2015	2_147130000 _147180000	50	-5.03 ± 0.6	E12
ENSGALG 00000053847	LncRNA		3_82190000 _82220000	30	-4.87	E11
ESRRG	Novel susceptibility gene. It is associated with BMI and fasting glucose level. It is located within a QTL region for abdominal fat weight.	Shan-Shan *et al.*, 2016	3_20100000 _20120000	20	-5.07	E9
TACR3	Protein works as a mediator of signals across the cellular membrane. Selective sweep detected for Ethiopian and Sri Lankan chicken. It was identified as a regulator of gonadotropin-releasing hormone secretion, which regulates sexual maturity.	Lagerström *et al.*, 2006 Lawal *et al.*, 2018 Topaloglu *et al.*, 2009	4_39150000 _39220000	70	-5.28 ± 0.2	E3
BEGAIN	Gene is involved in protein-protein interaction across chemical synapses. It regulates the development and differentiation of adipose and muscle.	Shin *et al.*, 2010	5_48990000 _49080000	90	-6.16 ± 0.1	E9
TSHR	Gene is identified as the most striking selective	Rubin *et al.*, 2010	5_41000000 _41020000	20	-5.57	E9

	sweep found in all domestic chicken. It regulates photoperiod response and seasonal reproduction. Therefore has been important for chicken domestication.	Karlsson *et al.*, 2016	5_41000000 _41030000	30	-6.04± 0.7	E3
			5_41000000 _41030000	30	-5.45 ± 0.3	E2
			5_41000000 _41030000	30	-5.38 ± 0.7	E1
GTF2A1	Gene within the same linkage disequilibrium (LD)-block as TSHR. It was reported as a candidate gene for egg production.	Karlsson *et al.*, 2016 Liu *et al.*, 2011	5_41000000 _41040000	40	-5.33 ± 0.7	E5
			5_41000000 _41040000	40	-5.19 ± 0.9	E10
			5_40990000 _41040000	50	-5.10 ± 0.7	E6
NRXN3	The gene has a significant polymorphism after heat stress when resistant versus susceptible to heat stress chicken are compared.	Zi-Xuan *et al.*, 2019	5_40540000 _40560000	20	-5.107	E9
ARHGAP22	It is a protein-coding gene involved in endothelial cell angiogenesis and increased capillarity permeability. It is associated with diabetic retinopathy risk.	Huang *et al.*, 2011	6_19150000 _19170000	20	-5.207	E3
ZZZ3	It is a gene with an important role in chromatin organisation. The gene modulates histone acetylation and gene activation.	Mi *et al.*, 2018	8_19380000 _19410000	30	-5.226 ± 0.5	E9
RNF7	(lncRNA) involved in lipid metabolism. Association found with risk of developing liver fibrosis or cirrhosis	Huang *et al.*, 2017 Kupcinskas *et al.*, 2017	9_10420000 _10460000	40	-5.22 ± 0.1	E4
GRK7	The gene expresses a retina-specific kinase involved in the photoresponse's shutoff and adaptation to changing light conditions.	Lamb *et al.*, 2018				
ATP1B3	Gene is related to homeostasis, muscle contraction, and the transport of small molecules. Associated with heat stress in chicken and shows low expression under low-temperature stress.	Zi-Xuan *et al.*, 2019 Changgeng *et al.*, 2015				
RASA2	Control of cellular proliferation and differentiation. SNPs surrounded the gene are associated with lymphocyte 9H/L ratio, which is a disease resistance trait.	Zhu *et al.*, 2019	9_10360000 _10420000	70	-5.05 ± 0.4	E12
HYDIN	Cilia motility protein. The located region is important for growth traits (testis weight).	Zhang *et al.*, 2020	11_1550000 _1570000	20	-5.35	E9

CACNA2D3	The upregulated gene participates in activating and signalling the resistant pathway. It is associated with protection against *Salmonella.*	Li *et al*., 2018	12_8120000 _8150000	30	-5.07 ± 0.4	E12
BCDO2	The gene has an important role in oxidising carotenoids during the biosynthesis of vitamin A. Yellow skin locus associated with chicken domestication.	Eriksson *et al*., 2008	24_6150000 _6180000	30	-5.02 ± 0.1	E10
TEX12	Encodes the testis-expressed sequence 12 protein. Gene within the same linkage (LD)-block as BCO2.	Gazda, 2019				
IL18	It is an interleukin protein with pro-inflammatory activity in spleen cells. The gene can modulate innate and adaptive immunity, and its dysregulation can cause autoimmune or inflammatory diseases.	Fabbi *et al*., 2015				
SDHD	Mitochondrial inner membrane protein. It plays an important role in the respiratory chain transporting electrons.	Hirawake *et al*., 1999				

(a) S.D. values are given when a region has multiple overlapping windows.
(b) Eco: ecotype.

Interesting associations and functions related to immune response, temperature stress, food intake, improvement, and gene regulation were found for the genes overlapping sweeps with extreme *Hp* signals (Table 5.3). A well-known human selected gene is *BCDO2,* responsible for the yellow skin trait in chickens (Eriksson *et al.*, 2008). It was detected in ecotypes E10 (*ZHp* = -5.02). Three more genes overlap the *BCDO2* sweep region. These are the *TEX12* gene expressed in testis (Gazda, 2019), *IL18,* which modulates the innate and adaptive immunity (Fabbi *et al.*, 2015), and *SDHD* with an important role in the mitochondrial respiratory chain (Hirawake, 1999). Studies of carotenoid ornamentation in canaries (Gazda, 2019) found a strong linkage and a consistent parallel expression of these four genes (*BCDO2*, *TEX12, IL18,* and *SDHD*).

GRK7 gene is involved in the adaptation to changing light conditions (Lamb *et al.,* 2018). Although this gene overlapped a strong signalling sweep only in E4 (*ZHp* = -5.22), it is also detected in E1, E4, E6, and E12 ecotypes. Another interesting gene is *TACR3*. It produces a gonadotropin hormone regulating sexual maturity (Lagerström *et al.,* 2006; Topaloglu *et al.,* 2009). Although *TACR3* shows its highest signals for E3 (*ZHp* = -5.28), it is also detected in candidate regions for E1, E5, E6, E8, and E12 (*ZHp* ≤ -4).

Likewise, a neurobehaviourally relevant gene *TSNARE1* shows a high signal in E12 (*ZHp* = -5.02), and candidate regions harbouring the gene were also identified in E1, 2, 3, 4, 5, 6, 9, and 10. *TSNARE1* gene shows parallel fixation in different commercial lines (Qanbari *et al.,* 2015). A very interesting gene overlapping a high signal sweep in E5 (*ZHp* = -5.19) is *HHLA1*, identified as a candidate for the feralisation of Kauai Island’s chicken population (Johnson *et al.*, 2015). This trait may be of great interest due to the scavenging nature and upbringing of the indigenous village chicken.

Three different genes related to immune response were identified. First, the *IL1RL2* gene (detected in E5 with *ZHp* = 5.05) was previously reported as a candidate gene for immune response and disease resistance in brown leghorn, Rhode Island red, and Wellcome inbreed lines of chicken (Kaiser *et al.,* 2008). Second, the *RASA2* gene (detected in E12 with a *ZHp* = -5.05) controls cellular proliferation and differentiation associated with lymphocyte 90/L ration, a disease resistance trait (Zhu *et al.,* 2019). Finally, the *CACNA2D3* gene (detected in E12 with a *ZHp* = -5.07; also detected in E10) participates in the signalling resistant pathway and has been associated with protection against *Salmonella* in chickens (Li *et al.,* 2018).

Two genes related to temperature stress (*ATP1B3, NRXN3*) and one to the tropical desert climate (*OC90*) were also identified. The *ATP1B3* gene function is related to homeostasis, muscle contraction, and transport of small molecules. It was reported as being involved in controlling low-temperature stress in Tilapia (Changgeng *et al.*, 2015) and as a candidate gene for controlling heat stress in broiler chicken (Zi-Xuan *et al.*, 2019). *ATP1B3* showed a high signal in E4 (*ZHp* = -5.22) and was also detected in E1, E6, and E12 (*ZHp* ≤ -4). These ecotypes are found in cool-humid and sub-humid parts of Ethiopia, with mean minimum temperatures in the coldest month ranging between 5.5 and 11.6°C, and mean maximum temperatures in the warmest month between 25.3 and 33.1°C. Therefore, the gene may be involved in adaptation to the large temperature difference experienced by these ecotypes. The other ecotypes (2, 3, 5, 7, 8, 9, 10, and 11 with *ZHp* of -1.69 ± 2.05) do not have such large variations in temperature and did not show signatures of selection in these regions.

The gene *NRXN3* has revealed nucleotide polymorphism associated with body temperature changes and heat stress in broiler chickens (Zi-Xuan *et al.*, 2019). This gene in a genomic region with a high *ZHp* signal of -5.1 for E9 was also detected in E6. For E9, the mean value for maximum temperature in the hottest month is 28.9°C, and for E6 33.1°C. The minimum temperature in the coldest month for E9 is 7.3°C and for E6 11.6°C, respectively. These ecotypes also experience medium to high values for precipitation seasonality (82 to 141 mm/m^2). This interaction between temperature and humidity impacts thermal comfort and performance in broiler chicken (Purswell, 2012). The other ecotypes' mean temperature ranges from 25.3°C to 33°C with a precipitation seasonality medium to low (100 to 50 mm/m^2).

The gene *OC90* involved in the arachidonic acid metabolism plays an important role in pain and inflammatory responses (Stelzer *et al.*, 2016). Only E5 showed a high signal at this gene (*ZHp* = -5.19), whereas all other ecotypes did not show any candidate windows for this genomic region. E5 displays minimum temperatures of 8.7 °C, a maximum temperature of 28.1°C, and precipitations between 116 to 505 mm/m^2.

Two genes related to adipose deposition and differentiation, *BEGAIN* (Shin *et al.,* 2010) and *ESRRG* (Shan-Shan *et al.,* 2016), showed high *ZHp* signals in E9 (-6.16 and -5.07, respectively). No other ecotypes presented significant sweeps at the *BEGAIN* gene. In addition, the *DRD3* gene was detected with a high signal in E3 (*ZHp* = -5.01). It is a dopamine receptor that plays an important role in systemic arterial blood pressure and is strongly related to chicken aggressive behaviour (Li *et al.,* 2016). The gene, however, was not detected in any other ecotypes. Other genes with multiple functions that show high signals in different ecotypes are *ARHGAP22* (E3), *ZZZ3* (E9), *RNF7* (E4), and *HYDIN* (E9) (Table 5.3).

5.3.3 Important sweep regions and genes detected with *iHS* analysis

Genes overlapping sweeps with the *iHS* highest scores from each of the twelve ecotypes were investigated (Table 5.4). These include 17 genes from several sweep regions (n = 12) in ecotypes E1 and E11 with |*iHS*| ≥ 3.5 (Figure 5.7). Fourteen of these genes are located on chr10. For ecotypes E2 to E10, none of the regions showed such signal threshold; therefore, for these, only the *iHS* = 3.5 sweeps were investigated here. (Table 5.4).

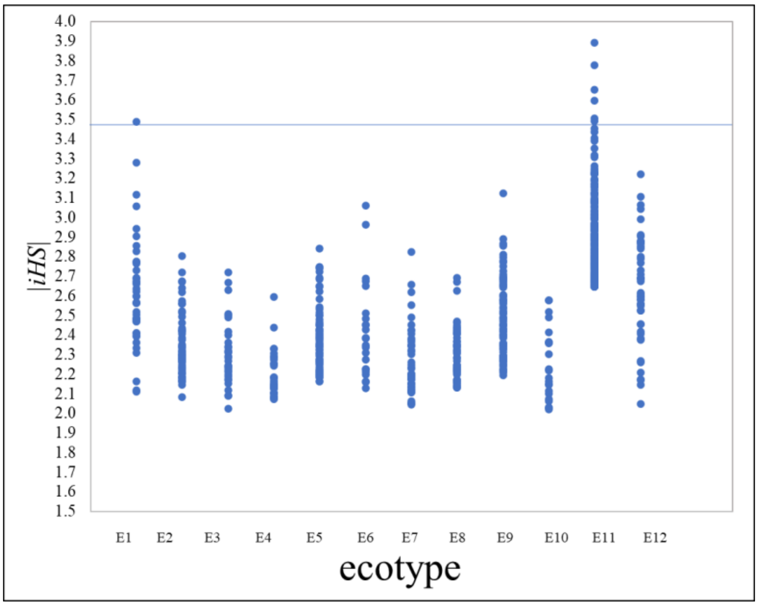


Figure 5.7 |*iHS*| values for candidate genome regions in each ecotype.
Genome regions showing sweeps with signals over a threshold |*iHS*| ≥ 3.5 (blue line) were further investigated (See Supplementary Table S5.6 – sub-tables 1 to 12).

The genes presented in Table 5.4 have several molecular functions being involved in transport, transcription, and signalling. Three genes related directly to phenotypic traits were identified. Two of these are associated with pigmentations of the plumage (*AGBL1* for the orange and red plumage colours (Xiaotong *et al.*, 2020), and *MYO5A* associated with black plumage in chicken (Mao *et al.*, 2019)) and one to dwarfism (*ACAN*) (Eberth *et al.*, 2018; Stavber *et al.*, 2020). These three genes were detected in chromosome 10 in ecotype E11. *MYO5A* and *AGBL1* did not show any |*iHS*| ≥ 2 or close to this value in any other ecotype. This finding might be reflecting an ongoing artificial selection in E11 (Arabo population) with Amhara farmers selecting birds for plumage colour (Addisu *et al.*, 2013). Also, E11, compared with other ecotypes, display high temperatures during the warmest months (32.5°C). Previous studies have shown an important impact of high temperatures on body volume reduction (Smith *et al.*, 2009; Yeates, 2009; Steinbacher *et al.*, 2017). In our case, this may translate that this gene is involved in the control of chicken size. By extension, it may explain the ongoing selection at the *ACAN* gene region for this ecotype.

Table 5.4 Genes overlapping extreme signal (|*iHS*| ≥ 3.5) or the strongest signalling *iHS* sweep in each ecotype.

GENE	Gene function / association	Reference	chr_start_stop	kb	*uiHS* ± s.d.^{a}	Eco^b
ENSGALG 00000050632	LncRNA		2_30320000 _30350000	30	2.71	E3
YES1	Known as Yamaguchi sarcoma virus oncogene with tyrosine kinase activity. Associated with lung cancer and as a predicted biomarker for therapy in humans.	Garmendia *et al.*, 2019	2_101930000 _101960000	30	3.65 ± 0.18	E11
CLUL1	It is a novel gene expressed predominantly in the retina. It is implicated in the pineal gland function between laying and broodiness periods in geese.	Zhang *et al.*, 2000 Yuan *et al.*, 2020				
ENSGALG 00000032651	Uncharacterized LOC112531902		2_139090000 _139140000	50	2.69	E8
TARBP1	Protein coding gene with methyltransferase function	Wu *et al.*, 2008	3_38540000 _38570000	30	3.05	E6
COA6	Involved in the mitochondrial respiratory chain as a member of the cytochrome c oxidase subunit family	Stroud *et al.*, 2015				
SLC35F3	It has a role as a thiamine transporter with implications in human hypertension	Zhang *et al.*, 2014				
MANEA	It is associated with panic disorder in humans. Show environmental-genomic association with low precipitation.	Jensen *et al.*, 2014 Gheyas *et al.*, 2021	3_73170000 _73220000	50	3.48 ± 0.33	E1
DPP10	It plays an important role in the neuronal excitability	Jensen *et al.*, 2014	7_29240000 _29310000	70	2.59	E4
GTDC1	Protein coding gene with transferase activity involved in neurodevelopmental disorders	Aksoy *et al.*, 2017	7_33250000 _33280000	30	3.21	E12
KPNA4	Mainly involved in intracellular protein traffic	Li *et al.*, 2011b	9_22290000 _22340000	50	2.57	E10

gga-mir-1692	miRNA					
SCARNA7	Also known as U90. It plays an important role in the methylation of RNAs as part of Box C/D RNA.	Morais *et al.*, 2021				
TRIM59	Protein coding gene with a role as regulator for innate signalling pathways	Boudinot *et al.*, 2011				
SMC4	It plays an important role in changes in chromosome structure during mitotic segregation	Kimura *et al.*, 2001				
CHRNB4	Protein-gene coding for subunits of nicotinic acetylcholine receptors.	Arvaniti *et al.*, 2016	10_4050000_4070000	20	3.89	E11
CHRNA5	Nicotine receptor family protein. Polymorphism is related with early schizophrenia and executive function.	Han, *et al.*, 2019	10_4090000_4110000	20	3.77	11
PSMA4	Proteosome involve in protein homeostasis. Down-regulated gene when embryo cells are infected with Marek's disease virus.	Zhou *et al.*, 2018				
HYKK	Protein-gene with transferase activity. Significant association with lung cancer susceptibility and pneumonia risk.	Wang *et al.*, 2017 Campos *et al.*, 2020				
ARPP19	Protein-gene coding phosphatase protein with an important role in the timing of meiosis.	Lemonnier *et al.*, 2020	10_9180000_9290000	110	3.30 ± 0.29	11
MYO5A	Myosin protein. It is a candidate gene for black plumage.	Mao *et al.*, 2019				
TM6SF1	Transmembrane protein-gene. Associated with hepatocellular carcinoma.	Tang *et al.*, 2019	10_11660000_11710000	50	3.45 ± 0.54	11
HDGFL3	Protein coding gene enhancing DNA synthesis and with an important role in cell proliferation	Haouzi *et al., 2018*				
ACAN	The encoded protein is an integral part of cartilaginous tissue. Alleles of this gene are associated with dwarfism in miniature horses and humans with short stature.	Eberth *et al.*, 2018 Stavber *et al.*, 2020	10_13530000_13610000	80	3.43 ± 0.54	11

ISG20	Interferon protein-coding gene. Regulates influenza A virus replication	Chai *et al.*, 2018				
gga-mir-1720	miRNA					
gga-mir-7-2	miRNA					
NTRK3	Neurotrophic receptors with roles in cell differentiation.	Ogura *et al., 2021*	10_13650000_13670000	20	3.40	11
			10_13780000_13820000	40	3.39 ± 0.43	11
			10_13710000_13770000	60	3.35 ± 0.25	11
AGBL1	Candidate gene to pheomelanin pigmentation in chicken.	Xiaotong *et al.*, 2020	10_14250000_14290000	40	3.59 ± 0.16	11
AGO4	Protein coding gene paying and integral role in RNA interference and RNA-mediated gene-silencing	Hauptmann *et al.*, 2014	23_4350000_440000	50	3.12	E9
CLSPN	Candidate gene that influences the function of the ovary and uterus	Yuan *et al.*, 2015				
C1orf216	Chromosome 1 Open Reading Frame					
TEAD3	Protein coding gene involved in organ size control and tumour suppression	Zhao *et al.*, 2008	26_60000_11000	50	2.82	E7
TULP1	Protein coding required for normal photoreceptor function and for the long-term survival of photoreceptor cells	North *et al.*, 1997				
FKBP5	Stress response-related gene	Lotvedt *et al.*, 2017				
DDIT3	Multifunctional transcription factor in endoplasmic reticulum stress response	Yamaguchi & Wang, 2004	33_4120000_4140000	20	2.8	E2
SLC2A1	Glucose transporter	Wagstaff *et al.*, 1995	33_6110000_6130000	20	2.84	E5

(a) S.D. values are given when a region has multiple overlapping windows.
(b) Eco: ecotype.

The gene *MANEA*, associated with panic disorder in humans (Jensen *et al.,* 2014), was identified in E1. This gene has shown an environmental genomic association with low precipitation in Ethiopian indigenous chicken (as found in chapter 4). The average value for the minimum precipitation in the driest month for E1 is 79.9 mm/m^2. However, no other ecotypes reveal signatures of selection in the same region, even though some have lower precipitation than E1 (Supplementary Table S5.8).

Other identified genes have been related to various functions (Table 5.4), e.g., *YES1* and *HYKK* associated with respiratory issues; *ISG20* and *PSMA4* involved in regulating virus replication; or *CLUL1* implicated in laying and broodiness in birds.

5.3.4 Shared sweep regions detected by *Hp* and *iHS*

In this work, two approaches were used to detect different kinds of selection events. While the *Hp* method identifies regions that have either reached fixation or near fixation (Rubin *et al.*, 2010), the *iHS* approach aims to detect ongoing selection (Voight *et al.*, 2006). Though the detection power of these methods is different, exploring selective sweeps detected simultaneously by both approaches could be interesting for two reasons. If a region is detected commonly within the same ecotype, that would indicate that it is close to reaching fixation but is not yet fixed. Also, if detected by both methods, it may act as a validation for the sweep. Moreover, if *ZHp* detects a region in one ecotype while by *iHS* in another, it may pinpoint different levels of selection in different ecotypes.

Taking all sweeps (detected by either method) from all twelve ecotypes, a correlation was calculated between *Hp* and unstandardised *iHS* (-0.043 ± 0.06, *P*-value = 0.00), and *ZHp* and standardised *iHS* (-0.18 ± 0.06, *P*-value 0.00) (Supplementary Figures S5.2, S5.3). It shows either no correlation or a low level of negative correlation between the two approaches. However, a low number of sweep regions (n = 27) were detected by both *Hp* and *iHS*, either within the same ecotype or among different ecotypes (Figure 5.8). No shared sweeps were observed in E4 and E7, whereas E3 and E5 displayed multiple shared sweeps between the two approaches (4 and 5, respectively). E6, E11, and E12 displayed only one sweep detected by both methods.

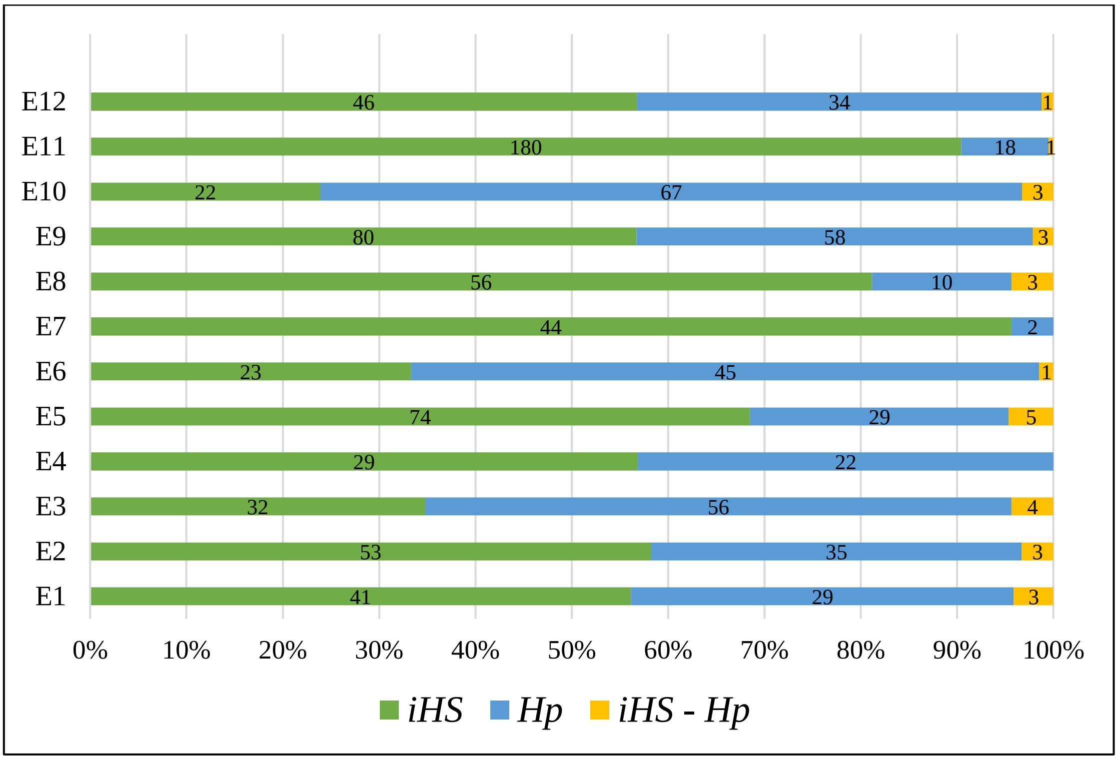


Figure 5.8 Detected sweep regions detected by *iHS*, *Hp*, and *iHS-Hp* simultaneously by ecotype.

Out of 27 commonly detected candidate sweep regions, eight were detected in the same ecotype (Figures 5.8). These regions overlap with five genes (*EFCH2*, *TSNARE1*, *TACR3*, *CHUCK*, and *AGTR1*) with multiple functions related to neuronal functions, inflammatory signals, and angiotensin function. In addition, six additional genes overlap the same region in chromosome 23 (*TCEB3*, *PITHD1*, *LYPLA2*, *GALE*, *HMGCL*, *FUCA1*). They are involved in diverse metabolic processes (Table 5.5).

Only E5 showed a sweep in the region chr2_147370000_147410000 simultaneously detected by *Hp*, and *iHS* (Table 5.5), which overlaps the gene *TSNARE1* associated with domestication (Qanbari *et al.*, 2015). Genes detected by each method found in unique specific windows per ecotype are highlighted in Supplementary Tables 5S.5 and 5S.6.

Further investigations on those genes will be needed to explore possible unique genetic prints at the ecotype level.

Table 5.5 Candidate sweep regions detected by *Hp* and *iHS* across ecotypes.

Position	**kb**	**Ecotypes**			**Genes**	**Function**	**Ref**
		Hp	***iHS***	***Hp-iHS***			
1_32940000_32960000	20	E10		E3	*ENSGALG00000049735*		
1_76440000_76470000	30	E3	E9		*ENSGALG00000011930* *ENSGALG00000026901*	OVSTL - uncharacterized protein microRNA 6606	NCBI
1_112390000_112440000	50			E2	*EFHC2*	Involved in neuron development	Gaudet *et al.*, 2011
1_120930000_120960000	30	E2, E9	E5		*ENSGALG00000049135*		
1_162440000_162460000	20	E5, E10			*ENSGALG00000047522*		
2_78320000_78340000	20			E10	*ENSGALG00000035925*	LOC428499	NCBI
2_147370000_147410000	40			E5	*TSNARE1*	Neurobehavioral gene	Quanbari *et al.*, 2015
439159999_39180000	30			E12	*TACR3*	Regulates gonadotropin-releasing hormone secretion	Topaloglu *et al.*, 2009
5_11710000_11730000	20	E11	E8		*ENSGALG00000053062*		
6_9980000_10000000	20		E5, E8	E3	*CHUK*	It plays an important role in signalling pathway activated by inflammatory cytokines, bacterial or viral products	Razani *et al.*, 2010
7_29280000_20310000	30	E2	E1		*DPP10*	Involved in neuronal excitability	Jerng *et al.*, 2007
912520000_12560000	40	E1, E8		E9	*AGTR1*	Angiotensin receptor. Contributes to pulmonary hypertension in chicken	Burks, 2011
10_6310000_6340000	30	E3	E5		*MPHOSPH10*	It plays an important role during mitosis. It is associated with growth curve parameters in chickens.	Moroudi *et al.*, 2021
					APBA2	It is involved in neurexin and neuroligin pathways in humans. It is associated with growth curve parameters in chickens.	Moroudi *et al.*, 2021
1118320000_18340000	20	E5	E3		*ZNF469*	Involve in transcriptional regulation. Associated with corneal thickness	Lu *et al.*, 2013
14_4960000_4980000	20	E12	E1		*USP22*	Involved with Telomere maintenance	Atanassov *et al.*, 2009

23_5730000_ 5750000	20			E6	*TCEB3*	General transcription elongation factor acting on RNA polymerase II transcription	Aso *et al.*, 1996
					PITHD1	Involved in megakaryocyte differentiation	Lu *et al.*, 2015
					LYPLA2	Acts on biological membranes to regulate lysophospholipids	Wepy *et al.*, 2019
					GALE	Involved in glucose and galactose metabolism	Timson, D. 2006
					HMGCL	Plays an important role in ketone body formation	Pié *et al.*, 2007
					FUCA1	Lysosomal enzyme involved in the degradation of glycoproteins and glycolipids	Ip *et al.*, 2002

5.3.5 RDA test and detection of candidate outliers SNPs

Hp and *iHS* approaches indicated candidate selective sweep regions within each ecotype as potentially associated with environmental adaptation. Since ecotypes defined in our study are a construct of a complex interaction of multiple environmental variables, we need to dissect the genotype-environment association further. Here, we employ Redundancy Analysis (RDA) for this purpose. RDA is a powerful method to detect genetic variants associated with agro-climatic variables. It is a multivariate linear regression approach, allowing simultaneous interrogation of many response variables (e.g. SNP genotype data) with many predictors (e.g. environmental variables) (Capblancq *et al.*, 2018). An original number of 277,619 SNPs overlapped the candidate sweep regions from *Hp* and *iHS* analyses. After LD-pruning, 44,202 SNPs were used. Correlation analyses were performed using the six main agro-climatic predictors to define the ecotypes (Chapter 3).

The RDA model was highly significant (F-statistics = 0.001) for the null hypothesis (no linear relationship exists between the SNP data and the environmental predictor). The model explains 1.9% of the genomic variance for the considered environmental parameters. The small proportion of genetic variance explained by the RDA model reflects that only a small proportion of the SNPs coming from the sweep regions have a linear relationship with environmental parameters. Each of the six RDA axes explained 10% to 30% of the variance captured by the RDA model (Figure 5.9c). Only the first five RDA axes were significant (P-value $\leq 0.1 - 0.01$), and these together accounted for 90% of the total captured variance by the model.

The SNP loadings (Figure 5.9d) for the significantly constrained axes were used to determine outliers or local adaptation candidates. The distribution of the chicken samples across the investigated ecotypes is presented (for RDA1 *vs* RDA2 and RDA1 *vs* RDA3) in ordination plots (Figures 5.9a-b).

RDA projection plots (Figure 5.9a-b) use scale values, where the SNPs scores were scaled by the eigenvalue's square root (Forester, 2019). As the explanatory variables are not dimensionally homogenous, this standardisation allows direct comparison of regression coefficients (ter Braak, 1994). Individuals' dispersion across the ordination space represents their relationship with the RDA axes, resulting from linear combination to the predictor variables employed. Some interesting relationships can be identified. For example, RDA1 *vs* RDA2 plot (Figure 5.9a) shows that individual genotypes from E1 are positively related to precDQ and LandUse, whereas E3 and E8 are negatively correlated with these two variables. Similarly, E12 genotypes are positively correlated with minTemp and negatively correlated with precWQ, whereas an opposite scenario for E7 is observed (Figure 5.9a). By contrast, E9 individuals are positively associated with LandUse in RDA axis 3 (Figure 5.9b).

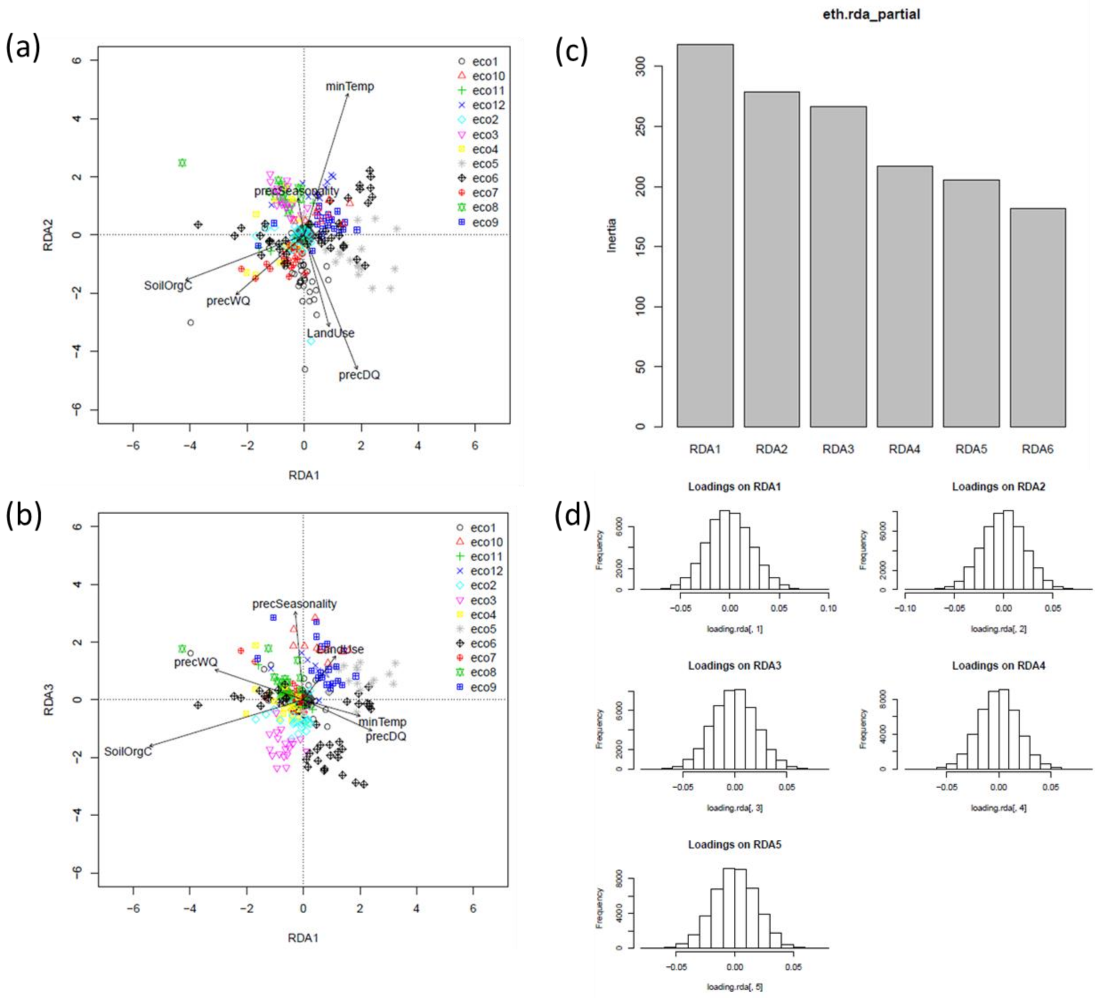


Figure 5.9 (a) Projection of SNPs and environmental variables into the RDA space RDA1-RDA2 and (b) RDA1-RDA3. (c) The six RDA axes explain variance (inertia). (d) Distribution of SNP loadings for significantly constrained axes (RDA1 – RDA5). The % of variation explained are: RDA1 (32.7%), RDA2 (19%), and RDA3 (18.1 %).

Interestingly, individuals from E6 are dispersed in subgroups for RDA1 vs. RDA2 and axes RDA1 vs. RDA3. Ecotype 6 comprises four different kebeles of populations (Mihquan, HadushAdi, Hugub, and Jarso), though MaxEnt model prediction clustered them together. For other ecotypes, no sub-clusters are evident when RDA ordination plots are explored.

Given a normal distribution of the SNP loadings in all axes (Figure 5.9d), SNPs that exceed SD > 3 (p-value = 0.0027) in both tails were extracted as outliers. With this threshold, 616 SNPs were found as outliers (Supplementary Table S5.9). The number of outliers varied in relation to the strongest correlated environmental variable. For minTemp, 113 outliers were identified, 72 for precSeasonality, 82 for precDQ, 104 for precWQ, 130 for SoilOrgC, and 115 for Land Use. The strength of the environmental correlation of the outlier SNPs was generally low to moderate ($r^2 \approx 0.1 - 0.4$) (Figure 5.10). The maximum correlation value was ~ 0.42, identified for precWQ.

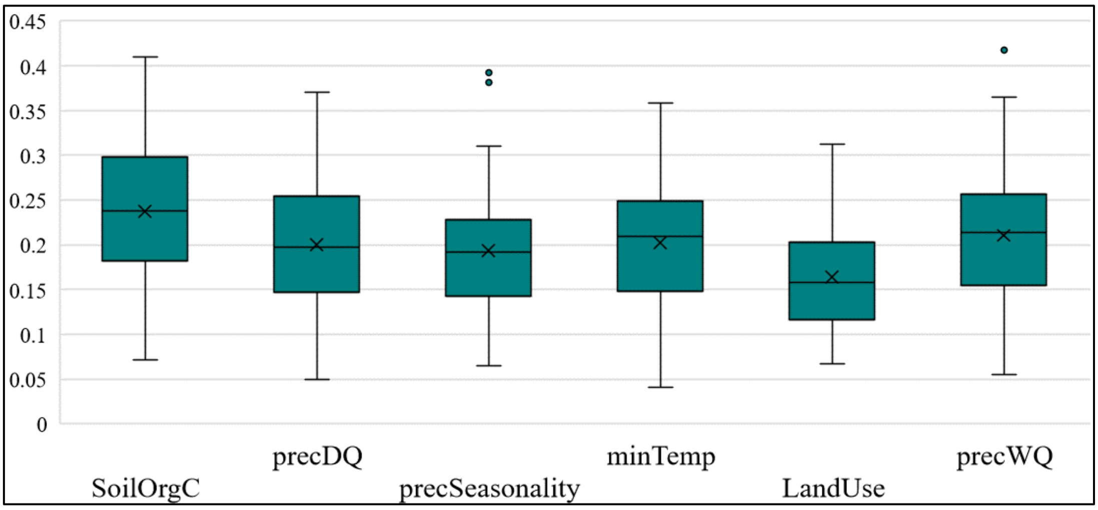


Figure 5.10 RDA correlation boxplot for identifying outliers.

The RDA analysis was performed to correlate SNP genotypes from individual samples with environmental parameters without directly considering the ecotype effect. Therefore, further investigations were made to check if the allele frequencies of ecotypes outlier SNPs vary consistently with environmental parameters. Table S5.8 presents outlier SNPs results with RDA r values ≥ 0.3 (but $r \geq 0.28$ for precDQ and Landuse). Remarkably, for most outliers, a moderate correlation ($r \approx 0.4 - 0.7$) was found between allele frequency and the ecotype average of the predictor value, showing a possible environmental cline for the SNPs associated with the investigated agro-climatic variables. Allele frequencies and correlation plots for one representative outlier per predictor are presented as an example in Figure 5.11.

In most cases, the difference between the minimum and maximum values of the allele frequencies across the ecotypes is not wide. However, a clinal increment or decrement is evident when the environmental values sequentially change among ecotypes. Correlation coefficients between allele frequencies and mean values of agro-climatic predictors support this relation (Figure 5.11). Even though the correlation is not completely linear in all cases, it still shows an interesting relationship.

Finally, further investigation was made on 32 outlier SNPs which showed relatively large environmental correlation values ($r \geq 0.3$) (Table 5.6). This resulted in two outliers for minTemp, four for precSeasonality, five for precWQ, and 14 for SoilOrgC. No outliers above the threshold ($r \geq 0.3$) were found for precipitation in the driest quarter (precDQ) and LandUse. Thus, $r \geq 0.29$ and 0.28 were respectively applied for selecting outliers for these environmental variables. With this, four outliers were chosen for precDQ and three for LandUse. Notably, all these SNPs came from *iHS*-detected sweep regions. Genes represented by these SNPs were explored to understand their possible functions in relation to the agro-climatic variables (Table 5.6).

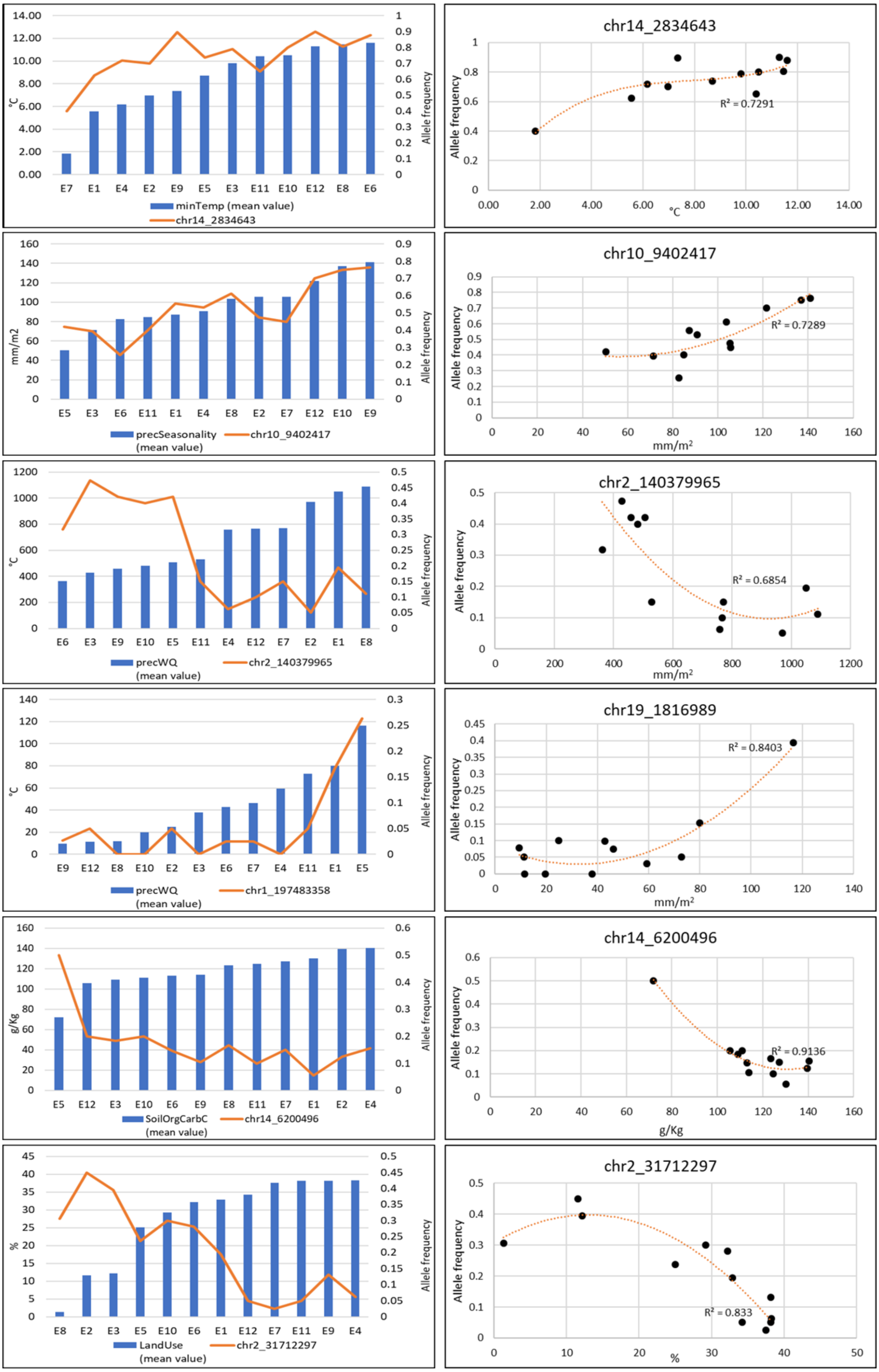


Figure 5.11 Allele frequency fluctuation and correlation with agro-climatic predictors (mean value) across ecotypes.

One outlier per predictor is shown as an example. Complete information can be found in Supplementary Table 5S.10.

Among the investigated SNPs/genes (Table 5.6), three genes (viz. *LDLRAD3*, *TNIP2*, and *GPX7*), associated with precipitation variables (precSeasonality, precWQ, and precDQ), have previously been reported to have an association with heat stress in chicken in agreement with the ability to cope with heat stress being dependent on water availability (i.e. the level of rainfall for scavenging chickens).

Small changes in rainfall combined with weather can alter the habitat moisture, particularly on the tropical rain belt, which is a dominant feature of the African climate (Jackson *et al.,* 2020). Therefore, the animal's thermal and suffocation sensation increases, although the maximum temperature ranges are not the highest in the ecotypes where the outliers were detected (E12, 11, and 7 with 30°, 32°, and 20°C respectively), but high enough to interact with precipitation and humidity to create physiological stress. Given these associations, the identified outliers are good candidates for future studies on adaptation to humid environments.

The gene *MAPK4*, a candidate for *Salmonella* resistance (Table 5.6), is associated with precSeasonality. Seasonal climate changes have proved to be related to the prevalence of infectious diseases in tropical areas. Accordingly, this predictor becomes interesting for GEA studies in search of associations with disease resistance or susceptibility.

Only one gene, candidate for immune response and heat stress (*PTPRZ1*), was not associated with a climatic predictor but with the SoilOrgC. However, since our work was performed with indigenous village chickens (predominantly scavengers and exposed to a broad range of pathogens through their food source), this finding is relevant, particularly for the immune response association.

Table 5.6 Selected RDA SNPs outliers overlapping protein-coding genes.

Associated predictor	ID_SNP outliers	Gene	Ecotype*	RDA cor	Function / association	Reference	QTLdb
minTemp	4_88536330	*PANK2*	E11	0.305	Involved in homeostasis for normal neuronal and vascular development.	(Zizioli *et al.*, 2016)	
minTemp	2_78280162	*MARCHF6*	E10	0.308	Metabolic integrator in cholesterol synthesis.	(Scott *et al.*, 2020)	
precSeasonality	1_71668364	*LRP6*	E2	0.308	Regulates muscle fibre development impacting shear force in chickens.	(Chao *et al.*, 2019a)	
precSeasonality	10_9402417	*MAPK4*	E11	0.302	It is associated with early body temperature and *Salmonella pullorum* resistance in chickens. Differential expression after hypoxia stress.	(Li *et al.*, 2020) (Yuan *et al.*, 2020)	
precSeasonality	4_84643310	*CTBP1*	E11	0.310	It modulates hypoxia-induced apoptosis in pituitary cells.	(Dorman *et al.*, 2012)	
precSeasonality OrgSoilC	5_19554343	*LDLRAD3*	E12	0.307 0.334	Associates with heat stress in broiler chicken.	(Jastrebski *et al.*, 2017)	
precWQ	2_117749087	*RDH10*	E2	0.417	An upregulated gene for abdominal adipose tissue of chickens.	(Huanxian *et al.*, 2012)	
precWQ	10_13599109	*DET1*	E11	0.348	Upregulated gene in domestic chicken breeds in comparison to wild species RJF.	(Quinghe *et al.*, 2012)	
precWQ	21_4353990	*FBXO42*	E8	0.310	Binding protein	(Sun *et al.*, 2009)	Gizzard_weight
precWQ	4_3627776	*HS6ST2*	E10	0.330	It is associated with growth and carcass traits in chicken.	(Wang *et al.*, 2018)	Body_weight_28 _days
precWQ	4_82439583	*TNIP2*	E11	0.306	It is associated with growth response to environmental stress, inflammation/oxidative/metabolic stress.	(Fleming *et al.*, 2016)	Feather_colour _extended_black
precDQ	1_197483358	*DCHS1*	E7	0.371	Calcium-dependent cell-cell adhesion molecules.	(Durst *et al.*, 2015)	
precDQ	20_1187011	*ERGIC3*	E4 E7	0.294	Cytoplasmic transport protein.	(Breuza *et al.*, 2004)	
precDQ	8_24798255	*GPX7*	E7	0.296	It is associated with antioxidant activity caused by heat stress in chickens.	(Del Vesco *et al.*, *2015*)	Feathered_feet
precDQ SoilOrgC	19_1816989	*AUTS2*	E8	0.344 0.372	It is implicated in neurological development.	(Gao *et al.*, 2014)	Feed_intake

SoilOrgC	1_23065714	*PTPRZ1*	E7	0.302	Associates with heat stress response and immune stimulation in chickens.	(Monson *et al.*, 2018)	Feather_colour_ extended_black
	1_23140159		E7	0.307			
	1_23141490		E7	0.405			
	1_23164664		E7	0.335			
SoilOrgC	10_7190022	*MYO1E*	E11	0.332	Involved in intracellular movement and membrane trafficking.	(Senna-Cherchi *et al.*, 2011)	
SoilOrgC	19_1819556	*LMF1*	E8	0.363	Involved in lipoprotein liver metabolism in chickens.	(He *et al.*, 2018)	
SoilOrgC	3_16891572	*GPCPD1*	E12	0.410	Important gene in determining muscle growth and feed efficiency in broilers.	(Kong *et al.*, 2017)	
SoilOrgC	3_38560061	*TARBP1*	E6	0.306	RNA binding protein.	(Christensen *et al.*, 2007)	Comb_weight
SoilOrgC	5_35557653	*NPAS3*	E7	0.361	This gene appeared to be responsible for domestication-related behavioural aspects.	(Hou *et al.*, 2020)	
SoilOrgC	5_39594296	*ADCK1*	E8	0.327	Involved in mitochondrial functions.	(Yoon *et al.*, 2019)	
SoilOrgC	6_28550216	*NHLRC2*	E12	0.307	Associated with body weight in chickens	(Li & Li, 2019)	Ovary_weight
SoilOrgC	7_30965068	*THSD7B*	E12	0.377	Candidate gene associated with albumin quality in chicken	(Qu *et al.*, 2019)	Ovary_weight Comb_weight
	7_30967748		E12	0.320			
SoilOrgC	8_6198062	*STX10*	E12	0.382	Involved in fusion events at the Golgi apparatus.	(Ganley *et al.*, 2008)	
LandUse	17_7056075	*GFI1B*	E11	0.312	Transcriptional regulator protein.	(Fuchs *et al.*, 1997)	Feather_colour_ extended_black
LandUse	17_7239853	*SLC2A6*	E11	0.296	A candidate gene for blood metabolites in chickens.	(Javanrough-Aliabad *et al.*, 2018)	
LandUse	2_31712297	*OSBPL3*	E4	0.282	Differentially expression of the gene on laying hen liver, closely related to fat metabolism.	(Li *et al.*, 2015)	

* The ecotypes from which the SNPs originally overlapped with sweep regions

We also find that some chicken production-related genes, e.g., *LRP6* and *HS6ST2,* are associated with precipitation variables, indicating the effects of precipitation variables on the growth and welfare of scavenging chickens. The largest number of high correlation outliers were detected for the OrgSoilC variable (Table 5.6). Interestingly, most of the genes associated with these candidates are involved in various lipoprotein mechanisms (viz. *PTPRZ1, MYO1E, LMF1, GPCPD1, TARBP1, NPAS3, ADCK1, NHLRC2, THSD7B,* and *STX10*). In particular, three genes - *GPCPD1*, a metabolite interconversion enzyme, *NHLRC2*, a protein modifying enzyme, and *TSHD7B*, a cell adhesion protein - play an important role in food conversion protein metabolism, and they are candidates for muscle growth, body weight, and albumin quality. These outlier genes were detected from E12 and E8, which showed medium to high soil organic carbon (Supplementary Table S5.8).

These findings have important implications for breeding improvement efforts on African poultry used for backyard farming. For example, these genes can be targeted in a selection programme to improve the scavenging performance of chicken (e.g. improving feed conversion efficiency from scarcely available natural food). These findings also show that our innovative framework helps identify potentially more efficient ecotypes in utilising natural food. Furthermore, the biological functions associated with these genes confirm that our assumption about the significance of soil carbon content for the scavenging adaptation in indigenous chickens is justified, and ENM has successfully captured this important predictor.

Investigation of the molecular functions, biological processes, and cell components of these 28 protein-coding genes (using PANTHER Classification System) show their involvement in wide-ranging molecular functions (Figure 5.12). Eight genes (*GFI1B, MYO1E, DET1, DCHS1, STX6, NPAS3, TNIP2,* and *OSBPL3*) are related to binding functions. Eight genes (*MYO1E, GPX7. TARBP1, PANK2, RDH10, HS6ST2, PTPRZ1, and MAPK6*) are related to catalytic activity. Two genes (*DET1*, *STX6*) are involved in molecular adaptor activity. Two genes (*GFI1B* and *NPAS3*) have molecular regulatory functions, and one gene (*OSBPL3*) is involved in transporter activity. Few significant results with multiple genes involved were obtained from enrichment analysis (Supplementary Table S5.11). After correcting for multiple testing, only two pathways showed *P*-value ≤ 0.05. These two are the insulin signalling pathway (*P*-value: 0.044 for *SREBF1*|*MTOR* genes) and herpes simplex virus-1 infection (*P*-value: 0.044 for *MAVS*|*MTOR* genes). These results may have not included other real causal SNPs as a previous LD pruning approach was applied removing potential targets for the analysis. However, the SNPs tested can act as tag SNPs (of the causal loci) for any future work.

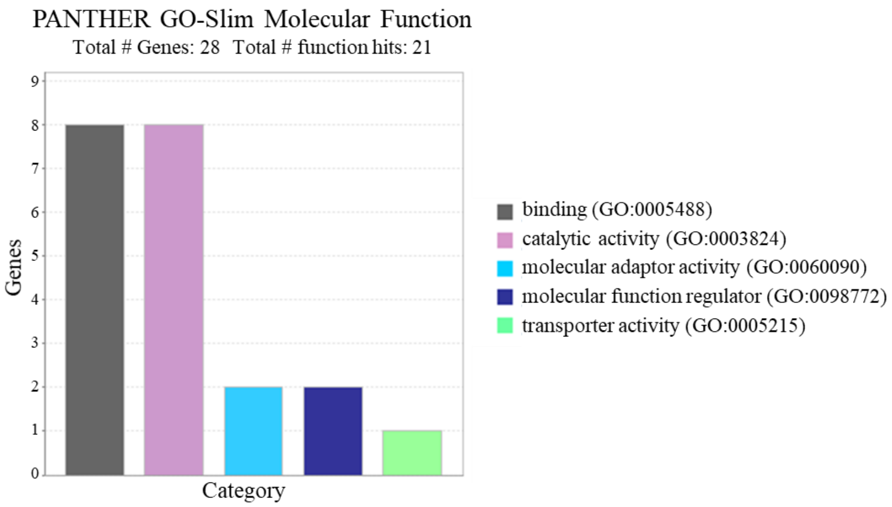


Figure 5.12 Molecular functions for 28 RDA outlier investigated. (panther.org)

5.4 Discussion

This chapter presents a novel approach to study environmental adaptation. Previously, different studies have investigated genomic adaptation to particular climatic or environmental factors but have not included detailed characterisation of the environment to define ecotypes. In this research, based on the habitat characterisation and applying ENM, ecotype definition was produced to cluster chicken populations based on the environment (Chapter 3). These ecotypes capture the complex interactions of major environmental variables within the local agro-ecology and offer the opportunity to explore their effects on the genome level. Exemplified by Ethiopian village chicken, this novel approach provides a new tool for identifying, characterising, utilising, and conserving livestock diversity. Furthermore, this is the first study on a livestock species that simultaneously considers a large spectrum of environmental variables.

The ecotypes were then used for dissecting the genetic basis of environmental adaptation using multiple complementary approaches: *Hp* for capturing selective sweep regions possibly near fixation (Rubin *et al.,* 2010), *iHS* to detect sweeps that are yet to reach fixation (Voight *et al.,* 2006), and finally, RDA to identify the linear relationship of genotypes with specific environmental variable gradients.

Among the various environmental selection pressures which have shaped the animal genome, climatic variables have been particularly under the research scope (Frichot *et al., 2013;* Lv *et al.,* 2014). For instance, studies on thermotolerance responses have followed the increasing concerns of the impact of climatic change on livestock production (Shilin *et al.,* 2020). Thus, an important advantage of the approach applied here is the prior identification of the environmental variables with a major impact on an ecotype as a preliminary step to genomic analysis. Moreover, combining sweep analysis with RDA permits identifying regions under positive selection and determines which variables or closely associated linked variables have been selected.

Several studies have discussed the different strategies to identify selection signatures in livestock species (de Simoni Gouveia *et al.,* 2014; Eusebi *et al.,* 2020; Saravanan *et al.,* 2020). Also, demographic events, recombination rate, and selection strength may have a differential influence on specific populations (e.g. Weigand & Leese, 2018). The genomic approach used by the present study based on the constructed ecotypes, instead of investigating individual populations, offers many benefits. First of all, clustering multiple populations based on environmental similarity permitted focusing on the genomic variation associated with specific ecological drivers, reducing false-positive candidates from demographic events (e.g. genetic flow, founder effect, bottlenecks, etc.). Second, combining multiple populations in most ecotypes is expected to reduce population structure in the signature analysis. Finally, it has allowed us to explore the genotype-environment association across a complex environmental landscape instead of using a simplified association between a single predictor and a genotype.

We have used a ranking-based outlier approach for finding candidate sweep regions. Since our analysis is not based on individual SNPs but based on window approach where information from multiple SNPs is combined, this should reduce false positives. Besides we have applied quite stringent criteria in sweep calling ($P<0.01$ and then further filtration based on tests statistics). This study detected a smaller number of selective sweeps that have reached or near fixation (based on the *Hp* approach). Interestingly, no SNPs from *Hp*-based sweeps were outliers with a strong correlation to environmental variables in RDA. This is probably because these SNPs are fixed in most or many ecotypes and the results of old selection events (e.g. improvement throughout domestication (viz. *BCDO2, TSHR*, *GTF2A1, TSNARE1*)). These will not vary in frequencies with environmental gradient and thus could not be captured by RDA.

The *iHS* method spotted most outliers in the RDA. What makes these findings more interesting is that RDA candidate SNPs are under selection over many ecotypes that are not necessarily close at geographic level but are affected by similar ranges of the response variables associated with them. Therefore, the *iHS* detection method followed by RDA is a more effective approach for detecting environmental association. This finding is relevant since the results obtained here permitted identifying the best methods to be applied in the proposed framework to study environmental adaptation.

5.5 Conclusion

Considering the complex interaction of agro-climatic variables in specific ecotypes, this study shows a new direction of analysing environmental adaptation by first classifying populations into ecotypes and then using the defined ecotypes to investigate signatures of selection and genotype-environment association. Moreover, we presented here an efficient signature of selection approach for identifying genomic regions and candidate-associated genomic markers associated with environmental adaptation. These results have important implications for conserving important genetic diversity, for management decisions about breed suitability for specific regions, and breeding improvements, especially regarding climate resilience traits.

Chapter 6

Conclusions and Perspectives

This final chapter highlights the major achievements from this PhD research along with their implications. The chapter also discusses the limitations of the work and provides some future research directions.

The main goal of this thesis was to dissect the genetic basis of tropical environmental adaption in African indigenous chickens. Although genetic adaptation has been an important research interest to understand how species respond to environmental conditions (Fox *et al.,* 2018; Julian & Moore, 2019), large-scale genomic studies on indigenous livestock, particularly on chicken, are still very limited. Previous studies on livestock were often limited in their scope due to small sample size, lack of genome sequence data, or not having extensive coverage of the genome (e.g. used only genotype data from SNP chips) (Gheyas *et al.,* 2015; Fleming *et al.,* 2016; Fleming *et al.,* 2017; Strillacci *et al.,* 2018; Elbeltagy *et al.,* 2019; Luo *et al.,*2020a).

Besides, dissecting environmental adaptation itself is a very challenging task. Many intervening factors can contribute to this, such as a wide phenotypic diversity across environmental ranges, making it difficult to identify ecotypes (Ehsaninia *et al.,* 2020). Also present are complex interactions between environmental factors (LaFleche & Waterman), gene-environment interactions (Gilad *et al.,* 2006), and interactions among genes (Sandoval-Castillo *et al.,* 2020). However, the need for research aiming to elucidate genotype-environment interactions in livestock species is increasingly becoming more relevant to our world's rising climate crisis. Therefore, methodological innovations are required, along with large environmental and genomic data to comprehend the current and future species' responses to modified environmental conditions. This PhD study attempted to integrate interdisciplinary analytical approaches (ecological and genomic) to improve the power of the analysis.

Employing ENM, this study proposes a new framework for defining livestock ecotypes (Chapter 3), including characterization of their habitats as well as genomes in relation to environmental pressures. No previous studies have done this before to such a level of detail in livestock. Although the method has been developed using Ethiopian indigenous chickens, it will be similarly applicable to any other species. Moreover, to dissect the genetic basis of adaptation, a rather comprehensive strategy was employed involving comparative genomics of populations from extreme environments (*Fst, XP-EHH* in Chapter 4*)*, within ecotype analysis (*iHS*, and *Hp* in Chapter 5), and combined analysis of all populations across the environmental gradients (RDA in Chapter 5). These analyses permitted identifying candidate genes related to important environmental stressors in the African chicken's agro-ecologies. While RDA identifies only candidates showing linear association across environmental gradients (Capblancq *et al.,* 2018), the other analyses can capture linear and non-linear associations (Luo *et al.,* 2020b). Particularly, extreme population comparison can identify regions under positive selection due to extreme environmental pressures, representing rapid response to stress factors. Therefore, the genomic analysis approaches presented here are complementary as well as confirmatory in certain cases. Their combined use in the present study has increased the power of detecting candidate regions/genes/variants, showing different types of environmental association (linear and non-linear), different stages of selection (already fixed and ongoing), and different selection strengths (strong to weak). Specifically, our study has shown that using RDA and candidate selective sweeps permitted us to identify signals with weak to moderate effects. This is particularly important for studying local adaptation in environments where climate change has the biggest impact, in other words, where environmental stressors can rapidly change and create complex genotype-environmental interactions over time.

Using this comprehensive analysis, a set of candidate genes have been identified in relation to different environmental stress factors, including adaptation to extreme temperatures, rainfall patterns, and high altitude. Remarkably, adaptation to scavenging conditions has also been identified using environmental proxies that affect the abundance and availability of food for foraging chickens. This is the first study to report candidate genomic regions for scavenging adaptation, which will have important implications for smallholder poultry farmers raising chickens in scavenging/semi-scavenging conditions. To the best of our knowledge, this is by far the largest tropical adaptation study on chicken involving whole-genome sequence data as well as the most comprehensive one in terms of methodological applications.

Tackling false discovery rate (FDR) constitutes an important component in any genomic analysis, like the selection signature analyses applied in our study. Different approaches are available to tackle the issues with some widely used approaches including taking either the top 1% (or 5%) regions as candidates or applying some stringent thresholds to determine outliers. In our study, we have combined both, first by applying an empirical *P-value* < 0.01, corresponding to the selection of the top 1% ranked windows, - followed by applying further stringent filtration on specific selection signature metric. Using the empirical $P < 0.01$, we detected about 910 - 930 windows as candidate sweeps in different analyses as about 91,000 - 93,000 windows were analysed genome-wide. It translates into approximately nine windows as false positives (FDR) in each analysis. However, we expect the number of false positive to be much lower as further filtration thresholds were applied based on standardized values or Z-scores of different matrices (*Hp*, *Fst*, *XPEHH* and *iHS)*. For normally distributed variables, the Z score itself can provide a prediction of false positive rate (based on Z-score probability table) but in our case most distributions were not normal. Therefore no attempt is made here to estimate

FDR based on the Z-score. The calculation of Z scores, however, was necessary for applying the same metric-specific threshold across different populations or ecotypes.

The study, however, is not without limitations and further analysis will be required to fill the gaps arising from such limitations. For instance, the study could not include any phenotypic data, and therefore, it was not possible to assess the phenotypic response of tropical adaptation in the studied samples. Hence, future studies need to incorporate relevant phenotypic data (whole animals and cellular/tissue studies) to assess their association with the candidate genes/regions and environmental stressors. The current study performed analysis based on SNP variants only, mainly because they provide a dense coverage of the genome which is important for selection signature analysis when sample size is limited. However, there are other types of variants in the genome which can be functionally important such as short Insertions and Deletions (INDELS), long Structural Variants (SVs), and retroviral elements, which are also widespread in the chicken genome and can have important adaptive roles (Ratan *et al.*, 2015; Platt II *et al.*, 2018; Heller & Vingron, 2019). For instance, previous studies have shown the importance of SVs in different traits, including adaptive traits with ecological and evolutionary implications (Wellenreuther *et al.*, 2019). They may contribute to important processes and profoundly affect gene functionality (Ha *et al.*, 2009). SVs can together cover larger regions than SNPs. They are present at significant frequencies in many populations, and it may be expected that the advances in genomics and bioinformatics will catalyse a growing number of studies reporting the involvement of complex structural variants with particular traits (Xia *et al.*, 2017; Sanchis-Juan *et al.*, 2018). While all these variants are valuable, particularly calling and genotyping SVs and INDELS is more challenging and time consuming, which is one of the reasons why those variants were no included in this study. In case of SNPs, we applied a sophisticated VQSR approach which is not possible for SVs and INDELS as there are no validated set of this type of variants yet. Besides, calling SVs would require long read data for validation. Besides their limitations, is important that future studies can include this type of variants.

This study provides a number of candidate genes detected by different methods. However, these need to be validated in future studies. For example, we have reported common detected genes in environmental extreme groups (chapter 4) and ecotypes (chapter 5), but also, it is necessary to confirm their selection in other independent populations or species living under similar environmental stressors. Their expression can be then validated *via* gene expression studies (transcriptomics, metabolomics, qPCRs, etc.) or following gene editing approaches (e.g. CRISPR/Cas9) (Chao *et al.*, 2019b) and *via* genomic feature models (GFM) (Rohde *et al.*, 2018; Sehgal *et al*., 2019).

The findings and tools generated in the course of this PhD work have many implications for livestock breeding, conservation and management. The proposed framework exemplifies an attractive opportunity for performing a detailed exploration of the environment and a comprehensive analysis of genetic variants involved in agro-climatic adaptation. These might represent a crucial tool for smallholder farming productivity, particularly those relying on indigenous or/and native populations at the local level. For example, the ACGG programme (https://www.ilri.org/research/projects/african-chicken-genetic-gains-acgg) and chicken breeding companies working with ACGG could use this genomic information to develop improved breeds with better adaptation providing smallholder farmers with improved genetic resources.

Furthermore, by integrating ecological concepts with genomic approaches, the methods presented in the thesis open a further opportunity for interdisciplinary research, with important implications at the conservation and management levels. The approach can be employed to study the impact of climate change on ecotypes and explore ecotype habitats and disease pre-disposition based on environmental characteristics. It contributes to potentially overcoming some of the challenges in landscape genetics applications (Balkenhol *et al.*, 2016) by integrating into a single conceptual framework all aspects of the research to investigate the landscape characteristic’s effect on evolutionary processes. Moreover, with the vast generation of genetic sequencing data and the free accessible climatic data for the last decades, our study illustrates the opportunity of developing conservation and sustainable production programs for livestock species incorporating this information for management purposes.

References

A

Abioja, M.O., & Abiona, J.A. 2021. Impacts of climate change to poultry production in Africa: Adaptation options for broiler chickens. *In: African Handbook of Climate Change Adaptation.* (eds.), W. Leal Filho et al. Springer. Switzerland.

Addisu, J., Hailu, M, & Zewdu, W. 2013. Indigenous chicken production system and breeding practice in North Wollo, Amhara Region, Ethiopia. *Poultr Fish Wild Sci,* 1 (2): 1000108. doi:10.4172/pfw.1000108.

Afsal, A., Sejian, V., Bagath, M., *et al*. 2018. Heat stress and livestock adaptation: neuroendocrine regulation. *Int J Vet Anim Med,* 1 (2): 108.

Akey, J.M. 2002. Interrogating a High-Density SNP map for signatures of natural selection. *Genome Res,* 12, 1805-1814. doi: 10.1101/gr631202.

Aksoy, I., Utami, K.H., Winata, C.L., *et al*. 2017. Personalized genome sequencing coupled with iPSC technology identifies *GTDC1* as a gene involved in neurodevelopmental disorders. *Hum Mol Genet,* 26 (2): 367-382. doi:10.1093/hmg/ddw393.

Alemu, D., Degefe, T., Ferede, S., *et al*. 2008. Overview and background paper on Ethiopia’s Poultry sector: Relevance for HPAI Research in Ethiopia. Africa/Indonesia Team Working Paper.

Alexander, D.H., Novembre, J., & Lange, K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res,* 19: 1655-1664.

Al-Jumaili, A.S., Boudali, S.F., Kebede, A., *et al.* 2020. The maternal origin of indigenous domestic chicken from the Middle East, the north and the horn of Africa. *BMC Genet,* 21 (30). doi:10.1186/s12863-020-0830-0.

Al-Zahrani, K., Baron, K., & Sabourin, L. 2013. Ste20-like kinase SLK, at the crossroads. *Cell Adh Migr,* 7 (1): 1-10. doi:10.4161/cam.22495.

Ardlie, G.D., Kruglyak, L., & Seielstad, M. 2002. Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet,* 3: 299-309.

Asefa, M., Cao, M., He, Y., *et al*. 2020. Ethiopian vegetation types, climate and topography. *Plant Div,* 42 (4): 302-311. doi:10.1016/j.pld.2020.04.004.

Aso, T., Haque, D., Fukudome, K., *et al*. 1996. A human cDNA encoding the 110-kDa A subunit of RNA polymerase II transcription factor elonging. *Gene,* 168 (2): 277-8. doi:10.1016/0378-1119(95)00750-4.

Arvaniti, M., Jensen, M., Soni, N., *et al*. 2016. Functional interactions between Lypd6 and nicotinic acetylcholine receptors. *J Neurochem.* 138 (6): 806-20. doi:10.1111/jnc.13718.

Atanassov, B., Evrard, Y., Multani, A., *et al*. 2009. *Gcn5* and *SAGA* regulate shelterin protein turnover and telomere maintenance. *Cell,* 35: 352-364.

B

Balkenhol, N., Waits, L.P., & Dezzani, R.J. 2009. Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. *Ecography,* 32: 818-830.

Balkenhol, N., Cushman, S.A., Storfer, A., *et al*. 2016. Introduction to landscape genetics - concepts, methods, applications. In *Landscape Genetics: concepts, methods, applications,* by N., Cushman, S.A., Storfer, A.T., and Waits, L.P. Balkenhol, 1-7. U.K.: John Wiley & Sons Ltd.

Balloux, F. & Lugon-Moulin, N. 2002. The estimation of population differentiation with microsatellite markers. *Mol Ecol,* 11: 155-165.

Barra, V., & Fachinetti, D. 2018. The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA. *Nature Comm,* 9: 4340.

Batjes, N.H. 2000. Global Data Set of Derived Soil Properties, 0.5 - degree grid (ISRI-WISE). ORNL DAAC. Oak Ridge, Tennessee, USA. Accessed January 2019. HTTP://dx.doi.org/10.3334/ORNLDAAC/546.

Benjamini, Y., & Speed, T.P. 2012. Summarizing and correcting the GC content bias in high-throughput sequencing. *Nucleic Acids Res*, 40(10): e72. doi: 10.1093/nar/gks001.

Bennett, E., Mandel, U., Clausen, H., *et al*. 2012. Control of mucin-type O-glycosylation: A classification of the polypeptide GalNAc-transferase gene family. *Glycobiology*, 22 (6): 736-756. doi:10.1093/glycob/cwr182.

Bernard, B., Nigam, N., Burkitt, K., *et al*. 2021. *SMYD3*: a regulator of epigenetic and signalling pathways in cancer. *Clin Epigen*, 13: 45. doi:10.1186/s13148-021-01021-9.

Bettridge, J., Psifidi, A., Terfa, Z., *et al*. 2018. The role of local adaptation in sustainable village chicken production. *Nat Sustain*, 1 (10): 574-582. doi:10.1038/s41893-018-0150-9.

Bhakta, M.S., Jones, V.A., & Vallejos, C.E. 2015. Punctuated distribution of recombination hotspots and demarcation of pericentromeric regions in *Phaseolus vulgaris* L. *PLoS One*, 10(1): 1-20.

Biswas, S., & Akey, J. 2006. Genomic insights into positive selection. *Trends Genet*, 22 (8): 437-446. doi:10.1016/j.tig.2006.06.005.

Boettcher, P.J., Hoffman, I., Baumung, R., *et al.* 2015. Genetic resources and genomic for adaptation of livestock to climate change. *Front Genet*, 5: 1-3. doi:10.3389/fgene.2014.004161.

Boschiero, C., Monteiro-Moreira, G.C., Gheyas, A.A., *et al.* 2018. Genome-wide characterization of genetic variants and putative regions under selection in meat and egg-type chicken lines. *BMC Genom*, 19, 83. doi:10.1186/s12864-018-4444-0.

Boudinot, P., van der Aa, L.M., Du Pasquier, L., *et al*. 2011. Origin and Evolution of *TRIM* proteins: New insights from the complete *TRIM* repertoire of Zebrafish and Pufferfish. *PLOS One* 6 (7): e22022. doi:10.1371/journal.pone.0022022.

Bouret, V., Dionne, M., & Bernatchez, L. 2014. Detecting genotypic changes associated with selective mortality at sea in Atlantic salmon: polygenic multilocus analysis surpasses genome scan. *Mol Ecol,* 23: 4444-4457.

Bradie, J., & Leung, B. 2017. A quantitative synthesis of the importance of variables used in MaxEnt species distribution models. *J Biogeogr,* 44: 1344-1361.

Breuza, L., Halbeisen, R., Jeno, P., *et al.* 2004. Proteomics of endoplasmic reticulum - Golgi intermediate compartment (ERGIC) membranes from brefeldin A-treated HepG2 cells identifies *EGIC-32,* a new cycling protein that interacts with human Erv46. *J Biol Chem,* 279 (45): 47242-53. doi:10.1074/jbc.M406644200.

Browning, S.R., & Browning, B.L. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet,* 81: 1084-1097.

Burks, J.R. 2011. Sequence analysis of the angiotensin II type 1 receptor (AGTR1) gene for mutations contributing to pulmonary hypertension in the chicken (Gallus gallus). *The University of Arkansas Undergraduate Res J,* 12: Art.9.

Burt, D.W. 2005. Chicken genome: current status and future opportunities. *Genome Res,* 15:1692-1698. doi: 10.1101/gr.4141805.

C

Campos, A., Fang Kho, P., Vázquez-Prada, K., *et al*. 2020. Genetic susceptibility to pneumonia: A GWAS meta-analysis between UK Biobank and FinnGen. *midrib,* doi:10.1101/2020.06.22.20103556.

Capblancq, T., Luu, K., Blum, M., *et al*. 2018. Evaluation of redundancy analysis to identify signatures of local adaptation. *bioRxiv.* doi:10.1101/258988.

Capeau, J. 2007. The story of adiponectin and its receptors AdipoR1 and R2: to follow. *Journal of Hepatology,* 47: 736-738. doi:10.1016/j.jhep.2007.06.002.

Cardoso-Leite, R., Vilarinho, A.C., Novaes, M.C., *et al*. 2014. Recent and future environmental suitability to dengue fever in Brazil using species distribution model. *Trans R Soc Trop Med Hyg,* 108: 99-104.

Carpenter, D., Ringrose, C., Leo, V., *et al*. 2009. The role of CACNA1S in predisposition to malignant hyperthermia. *BMC Med Genet,* 13 (10): 104. doi:10.1186/1471-2350-10-104.

Cedraz, H., Gromboni, J.G.G., Garcia, A.A., *et al.* 2017. Heat stress induces expression of HSP genes in genetically divergent chickens. *PLoS One,* 12.

Chai, W., Li, J., Shangguan, W., *et al*. 2018. Lnc-ISG20 inhibits influenza A virus replication by enhancing *ISG20* expression. *J Virol,* 92 (16).

Changgeng Y, Ming J, Hua W., *et al*. 2015. Analysis of differential gene expression under low-temperature stress in Nile tilapia (*Oreochomis niloticus*) using digital gene expression. *Gene,* 546: 134-140.

Chao, L., Niu, S., Yan, S., *et al*. 2019a. Low-density lipoprotein receptor-related protein 1 regulates muscle fibre development in cooperation with related genes to affect meat quality. *Poultr Sci,* 98 (9): 3418-3425.

Chao, T., Liu, Z., Zhang, Y., *et al*. 2019b. Precise and rapid validation of candidate gene by allele specific knockout with CRIPSR/Cas9 in wild mice. *Front Genet,* 10: 124. doi: 10.3389/fgene.2019.00124.

Charlesworth, B. 2007. A hitch-hiking guide to the genome: A commentary on 'The hitch-hiking effect of a favourable gene' by John Maynard Smith and John Haigh. *Genet Res,* 89: 389-390.

Christensen, H., Daher, A., Soye, D., *et al*. 2007. Small interfering RNAs against the TAR RNA binding protein, *TRBP*, a Dicer cofactor, inhibit human immunodeficiency virus type 1 long terminal repeat expression and viral production. *J Virol,* 81 (10): 5121-31. doi:10.1128/JVI.01511-06.

Clifford, M., Athar, W., Leonard, C., *et al*. 2014. EphA7 signalling guides cortical dendritic development and spine maturation. *PNAS,* 111 (13): 4994-4999. doi:10.1073/pnas.1323793111.

Cong R.G., & Brady, M. 2012. The interdependence between rainfall and temperature: coupla analyses. *Sci World J,* doi:10.1100/2012/405675.

Coop, G. Witonsky, D., Di Rienzo, A., *et al*. 2010. Using environmental correlations to identify loci underlying local adaptation. *Genetics,* 185: 1422-1423.

Coullin P, Crooijmans RP, Groenen MA, *et al*. 2002. Assignment of FUT8 to chicken chromosome band 5q1.4 and to human chromosome 14q23.2→q24.1 by in situ hybridization. Conserved and compared synteny between human and chicken. *Cytogenet Genome Res,* (97): 234-238.

Csala, A., Voorbraak, F., Zwinderman, A., *et al*. 2017. Sparse redundancy analysis of high-dimensional genetic and genomic data. *Bioinformatics,* 33 (20): 3228-3234. doi:10.1093/bioinformatics/btx374.

D

DAGRIS. 2007. Domestic Animal Genetic Resources Information System (DAGRIS). Edited by Y. Mamo, B. Asrat & T. Dessie Eds. S. Kemp. Addis Ababa. Accessed September 13th, 2020. http://dagris.ilri.cgiar.org.

Dai, Q., Ren, A., Westholm, J., *et al*. 2015. Common and distinct DNA-binding and regulatory activities of the BEN-solo transcription factor family. *Genes Dev,* 29 (1): 48-62. doi:10.1101/gad.252122.114.

Danecek, P., Auton, A., Abecasis, G., *et al*. 2011. The variant call format and VCFtools. *Bioinformatics,* 27 (15): 2156-2158.

Derks, M., Megens, H-J., Bosse, M., *et al*. 2018. "A survey of functional genomic variation in domesticated chickens. *Genet Sel Evol* 50: 17. doi:10.1186/ s12711-018-0390-1.

Dessie, T. 2011a. Why chicken research for development? *ILRI*, International Livestock Research Institute, Issue Brief 6.

Dessie, T., Taye, T., Dana, N., *et al*. 2011b. Current state of knowledge on phenotypic characteristics of indigenous chickens in the tropics. *Poult Sci J,* 67 (3): 507-516. doi:10.1017/S0043933911000559.

Desta, T., Dessie, T., Bettridge, J., *et al*. 2013. Signature of artificial selection and ecological landscape on morphological structures of Ethiopian village chickens. *Anim Genet Resour,* 52: 17-29.

DePristo, M., Banks, E., Poplin, R., *et al*. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet,* 43: 491-498.

de Simoni Gouveia, J., Barbosa da Silva, M., Rezende Paiva, S., *et al*. 2014. Identification of selection signatures in livestock species. *Genet Mol Biol,* 37 (2): 330-342.

de Wit, J. & Ghosh, A. 2016. Specification of synaptic connectivity by cell-surface interactions. *Nat Rev Neurosci,* 17 (4). doi:10.1038/nrn.2015.3.

Del Vesco, A., Gasparino, E., de Oliveira-Grieser, D., *et al*. 2015. Effects of methionine supplementation on the expression of oxidative stress-related genes in acute heat stress-exposed broilers. *Brit J Nutr,* 113: 549-559. doi:10.1017/S000711451 4003535.

Diao, S., Huang, S., Chen, Z., *et al*. 2019. Genome-wide signatures of selection detection in three South China Indigenous Pigs. *Genes,* 10 (5): 346. doi:10.3390/ genes10050346.

Didychuck, A., Butcher, S., & Brow, D. 2018. The life of U6 small nuclear RNA, from cradle to grave. *RNA,* 24 (4): 437-460. doi:10.1261/rna.065136.117.

Dil Bahadur, R., & Akhter, A. 2018. Impact of climate-change risk-coping strategies on livestock productivity and household welfare: empirical evidence from Pakistan. *Haliyon,* e00797. doi:10.1016/j.heliyon.2018.e00797.

Dirami, T., Rode, B., Jollivet, M., *et al*. 2013. Missense mutations in *SLC26A8,* encoding a sperm-specific activator of *CFTR,* are associated with human asthenozoospermia. *Am J Hum Genet,* 92 (5): 760-6. doi:10.1016/j.ajhg.2013.03.016.

Dorman, K., Shen, Z., Yang, C., *et al*. 2012. *CTBP1* interacts with ikaros and modulates pituitary tumour cell survival and response to hypoxia. *Mol endocrinol,* 26 (3): 447-457.

Dunn W., & Milne, B. 2014. Implication of climatic heterogeneity for conservation of the Lesser Praire-Chichen (*Tympanuchs pallidinctus*). *Ecosphere,* 5 (5): 1-17.

Durst, R., Sauls, K., Peal, D., *et al*. 2015. Mutation in *DCHS1* causes mitral valve prolapse. *Nature,* 525 (7567): 109-13. doi:10.1038/nature14670.

E

Eberth, J.E., Graves, K.T., MacLeod, J.N., *et al*. 2018. Multiple alleles of *ACAN* associated with chondrodysplastic dwarfism in Miniature horses. *Anim Genet,* doi:10.1111/age.12682.

Elbeltagy, A.R., Bertolini, F., Fleming, D.S., *et al*. 2019. Natural selection footprints among African chicken breeds and village ecotypes. *Front Genet*, 10: 376. doi:10.3389/fgene.2019.00376.

Ehsaninia, J., Faye, B., & Hossein-Zadeh, G. 2020. Phenotypic diversity of camel ecotypes (*Camelus dromedarius*) in the south region of Kerman province of Iran. *Iran J Appl Anim Sci*, 10 (4): 735-745.

Elferink, MG., van As, P., Veenendaal, T., *et al.* 2010. Regional differences in recombination hotspots between two chicken populations. *BMC Genet*, 11.

Elith, J., Graham, C.K., Anderson, R.P., *et al*. 2006. Novel methods improve prediction of species distributions from occurrence data. *Ecography,* 29: 129-251.

Elith, J., Philips, S.J., Hastie, T., *et al*. 2011. A statistical explanation of MaxEnt for ecologists. *Divers Distrib,* 17: 43-57.

Eriksson J, Larson G, Gunnarsson U, *et al*. 2008. Identification of the yellow skin gene reveals a hybrid origin of domestic chicken. *PLoS Genet,* 4: e1000010. doi:10.1371/journal.pgen.1000010.

Espinoza-Derout, J., Wagner, M., Shahmiri, K., *et al*. 2007. Pivotal role of cardiac lineage protein-1 (*CLP-1*) in transcriptional elongation factor P-TEFb complex formation in cardiac hypertrophy. *Cardiovasc Res,* 75: 129-138.

Eusebi, P.G., Martinez, A., & Cortes, O. 2020. Genomic tools for effective conservation of Livestock breed diversity. *Diversity,* 12 (1): 8. doi:doi.org/10.3390/d12010008.

Evanno, G., Regnaut, R., & Goudet, J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol,* 14 (8): 2611-2626.

Eynard, S.E. 2018. Thesis: Using genomic information to conserve genetic diversity in livestock. Wageningen University.

F

Fabbi, M., Carbotti, G., & Ferrini, S. 2015. Context-dependent role of IL-18 in cancer biology and counter-regulation by IL-18BP. *J Leukoc Biol,* 97 (4): 666-675. doi:10.1189/jlb.5RU0714-360RR.

Fabregat, A., Jupe, S., Matthews, L., *et al*. 2018. The reactome pathway knowledgebase. *Nucleic Acids Res,* 46 (D1): D649-D655.

FAO. 2019. Poultry sector Ethiopia. FAO Animal Production and Health Livestock Country Reviews. No. 11. Rome.

FAO. 2020a. The State of Food Security and Nutrition in the World 2020. Rome: FAO.

FAO. 2020b. http://www.fao.org/emergencies/countries/detail/en/c/151593.

Fazzini, M., Bisci, C. & Billi, P. 2015. The climate of Ethiopia. *In Landscapes and Landforms of Ethiopia*, by P. Billi. Springer Science+Business Media Dordrecht. doi:10.1007/978-94-017-8026-1_3.

Fick, S.E., & Hijmans, R.J. 2017. Worldclim2: New 1-km spatial resolution climate surfaces for global land areas. *Int J Clim*, 37: 4302-4315.

Fisher, G., Nachtergaele, F., Prieler, S., *et al.* 2008. Global Agro-ecological Zones Assessment of Agriculture (GAEZ 2008). Luxemburg, Austria and Rome, Italy: IIASA and FAO.

Fleming, D., Koltes, J., Markey, A., *et al*. 2016. Genomic analysis of Ugandan and Rwandan chicken ecotypes using a 600k genotyping array. *BMC Genomics*, 407: 17. doi:10.1186/s12864-016-2711-5.

Fleming, D.S., Weigend, S., Simianer H., *et* al. 2017. Genomic Comparison of Indigenous African and Northern European Chickens Reveals Putative Mechanisms of Stress Tolerance Related to Environmental Selection Pressure. *G3* (*Bethesda*), 7(5):1525-1537. doi: 10.1534/g3.117.041228.

Forester, B., Jones, M., Joost, S., *et al*. 2016. Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Mol Ecol*, 25: 104-120. doi:10.1111/mec.13476.

Forester, B.R., Lasky, J.R., Wagner, H.H., *et al*. 2017. Comparing methods for detecting multilocus adaptation with multivariate genotype-environment associations. *Mol Ecol*, 27: 2215-2233. doi: 10.1111/mec.14584.

Forester, B. 2019. Detecting multilocus adaptation using Redundancy Analysis (RDA). Accessed November 24th, 2019. https://popgen.nescent.org/2018-03-27_RDA_GEA.html.

François, O., & Waits, L.P. 2016. Clustering and assignment methods in Landscape. *In Landscape Genetics*: concepts, methods, applications, by N., Cushman, S.A., Storfer, A.T., and Waits, L.P. Balkenhol, 114-128. Willey Blackwell. John Wiley & Sons Ltd.

Frichot, E., Schoville, S.D, Bouchard, G., *et al*. 2013. Testing for associations between loci and environmental gradients using latent factor mixed models. *Mol Biol Evol*, 30: 1687-1699.

Fox, RJ., Donelson, JM., Schunter, C., *et al*. 2018. Beyond buying time: the role of plasticity in phenotypic adaptation to rapid environmental change. *Philos Trans*, B 374: 20180174. doi:10.1098/rstb.2018/0174.

Fu, W., Dekkers, J., Lee, W.R., *et al*. 2015. Linkage disequilibrium in crossbred and pure line chickens. *Genet Select Evol,* 47 (11).

Fuchs, B., Wagner, T., Rossel, N., *et al.* 1997. Structure and erythroid cell-restricted expression of chicken cDNA encoding a novel zinc finger protein of the Cys + His class. *Gene,* 195 (2): 277-84. doi:10.1016/s0378-1119(97)00175-3.

Fukuda, M. 2011. TBC proteins: GAPs for mammalian small GTPase Rab? *Biosci Rep,* (31): 159-168. doi:10.1042/BSR20100112.

Furukawa, J., Inoue, K., Maeda, J., *et al*. 2015. Functional identification of *SLC43A3* as an equilibrate nucleobase transporter involved in purine salvage in mammals. *Sci Rep,* 5: 15057.

G

Gabriel, S.B., Shaffner, S.F., Nguyen, H., *et al.* 2002. The structure of haplotype blocks in the human genome. *Science,* 296: 2225-2229.

Galante, P.J., Alada, B., Muscarella, T., *et al*. 2017. The challenge of modelling niches and distribution for data-poor species: a comprehensive approach to model complexity. *Ecography,* 41: 726-736.

Galli, G., Honnens de Lichtenberg, K., Carrara, M., *et al*. 2012. Prdm5 regulates collagen gene transcription by association with RNA polymerase II in developing bone. *PLoS Genetics,* 8 (5): e1002711. doi:10.1371/journal.pgen.1002711.

Gambhir, V.A., Khatana, S., & Bedi, V.S. 2014. A brief review of high-altitude thrombosis. *Indian Journal of Vascular and Endovascular Surgery,* 1 (1): 20-23.

Gammoh, E., Arekat, M., Saldhana, L., *et al.* 2015. DENND1A gene variants in Bahraini Arab women with polycystic ovary syndrome. *Gene,* 560 (1): 30-3. doi:10.1016/j.gene.2015.01.034.

Ganley, I., Espinosa, E., & Pfeffer, S. 2008. A sintasin *10-SNARE* complex distinguishes two distinct transport routes from endosomes to the trans-Golgi in human cells. *J. Cell Biol,* 180 (1): 159-72. doi:10.1083/jcb.200707136.

Gao, Z., Lee, P., Stafford, J., *et al*. 2014. An *AUTS2* - polycomb complex activates gene expression in the CNS. *Nature.* 516 (7531): 349-54. doi:10.1038/nature13921.

García-Gómez, E., Sahana, G., Gutierrez-Gil, B., *et al*. 2012. Linkage disequilibrium and inbreeding estimation in Spanish Churra sheep. *BMC Genet,* 12: 43.

Garmendia, I., Pajares, M., Hermida-Prado, H., *et al*. 2019. YES1 drives lung cancer growth and progression and predicts sensitivity do Dasatinib. *Am J Respir Crit Care Med,* 200 (7): 888 - 899. doi:10.1164/rccm.201807-1292OC.

Gaudet, P., Livstone, M.S., Lewis, S.E., *et al*. 2011. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Brief Bioinform,* 449-462. doi:10.1093/bib/bbr042.

Gazda, M.A. 2019. Thesis: Genetics basis of simple and complex traits with relevance to avian evolution. Porto: Facultad de Ciencias da Universidade do Porto.

Getu, A., & Tadesse, A. 2014. A phenotypic and genetic characterized indigenous chicken ecotypes in Ethiopia. *Int J Genet,* 4: 4 - 10.

GFSAD30. 2017. Global Food Security Analysis-Support Data at 30 Meters Project. Accessed April 10th, 2018. https://www.usgs.gov/centers/wgsc/science/global-food-security-support-analysis-data-30-m-gfsad.

Gheyas, A., Boschiero, C., Eory, L., *et al*, D. 2015. Functional classification of 15 million SNPs detected from diverse chicken populations. *DNA Res*, 22 (3): 205-217. doi:10.1093/dnares/dsv005.

Gheyas, A., Vallejo-Trujillo, A, Kebede, A., *et al*. 2021. Integrated environmental and genomic analysis reveals the drivers of local adaptation in African indigenous chickens. *Mol Biol Evol*, preprint. doi:10.1093/molbev/msab156.

Gheye, E.F. 2000. The role of family poultry in poverty alleviation, food security and the promotion of gender equality in rural Africa. *Agriculture*, 29: 129-36.

Gilad, Y., Oschlack, A., & Rifkin, S.A. 2006. Natural selection on gene expression. *Trends Genet*, 22: 456-461.

Godber, O., & Wall, R. 2014. Livestock and food security: vulnerability to population growth and climate change. *Glob Change Biol,* 20: 3092-3102.

Goraga, Z., Wigend, S., & Brockman, G. 2011. Genetic diversity and population structure of five Ethiopian chicken ecotypes. *Anim Genet,* doi:10.1111/j/1365-2052/2011/02270.

Google. 2016. Google Earth Pro. “Ethiopia” Lat 9.017654 lon 39.988642 elev 987m. Borders and labels, places and terrain. Accessed January 9th, 2018.

GRCg6a (Genome Reference Consortium Chicken Build). 2018. Genome Reference Consortium. Washington University Genome Sequencing Center. https://www.ncbi.nlm.nih.gov/assembly/GCF_000002315.6.

Groenen, M., Wahlberg, P., Foglio, M., *et al*. 2009. A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genom Res,* 19: 510-519.

Groeneveld, L.F., Lenstra, J.A., Eding, H., *et al*. 2010. Genetic diversity in farm animals - a review. *Anim Genet,* 41 (1): 6 - 31. doi:10.1111/j.1365-2052.2010.02038.x.

Griffin, E., Hoyle, F. & Murphy, D. 2013. ‘Soil organic carbon.’ *In: Report card on sustainable natural resource use in agriculture*. Department of Agriculture and Food, Western Australia.

Gunaratne, G.S., Brailoiu, E., He, S., *et al*. 2021. Essential requirement for *JPT2* in NAADP-evoked Ca 2+ signalling. *Sci Signal,* 14 (675): eabd5605. doi:10.1126/scisignal.abd5605.

Guo, R., Chen, Y., Borgard, H., *et al*. 2020. The function and mechanism of Lipid molecules and their roles in the diagnosis and prognosis of breast cancer. *Molecules,* 25: 4864. doi:10.3390/molecules25204864.

H

Ha, M., Kim, E.D., & Chen, Z.J. 2009. Duplicate genes increase expression diversity in closely related species and allopolyploids. *PNAS*, 106: 2295-2300. doi:10.1073/pnas.0807350106.

Han, W., Zhang, T., Ni, T., *et al*. 2019. Relationship of common variants in *CHRNA5* with early-onset schizophrenia and executive function. *Schizophr Res*, 206: 407 - 412. doi:10.1016/j.schres.2018.10.011.

Haouzi, D., Boumela, I., Chebli, K., *et al*. 2018. Global, survival and apoptotic transcriptome during mouse and human early embryonic development. *Biomed Res Int,* 5895628. doi:10.1155/2018/5895628.

Hardison, R.C. 2012. Evolution of Haemoglobin and its genes. *Cold Spring Harb Perspect Med,* 2 (12): a011627. doi:10.1101/cshperspect.a011627.

Hartl, D.L. & Clark, A.G. 1997. Principles of Population Genetics. 3rd. Sunderland, MA.: Sinauer Associates, Inc.

Harvest Choice. 2015. AEZ (16-class, 2009). International Food Policy Research Institute, Washington, DC., and University of Minnesota, St. Paul, MN. Accessed March 13th, 2018. http://harvestchoice.org/data/AEZ16_CLAS.

Harvey, C.A., Chacón, M., Donatti, C.I., *et al*. 2014. Climate-Smart Landscapes: Opportunities and Challenges for Integrating Adaptation and Mitigation in Tropical Agriculture. *Conserv Let,* 7 (2): 77–90.

Hauptmann, J., Kater, L, Loffler, P., *et al*. 2014. Generation of catalytic human *AGO4* identifies structural elements important for RNA cleavage. *RNA,* 20 (10): 1532-8. doi:10.1261/rna.045203.114.

He, Q., Wang, S.Z., Leng, L., *et al*. 2018. Differentially expressed genes in the liver of lean and fat chickens. *Genet Mol Res,* 13 (4): 10823-10828.

Heller, D., & Vingron, M. 2019. SVIM: structural variant identification using mapped long reads. *Bioinformatics,* 35 (17): 2907-2915. doi:10.1093/bioinformatics/btz041.

Hengl, T., De-Jesus, J.M., MacMillan, R.A., *et al*. 2014. SoilGrids1km - Global Soil Information based on automated mapping. *PLoS ONE,* 9 (8): e105992.

Hijmans, R.J., Cameron, S.E., Parra, J.L., *et al*. 2005. Very high-resolution interpolated climate surfaces for global land areas. *Int J Clim,* 25: 1965-1978.

Hill, W.G. 1974. Estimation of Linkage Disequilibrium in randomly mating populations. *Heredity,* 33: 229-239.

Hirawake, H., Taniwaku, M., Tamura, H., *et al*. 1999. Characterization of the human SDHD gene encoding the small subunit of cytochrome b (cybS) in mitochondrial succinate - ubiquinone oxidoreductase. *Biochem. Biophys,* Acta 1412 (3): 295-300. doi:10.1016/s0005-2728(99)00071-7.

Hong, E.P. & Park, J.W. 2012. Sample size and statistical power calculation in genetic association studies. *Genomics Inf,* 2: 117-122. doi:10.5808/GI.2012.10.2.117.

Hou, Y., Qi, F., Bai, X., *et al.* 2020. Genome-wide analysis reveals molecular convergence underlying domestication in 7 birds and mammals. *BMC Genomics,* 21. doi:10.1186.s12864-020-6613-1.

Hu, Z.L., Park, C.A., Wu, X.L., *et al.* 2014. Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic Acids Res,* (41). doi:10.1093/nar/gks1150.

Huaiyu, M., Anushya, M., Dustin, E., *et al*. 2019. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res,* doi:10.1093/nar/gky1038.

Huang, Y.C., Lin, J-M, Lin, H.J., *et al*. 2011. Genome-wide association study of diabetic retinopathy in a Taiwanese population. *Ophthalmology,* 118 (4): 642-648. doi:10.1016/j.ophtha.2010.07.020.

Hwang, C., Chaurasia, S., Jackson, C., *et al*. 2013. Circadian rhythm of contrast sensitivity is regulated by a dopamine-neuronal *PAS*-domain protein 2-adenylyl cyclase 1 signalling pathway in Retinal Ganglion cells. *The Journal of Neuroscience,* 33 (38): 14989-14997. doi:10.1523/JNEUROSCI.2039-13.2013.

Huang, XQ., Zhang, XF., Xia, JH., *et al*. 2017. Tripartite motif-containing 3 (*TRIM3*) inhibits tumour growth and metastasis of liver cancer. *Chin J Cancer,* 36: 77. doi:10.1186/s40880-017-0240-5.

Huanxian, C., Zhao, C., Liu, R., *et al*. 2012. FSH stimulates lipid biosynthesis in chicken adipose tissue by upregulating the expression of its receptor *FSHR. J Lipid Res,* 53: 909-917.

Hui, T.Y., & Burt, A. 2020. Estimating linkage disequilibrium from genotypes under Hardy-Weinberg equilibrium. *BMC Genetics,* 21. doi:10.1186/s12863-020-0818-9.

Hurni, H. 1998. Agroecological belts of Ethiopia. Explanatory notes on three maps at a scale of 1:1,000,000. Adis Ababa - Ethiopia: Soil Conservation Research Programme.

Hwang, C., Chaurasia, S., Jackson, C., *et al*. 2013. Circadian rhythm of contrast sensitivity is regulated by a dopamine-neuronal *PAS*-domain protein 2-adenylyl cyclase 1 signalling pathway in Retinal Ganglion cells. *J Neurosci,* 33 (38): 14989-14997. doi:10.1523/JNEUROSCI.2039-13.2013.

I

Ip, P., Goh, W., Chan, K.W., *et al*. 2002. A novel *FUCA1* mutation causing fucosidosis in a Chinese boy. *J. Inherit Metab Dis,* 25 (5): 415-6. doi:10.1023/a:1020116220624.

J

Jackson, L.S., Finney, D.L., Kendon, E.J., *et al.* 2020. The effect of explicit convection on coupling between rainfall, humidity, and ascent over Africa under Climate change. *J Clim*, 8315-8337. doi:10.1175/JCLI-D-19-0322.1.

Jastrebski, S., Lamont, S., & Schmidt, C. 2017. Chicken hepatic response to chronic heat stress using integrated transcriptome and metabolome analysis. *PLoS One,* doi:10.1371/journal.pone.0181900.

Javanrough-Aliabad, A., Vaez-Towshizi, R., Masoudi, A.A., *et al.* 2018. Identification of candidate genes for blood metabolites in Iranian chickens using a genome-wide association study. *Brit Poultr Sci,* 1466-1799. doi:10.1080/00071668. 2018.1472743.

Jensen, K.P., Stein, M.B., Kranzler, B.Z., *et al.* 2014. The alpha-endomannosidase gene (*MANEA*) is associated with panic disorder and social anxiety disorder. *Transl Psychiatry,* 4: e353.

Jerng, H.H., Lauver, A.D., & Pfaffinger, P.J. 2013. DPP10 splice variants are localized in distinct neuronal populations and act to differentially regulate the inactivation properties of Kv4-based ion channels. *Mol Cell Neurosci,* 35 (4): 604-624. doi:10.1016/j.mcn.2007.03.008.

Jeschke, J., & Strayer, D. 2008. Usefulness of bioclimatic models for studying climate change and invasive species. *Ann N. Y. Acad Sci,* 1134: 1-24.

Jiménez-Valverde, A., Peterson, A.T., Soberon, J., *et al.* 2011. Use of niche models in invasive species risk assessments. *Biol Invasions,* 13: 2785-2797.

Jiménez-Valverde, A. 2012. Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Glob Ecol Biogeogr,* 21: 498-507.

Johnson, M., Gering, E., Willis, P., *et al.* 2015. Feralisation targets different genomic loci to domestication in the chicken. *Nat Comm,* doi:10.1038/ NCOMMS12950.

Joost, S., Bonin, A., Bruford, M.W., *et al.* 2007. A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Mol Ecol,* 16: 3955-3969.

Judith, D., Jefferies, H., Moeing, S., *et al.* 2019. ATG9A shapes the forming autophagosome through Arfaptin 2 and phosphatidylinositol 4-kinase IIB. *J Cell Biol,* 218 (5): 1634-1652. doi:10.1083/jcb.201901115.

Jueterbock, A., Smolina, I., Coyer, J.A., *et al.* 2016. The fate of the Arctic seaweed *Fucus distichus* under climate change: an ecological niche modelling approach. *Ecol Evol,* (6) 6: 1712-1724.

Julian, C.G., & Moore, L.G. 2019. Human Genetic adaptation to high altitude: evidence from the Andes. *Genes,* 10 (105). doi:10.3390/genes10020150.

Jury, M.R., & Funk, C.R. 2013. Climatic trends over Ethiopia: regional signals and drivers. *Int J Clim*, 33: 1924-1935.

K

Kaiser, P, Howell, J, Fife, M, *et al*. 2008. Integrated immunogenics in the chicken: deciphering the immune response to identify disease resistance genes. *In Animal Genomics for Animal Health*, by C Gay, P-P Pastoret & B Dodet (eds) M-H Pinard, 57-66. doi:10.1159/000317144.

Karimi, K., Esmarilizadeh-Koshkoiyeh A., & Gondro, C. 2014. Comparison of linkage disequilibrium levels in Iranian indigenous cattle using whole-genome SNPs data. *J Anim Sci.*, 57: 47-53.

Karimi, V., Karimi E., & Kshavarz, M. 2018. Vulnerability and adaptation of livestock producers to climate variability and change. *Rangel Ecol Manag*, 71 (2): 175-184. doi:10.1016/j.rama.2017.09.006.

Karlsson A., Fallahsharoudi, A., Johnsen, H., *et al*. 2016. A domestication related mutation in the thyroid stimulating hormone receptor gene (*TSHR*) modulates photoperiodic response and reproduction in chickens. *Gen Comp Endocrinol*, 228: 69-78.

Karpova, N., Bovinnec, Y., Fouix, S., *et al*. 2006. Jupiter, a new Drosophila protein associated with microtubules. *Cell Motil Cytoskeleton*, 65 (5): 301-12. doi:301-12. doi: 10.1002/cm.20124.

Kasambra, A. 2017. Package' ggpubr.' http://www.sthda.com/english/rpks/ggpubr.

Kassahun, B. & Poulton, C. 2014. The political economy of Agricultural extension policy in Ethiopia: Economic growth and political control. *Dev Policy Rev* 32: s197-s213.

Keambou, T.C., Hako, B.A., Ommeh, S., *et al*. 2014. Genetic diversity of the Cameron Indigenous chicken ecotypes. *Int J Poultr Sci*, 13 (5): 279-291.

Kebede, A. 2018. Thesis: Whole genome-based characterization of Indigenous chicken population in Ethiopia. Addis Ababa: Addis Ababa University.

Kenehisa, M., & Goto, S. 2000. KEEG - Kyoto Encyclopedia of genes and genomes.

Khalifa, H.H. 2003. Bioclimatology and adaptation of farm animals in a changing climate. *In Interactions between climate and animal production*, by N., Bernabucci, U., Khalifa, HH., Ronchi, B., and Nordone, A. Lacetera. The Netherlands: Wageningen Academic Publishers.

Kim, E.S., Cole, J.B., Hudson, H., *et al*. 2013. Effect of artificial selection on runs of homozygosity in U.S. Holstein cattle. *PLoS One*, 8: 380813.

Kim, H-K., Lee, H., Kuon, J-T., *et al*. 2015. A polymorphism in AGT and *AGTR1* gene is associated with lead-related high blood pressure. *J Renin Angiotensin Aldosterone Syst,* 16 (4): 712-9. doi:10.1177/1470320313516174.

Kimura, K., Cuvier, O., & Hirano, T. 2001. Chromosome condensation by a human condensin complex in Xenopus egg extracts. *J Biol Chem,* 276 (8): 5417-20. doi:10.1074/jbc.C000873200.

Kong, B.W., Hudson, N., Seo, D., *et al*. 2017. RNA sequencing for global gene expression associated with muscle growth in a single male modern broiler line compared to a foundational Barred Plymouth Rock chicken line. *BMC Genomics,* (18). doi:10.1186/s12864-016-3471-y.

Köhler-Rollefson, I. 2001. Community-based management of animal genetic resources - with special reference to pastoralists. Community-based management of animal genetic resources. Mbbane, Swaziland. 07-11.

Kranis, A., Gheyas, A., Boschiero, C., *et al*. 2013. Development of a high density 600K SNP genotyping array for chicken. *BMC Genomics*, 14, 59. doi:10.1186/1471-2164-14-59

Kupcinskas, J., Valantiene, I., Varkalaite, G., *et al*. 2017. *PNPLA3* and *RNF7* gene variants are associated with the risk of developing liver fibrosis and cirrhosis in an Eastern European population. *J Gastrointestin Liver Dis*, 26(1): 37-43. doi:10.15403/jgld.2014.1121.261.pnp.

L

LaFleche, L.J., & Waterman, J.M. 2020. Not playing by the rules: mixed support of ecogeographic rules in an arid-adapted African ground squirrel. *J Zool,* 312: 94-101.

Lagerström M.C., Hellström A.R., Gloriam D.E., *et al*. 2006. The G Protein-Coupled Receptor Subset of the Chicken Genome. *PLoS Comput Biol,* 2 (6): e54. doi:10.1371/journal.pcbi.0020054.

Lai, M.C., Chang, C.M., & Sun, H.S. 2016. Hypoxia induces autophagy through translational up-regulation of Lysosomal proteins in human colon cancer cells. *PLoS One,* 11.

Lamb, T.D., Patel, H.T., Chuah, A., *et al.* 2018. Evolution of the shut-off steps of vertebrate phototransduction. *Open Biol,* 8. doi:10/1098/rsob.170232.

Lawal, R.A., Al-Atiyat, R.M., Aljumaah, R.S., *et al*. 2018. Whole-Genome resequencing of Red Junglefowl and Indigenous Village Chicken Reveal New Insights on the Genome Dynamics of the Species. *Front Genet,* 1-17. doi:10.3389/fgene.2018.00264.

Lawler, A. 2016. Why did the chicken cross the world? The epic saga of the bird that powers civilization. NY: Atria Books.

Lewontin, R.C. 1964. The interaction of Selection and Linkage. *Genetics*, 49: 49-67.

Lee, J., Eschen-Lippold, L., Lassowskat, I., *et al*. 2015. Cellular reprogramming through mitogen-activated protein kinases. *Front Plan Sci,* doi:10.3389/fpls.2015. 00940.

Lee, W.H., Tsai, M.J., Chang, W.A., *et al*. 2018. Deduction of novel genes potentially involved in hypoxic AC16 human cardiomyocytes using next-generation sequencing in bioinformatics approaches. *Int J Mol Med,* 42: 2489-2502.

Lemonnier, T., Dupre, A., & Jessus, C. 2020. The G2-to-M transition from a phosphatase perspective: a new vision of the meiotic division. *Cell Div,* 15 (9). doi:10.1186/s13008-020-00065-2.

Li, S., Watanabe, T., Nakamura, S., *et al*. 1996. cDNA cloning and chromosomal mapping of a novel human *GAP (GAP1M),* a GTPase-activating protein or Ras. *Genomics,* 35: 625-627.

Li, H., & Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics,* 25: 1754-1760.

Li, Y., Sidore, C., Kang, H.M., *et al.* 2011a. Low-coverage sequencing: implications for design of complex trait association studies. *Genome Res.,* 21(6): 940-951. doi: 10.1101/gr.117259.110.

Li, C., Wang, X., Wang, G., *et al*. 2011b. Genome-wide expression analysis of roxarsone-stimulated growth of broiler chickens (*Gallus gallus*). *Comp Biochem Physiol,* 264-270. doi:10.1016/j.cbd.2011.05.003.

Li, H., Wang, T., Xu, C., *et al*. 2015. Transcriptome profile of liver at different physiological stages reveals potential model for lipid metabolism in laying hens. *BMC Genom,* 16 (763). doi:10.1186/s12864-015-1943-0.

Li, Z., Zheng, M., Abdalla, B. *et al*. 2016. Genome-wide association study of aggressive behaviour in chicken. *Sci Rep,* (6): 30981. doi:10.1038/srep30981.

Li, P., Fan, W., Everaert, N., *et al*. 2018. Messenger RNA Sequencing and Pathway Analysis Provide Novel Insights into the Susceptibility to *Salmonella enteriditis* infection in chickens. *Front Genet,* 9 (256). doi:10.3389/fgene.2018.00256.

Li, F.G. & Li, H. 2019. A time-dependent genome-wide SNP=SNP interaction analysis of chicken body weight. *BMC Genom,* 20. doi:10.1186/s12864-019-6132-0.

Li, X., Nie, C., Liu, Y., *et al*. 2020. The genetic architecture of early body temperature and its correlation with *Salmonella pullorum* resistance in three chicken breeds. *Front Genet,* doi:10.3389/fgene.2019.012870.

Lin, P.T., Gleeson, J.G., Corbo, J.C., *et al*. 2000. DCAMKL1 encodes a protein kinase with homology to doublecortin that regulates microtubule polymerization. *J Neurosci,* 20 (24): 9152-61. doi:10.1523/JNEUROSCI.20-24-09152.2000.

Liu W, Li D, Liu J, *et al.* 2011. A genome-wide SNP scan reveals novel loci for egg production and quality traits in white leghorn and brown-egg dwarf layers. *PLoS One,* 6 (12): 328600.

Liu, W., Liang, B., Liu, H., *et al*. 2017. Overexpression of non-SMC condensin I complex subunit G serves as a promising prognostic marker and therapeutic target for hepatocellular carcinoma. *Int J Mol Med,* 40 (3): 731-738. doi:10.3892/ijmm.2017. 3079.

Liu, J., Zhang, C., Wang, X., *et al*. 2020. Tumor suppressor p53 cross-talks with TRIM family proteins. *Gen Diseases*. Doi: 20.1016/j.gendis.2020.07.003.

Lobo, J.M., Jiménez-Valverde, A., & Real, R. 2008. AUC: a misleading measure of the performance of predictive distribution models. *Glob Ecol Biogeogr,* (2): 145-151.

Lotvedt, P., Fallahshahroudi, A., Bektic, L., *et al*. 2017. Chicken domestication changes expression of stress-related genes in brain, pituitary and adrenals. *Neurobiol Stress,* 7: 113-121. doi:10.1016/j.ynstr.2017.08.002.

Lozano-Jaramillo, M., Bastiaansen, J.W.M., Dessie, T., *et al*. 2018. Use of geographic information system tools to predict animal breed suitability for different agro-ecological zones. *Animal* 13: 1-8.

Lowry, D. 2012. Ecotype and the controversy over stages in the formation of new species. *Biological J,* 106 (2): 241-257.

Lu, Y., Vitart, V., Burdon, K., *et al*. 2013. Genome-wide association analyses identify multiple loci associated with central corneal thickness and keratoconus. *Nat Genet,* 45 (2): 155-163. doi:10.1038/ng.2506.

Lu, B., Sun, X., Chen, Y., *et al*. 2015. Novel function of *PITH* domain - containing 1 as an activator of internal ribosomal entry site to enhance *RUNX1* expression and promote megakaryocyte differentiation. *Cell Mol Life Sci,* 72 (4): 821-32. doi:10.1007/s00018-014-1704-2.

Lüling, R., John, H., Gudermann, T., *et al*. 2018. Transient receptor potential channel A1 (*TRPA1*) regulates sulphur mustard-induced expression of heat shock 70 kDA protein 6 (*HSPA6*) in vitro. *Cells,* 7 (9): 126. doi:10.3390/cells7090126.

Luo, W., Luo, C., Wang, M., *et al*. 2020a. Genome diversity of Chinese indigenous chicken and the selective signatures in Chinese gamecock chicken. *Sci Rep,* 10: 14532. doi:10.1038/s41598-020-71421-z.

Luo, L., Tang, Z-Z., Schoville, S.D., *et al*. 2020b. A comprehensive analysis comparing linear and generalized linear models in detecting adaptive SNPs. *Mol Ecol Res,* 21 (3): 733-744.

Lv, F.H., Agha, S., Kantanen, J., *et al*. 2014. Adaptations to climate-mediates selective pressures in sheep. *Mol. Biol. Evol,* 31: 3324-3343.

Lyimo, C.M., Weigend, A., JanBen-Tapken, U., *et al*. 2013. Assessing the genetic diversity of five Tanzanian chicken ecotypes using molecular tools. *S Afr J Anim Sci,* 43: 400-510.

M

Ma, Y., Ding, X., Qanbari, S., *et al*. 2015. Properties of different selection signature statistics and a new strategy for combining them. *Heredity,* 115 (5): 426-36. doi:10.1038/hdy.2015.42.

Marchesi, J.A., Ibelli, A.M., Peixoto, J.O., *et al*. 2019. Whole transcriptome analysis of the pectoralis major muscle reveals molecular mechanisms involved with white striping in broiler chickens. *Poultry Science,* 98: 590-601.

Maclean, C.A., Hong, N.P.C., & Prendergast, J.G.D. 2015. Hapbin: An efficient program for performing haplotype-based scans for positive selection in large genomic datasets. *Mol Biol Evol* 32: 3027-3029.

Magnuson, N., Wang, Z., Ding, G., *et al*. 2010. Why target PIN1 for cancer diagnosis and treatment? *Future Oncol,* 6 (9): 1461-1478. doi:10.2217/fon.10.106.

Mahendrasingam, S., Fettiplace, R., Alagramam, K., *et al*. 2017. Spatiotemporal changes in the distribution of *LHFPL5* in mice cochlear hair bundles during development and in the absence of *PCDH15*. *PLOS One,* 12 (10): e0185285. doi:10.1371/journal.pone.0185285.

Makuvaru, V., Walker, S, Maswre, T., *et al*. 2018. Smallholder farmer perceived effects of climate change on agricultural productivity and adaptation strategies. *Journal of Arid Environments,* 75-82. doi:10.1016/j.jaridenv.2018.01.016.

Manel, S., Schwartz, M.K., Luikart, G., *et al*. 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol Evol,* 18 (4): 189-197. doi:10.1016/S0169-5347(03)00008-9.

Manel S., & Holderegger, R. 2013. Ten years of landscape genetics. *Trends Ecol Evol,* 28: 614-621.

Mangistu, A. 2006. Ethiopia. Rome: Food and Agriculture Organization of the United Nations (FAO).

Manyelo, T.G, Salaledi, L., Hassan, Z.M., *et al.* 2020. Local chicken breed of Africa: Their description, uses and conservation methods. *Animals*, 10 (12): 2257. doi:10.3390/ani10122257.

Mao, H., Wang, X., Cheng, D., *et al*. 2019. Whole-genome SNP data unravel population structure and signatures of selection for black plumage of indigenous chicken breeds from Hiangxi province, China. *Anim Genet,* 50 (5): 475-483. doi:10.1111/age.12827.

McDermott, J.J., Staal, S.J., Freeman, H.A., *et al*. 2010. Sustaining intensification of smallholder livestock systems in the tropics. *Livestock Sci,* 130: 95-109.

McLaren, W., Gil, L., Hunt, S.E., *et al*. 2016. The Ensembl Variant Effect Predictor. *Genom Biol,* 17 (1): 122.

McKenna, A., Hanna, M., Banks, E., *et al*. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genom Res*, 20 (9): 1297-1303.

Meng, L., Wenteng, S., Zhu, Y., *et al.* 2018. Molecular characterization and expression analysis of strbp in Chinese tongue sole (*Cynoglossus semilaevis*). *Theriogenology*, 118: 225-232. doi:10.1016/j.theriogenology.2018.06.001.

Mengistu, A. 2006. ETHIOPIA: Country Pasture/Forage Resource Profiles. Rome: Food and Agriculture Organization of the United Nations (FAO).

Mi, W., Zhang, Y., Lyu, J., *et al*. 2018. The ZZ-type zinc of *ZZZE* modulates the ATAC complex-mediated histone acetylation and gene activation. *Nat Commun*, 9, 3759. doi: 10.1038/s42467-018-06247-5.

Mignon-Grasteau, S., Boissy, A., Bouix, J., *et al*. 2005. Genetics of adaptation and domestication in livestock. *Livestock Prod Sci*, 93: 3-14.

Miotto, B., Chibi, M., Xie, P., *et al*. 2014. The RBBP6/ZBTB38/MCM10 axis regulates DNA replication and common fragile site stability. *Cell Rep*, 7 (2): 575-587. doi:10.1016/j.celrep.2014.03.030.

MoA – Ministry of Agriculture. 2000. Agroecological zonation of Ethiopia. Addis Ababa, Ethiopia.

Monson, M., Van Goor, A., Aschwell, C., *et al*. 2018. Immunomodulatory effects of heat stress and lipopolysaccharide on the bursal transcriptome in two distinct chicken lines. *BMC Genom*, 19. doi:10.1186/s12864-018-5033-y.

Morales, N.S., Fernández, I.C. & Baca-Gonzalez, V. 2017. MaxEnt's parameter configuration and small samples: are we paying attention to recommendation? A systematic review. *PeerJ* 5: e3093. doi:10.7717/peerj.3093.

Morais, P., Adachi, H., & Yu, Y-T. 2021. Spliceosomal snRNA Epitranscriptomics. *Front Genet*, 12: 652129. doi:10.3389/fgene/2021.652129.

Moroudi, R.S., Mahyari, S.A., Torshizi, R.V., *et al*. 2021. Identification of new genes and quantitative trait loci associated with growth curve parameters in F2 chicken population using genome-wide association study. *Anim Genet*, 52: 171-184.

Msaki, A., Sanchez, A., Koh, L, *et al*. 2011. The role of RelA (p65) threonine 505 phosphorylation in the regulation of cell growth, survival, and migration. *Mol. Biol. Cell*, 22 (17): 3032-40. doi:10.1091/mbc.E11-04-0280.

Muchadeyi, F.C. 2007. Assessment of genetic diversity of Zimbabwe village chicken eco-types. Gottingen, Germany: Ph.D. Thesis.

Muchadeyi, F., Eding, H., Simianer, H., *et al*. 2008. Mitochondrial DNA D-loop sequences suggest a Southeast Asian and Indian origin of Zimbabwean village chickens. *Anim Genet*, 39: 615-622.

Muchadeyi, F., & Dzomba, E. 2017. Genomics tools for the characterization of genetic adaptation of low input extensively raised chickens. *Poultr Sci,* IntechOpen. doi:10.5772/65679.

Muhia, M., Thies, E., Labonte, D., *et al*. 2016. The kinesin *KIF21B* regulates microtubule dynamics and Is essential for neuronal morphology, synapse function, and learning and memory. *Cell Reports,* 5 (3): 968-977. doi:10.1016/j.celrep.2016.03.086.

Muscarella, R., Galante, P.J., Soley-Guardia, M., *et al.* 2014. ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for MaxEnt ecological niche models. *Methods Ecol Evol,* 5 (11): 1198-1205.

Mwacharo, J.M., Bjornstad, G., Mobegi, V., *et al*. 2011. Mitochondrial DNA reveals multiple introductions of domestic chicken in East Africa. *Mol Phyl Evol,* 58: 374-82.

Mwacharo, J.M., Nomura, K., Hanada, H., *et al*. 2013a. Reconstructing the origin and dispersal patterns of village chickens across East Africa: insights from autosomal markers. *Mol Ecol,* 22: 2683-2697. doi:10.1111/mec.12294.

Mwacharo, J.M., Bjornstad, G., Han, J.L., *et al*. 2013b. The history of African Village Chickens: an archaeological and molecular perspective. *Afr Archaeol Rev,* 30: 97-114. doi:10.1007/s10437-013-9128-1.

N

Nachury, M.V., & Mick, D.U. 2019. Establishing and regulating the composition of cilia for signal transduction. *Nat Rev Mol Cell Biol,* 20: 389-405.

NCBI. 2016. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res,* 44: D7-D19. doi:10.1093/nar/gkv1290.

Niyas, P.A., Chadainya, K., Shaji, S., *et al*. 2015. Adaptation of livestock to environmental challenges. *J Vet Sci Med Diag,* 4 (3). doi:10.4172/2325-9590.1000162.

North, M.A., Naggert, J.K., Yan, Y., *et al*. 1997. Molecular characterization of *TUB, TULP1,* and *TULP2,* members of the novel tubby gene family and their possible relation to ocular diseases. *PNAS.,* USA 94: 3128-33. doi:10.1073/pnas.94.7.3128.

Nosrati, M. 2017. The patterns of Linkage Disequilibrium in Livestock Genome. *Iran J Appl Anim Sci,* 7 (1): 1-7.

Nyoni, N.M.B., Grab, S. & Archer, E.R. 2018. Heat stress and chickens: climate risk effects on rural poultry farming in low-income countries. *Clim Dev,* doi:10.1080/17565529.2018.1442792.

O

O’Donnell, L., Panier, S, Wildenhain, J., *et al*. 2010. The *MMS22L-TONSL* complex mediates recovery from replication stress and homologous recombination. *Mol Cell,* 40: 619-631. doi:10.1015/j.molcel.2010.10.024.

Ogura, K., Somwar, R., Hmeijak, J., *et al*. 2021. Therapeutic potential of *NTRK3* inhibition in Desmoplastic small round cell tumour. *Clinic Cancer Res*, 27 (4). doi:10.1158/1078-0432.

Oksanen, J. 2015. Multivariate analysis of ecological communities in R: vegan tutorial. Accessed November 24th, 2019. https://cran.r-project.org/web/packages/vegan/vegan.pdf.

Olson, D.M., & Dinerstein, E. 2002. The global 200: priority ecoregions for global conservation. *Ann Mol Bot Gard*, 89: 100. doi:10.2307/3298564.

Osabe, D., Tanahashi, T., Nomura, K., *et al*. 2007. Evaluation of sample size effect on the identification of haplotype blocks. *BMC Bioinformatics*, 8: 200. doi:10.1186/1471-2105-8-200.

Ostman, B., Hintze, A., and Adami, C. 2012. Impact of epistasis and pleiotropy on evolutionary adaptation. *Proc Biol Sci*, 279:247-256.

P

Pagani, P., & Abebe, W. 2008. Review of the new features of the Ethiopian poultry sector: Biosecurity implications. By FAO.

Pampouille, E., Hennequet-Antier, C., Praud, C., *et al*. 2019. Differential expression and co-expression gene network analyses reveal molecular mechanisms and candidate biomarkers involved in breast muscle myopathies in chicken. *Sci Rep*, 9.

Parra, M., Liu, X., Derrick, S., *et al*. 2013. Molecular analysis on non-specific protection against murine malaria induces by BCG vaccination. *PLoS ONE*, 8 (7): e66115. doi:10.1371/journal.pone.0066115.

Patterson, N., Price, A.L., & Reich, D. 2006. Population structure and Eigen analysis. *PLOS Genet*, 2 (12): e190.

Pearson, R.G. 2010. Species' Distribution Modelling for Conservation Educators and Practitioners. *Lessons Cons*, 3: 54-89.

Pernes, G., Flynn, C., Lancaster, G., *et al*. 2019. Fat for fuel: lipid metabolism in haematopoiesis. *Clinic Trans Immun*, 8: e1098. doi:10.1002/cti2.1098.

Phillips, S.J., Anderson, R.P., & Shapire, R.E. 2006. Maximum entropy modelling of species geographic distributions. *Ecol Modell*, 190 (3/4): 231-259.

Pié, J., López-Vinas, E., Puisac, B., *et al*. 2007. Molecular genetics of HMG-CoA lyase deficiency. *Mol Genet Metabol*, 92: 198-209.

Pitt, J., Gillingham, P., Maltby, M., *et al*. 2016. New perspectives on the ecology of early domestic fowl: An interdisciplinary approach. *J Archaeol Sci* 74: 1-10.

Platt II, R.N., Vandewege, M.W., & Ray, D.A. 2018. Mammalian transposable elements and their impacts on genome evolution. *Chrom Res*, 26: 25-43. doi:10.1007/s10577-017-9570-z.

Prendergast, M.E., Buckley, X.M.M., Crowther, A.X., *et al*. 2017. Reconstructing Asian faunal introduction to eastern Africa from multi-proxy biomolecular and archaeological datasets. *PLoS One,* 12.

Price, A.L., Patterson, N.J., Plangem, R.M., *et al*. 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet,* 38: 904-909.

Promodou, N.V., Thompson, C.L., Osborn, D.P, *et al*. 2012. Heat shock induces rapid resorption of primary cilia. *J Cell Sci,* 125: 4297-4305.

Psfidi, A., Banos, G., Matika, O., *et al*. 2016. Genome-wide association studies of immune, disease and production traits in indigenous chicken ecotypes. *Genet Sel Evol*, 48 (74).

Puigserver, P., & Spiegelman, B.M. 2003. Peroxisome proliferator-activated receptor-gamma coactivator 1 alpha (*PGC-1* alpha): transcriptional coactivator and metabolic regulator. *Endocrinol Rev,* 24 (1): 78-90. doi:10.1210/er.2002-0012.

Purcell, S., Neale, B., Todd-Brown, K., *et al*. 2007. PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet,* 81.

Purswell, J., Dozier, W., Olanrewaju, H., *et* al. 2012. Effect of temperature-humidity index on live performance in Broiler chickens grown from 49 to 63 days of age. *9th International Livestock Environment Symposium.* Valencia, Spain: American Society of Agricultural and Biological Engineers (ASABE). ILES12-0265.

Q

Qanbari S, Seidel M, Strom TM, *et al*. 2015. Parallel Selection revealed by population sequencing in chicken. *Genome Biol Evol*, 7 (12): 3299-3306. doi:10.1093/GBE/EVV222.

Qanbari, S. 2020. On the extent of linkage disequilibrium in the genome of farm animals. *Front Genet*, 10:1304. doi: 10.3389/fgene.2019.01304.

Qu, L., Shen, M., Guo, J., *et al*. 2019. Identification of potential genomic regions and candidate genes for egg albumen quality by a genome-wide association study. *Arch Anim Breed,* 62 (1): 113-123. doi:10.5194/aab-62-113-2019.

Quinlan, A.R., & Kindlon, N. 2017. Bedtools. Accessed 11 24, 2019. https://bedtools.readthedocs.io/en/latest/index/html.

Quinghe Li, Wang, N., Zhou, D., *et al*. 2012. Gastrocnemius transcriptome analysis reveals domestication induced gene expression changes between wild and domestic chickens. *Genomics,* 100 (5): 314-319.

R

Rashamol, V. & Sejian, V. 2018. Climate resilient livestock production: way forward. *J Dairy Vet Sci*, 5 (5): 5556673. doi:10.19090/JDVS.2018.05.555673.

Ratan, A., Olson, T.L., Loughran, T.P., *et al*. 2015. Identification of indels in next-generation sequencing data. *BMC Bioinformatics*, 16 (42). doi:10.1186/s12859-015-0483-6.

Razani, B., Zarnegar, B., Ytterberg, A.J., *et al*. 2010. Negative feedback in noncanonical NF-kappaB signalling modulates NIK stability through IKKalpha-mediated phosphorylation. *Sci Signal*, 3 (23): ra41. doi:10.1126/scisignal.2000778.

R Core, Team. 2013. R: A language and environment for statistical computing. Prod. R Foundation for Statistical Computing. Vienna. http://www.R-project.org/.

Rees, J.S., Castellano, S., Andres, A.M. 2020. The genomics of human local adaptation. *Trends Genet* 36:415-428.

Rellstab, C., Gugerli, F., Eckert, A.J., *et al*. 2015. A practical guide to environmental association analysis in landscape genomics. *Mol Ecol*, 24: 4348-4370.

Rima, M., Daghsni, M., De Waard, S., *et al*. 2017. The B4 subunit of the voltage-gated calcium channel (*Cacnb4*) regulates the rate of cell proliferation in Chinese Hamster Ovary cells. *Int J Biochem Cell Biol*, 89: 57-70. doi:10.1016/ j.biocel.2017.05.032.

Rohde, P.D., Østergaard, S., Kristense, T.N., *et al*. 2018. Functional validation of candidate genes detected by genomic feature models. *G3*, 8 (5): 1659-1668. doi:10.1534/g3.118.2000082.

Rubin, C.J., Zody, M.C., Eriksson, J., *et al*. 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature*, 464: 587-591.

S

Sabeti, P.C., Varily, P., Fry, B., *et al*. 2007. Genome-wide detection and characterization of positive selection in human populations. *Nature*, 449: 913-918.

Sanarana, Y., Visser, C., Bosman, L., *et al*. 2016. Genetic diversity in South African Nguni cattle ecotypes based on microsatellite markers. *Trop Anim Health Prod*, 379-385.

Sanchis-Juan, A., Stephens, J., French, C.S., *et al*. 2018. Complex structural variants in Mendelian disorders: Identification and breakpoint resolution using short- and long-read genome sequencing. *Genome Medicine*, 10 (95). doi:10.1186/s13073-018-0606-6.

Sandoval-Castillo, J., Gates, K., Brauer, C.J., *et al*. 2020. Adaptation of plasticity to projected maximum temperatures and across climatically defined bioregions. *PNAS*, 117 (29): 17112-17121. doi:10.1073/pnas.1921124117.

Saravanan, K.A., Panigrahi, M., Kumar, H., *et al*. 2020. Selection signatures in livestock genome: A review of concepts, approaches and applications. *Livestock Sci*, 241(104257). doi: 10.1016/j.livsci.2020.104257.

Sarkar, S., Benerjee, P.K., & Selvamurthy, W. 2003. High altitude hypoxia: An intricate interplay of oxygen responsive macroevents and micromolecules. *Mol Cell Biochem,* 253: 287-305.

Schmid, M., Smith, J., Burt, DW., *et al*. 2015. Third report on chicken genes and chromosomes. *Cytogenet Genome Res,* 145: 78-179.

Scholtz, M.M., Furstenburg, D., Maiwashe, A., *et al.* 2010. Environmental-genotype responses in livestock to global warming: A southern African perspective. *S Afr J Anim Sci,* (5(1)): 408-413.

Scott, N., Sharpe, L., & Brown, A. 2020. The *E3* ubiquitin ligase *MARCHF6* as a metabolic integrator in cholesterol synthesis and beyond. *Mol Cell Biol Lip,* 1866 (1).

Sehgal, D., Baliyan, N., & Kaur, P. 2019. Progress towards identification and validation of candidate genes for abiotic stress tolerance in wheat. In *Genomics assisted breeding of crops for abiotic stress tolerance*. Vol. II. Sustainable Development and Biodiversity, vol21, by V., Sehgal, D., Kumar, A., Raina, S. (eds) Rajpal. Springer.

Sejian, V. 2012. Introduction. *In Environmental stress and amelioration in livestock production,* by V., Naqvi, S.M.K., Ezeji, T., Lakritz, J., and Lal, R. Sejian, 1-15. New York: Springer Verlag.

Senna-Cherchi, S., Burgess, K., Nees, S., *et al*. 2011. Exome sequencing identified MYO1E and NEIL1 as candidate genes for human autosomal recessive steroid-resistant nephrotic syndrome. *Kidney Int,* 80 (4): 389-96. doi:10.1038/ki.2011.148.

Shan-Shan D, Yan G, Dong-Li Z, *et al*. 2016. Epigenomic element analyses for promoters identify *ESRRG* as a new susceptibility gene for obesity-related traits. *Int J Obes,* 40 (7): 1170-1176. doi:10.1038/ijo.2016.4.

Shilin T, Xuming Z, Tashi P, *et al*. 2020. Genomic Analyses Reveal Genetic Adaptations to Tropical Climates in Chickens. *iScience,* 23: 101544. doi:10.1016/j.osco.2020.101644.

Shi, H., Wang, X., Lu, Z., *et al*. 2017. *YTHDF3* facilitates translation and decay of N6-methyladenosise-modified RNA. *Cell Res,* 27: 315-328. doi:10.1038/cr.2017.15.

Shin, S., Han, J.Y., & Lee, K. 2010. Cloning of avian Delta-like 1 homolog gene: The biallelic expression of Delta-like 1 homolog in avian species. *Poultr Sci,* 88 (5): 948-955. doi:10.3382/ps.2009-00572.

Singh, V., Singh, K., & Baum, K. 2018. The role of methionine sulfoxide reductases in oxidative stress tolerance and virulence of Staphylococcus aureus and other bacteria. *Antioxidants,* 7 (10): 128. doi:10.3390/antiox7100128.

Slate, J., David, P., Dodds, K.G., *et al*. 2004. Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. *Heredity,* 93: 255-265.

Smith, J.J., Hasiotis, S.T., Kraus, M.J., *et al*. 2009. Transient dwarfism of soil fauna during the Paleocene-Eocene Thermal Maximum. *PNAS*, 106: 17655-17660. doi:10.1073/pnas.0909674106.

Solinhac, R., Leroux, S., Galkina, S., *et al*. 2010. Integrative mapping analysis of chicken microchromosome 16 organization. *BMC Genom*, 11: 616.

Son, S., Park, S., Fernández-Estévez, M., *et al*. 2021. Autophagy regulation by acetylation - implications for neurodegenerative diseases. *Exp Mol Med*, 53: 30-41. doi:10.1038/s12276-021-00556-4.

Stamm, S., Gruber, S., Rabchevsky, A. *et al*. 2018. The activity of the serotonin receptor 2C is regulated by alternative splicing. *Hum Genet*, 136 (9): 1079-1091. doi:10.1007/s00439-017-1826-3.

Stavber, L., Hovnik, T., Ktonik, P., *et al*. 2020. High frequency of pathogenic ACAN variants including an intragenic deletion in selected individuals with short stature. *Eur J Endocr*, 182: 243-253. doi:10.1530/EJE-19-0771.

Steinbacher, P., Wanzenbock, J., Brandauer, M., *et al*. 2017. Thermal experience during embryogenesis contributes to the induction of dwarfism in whitefish Coregonus lavaretus. *PLoS ONE*, 12 (9): e0185384. doi:10.1371/journal.pone.0185384.

Stelzer G, Rosen R, Plaschkes I, *et al*. 2016. *GeneCards*. www.genecards.org.

Stephan, A., Shum, E., Hirsh, S., *et al*. 2009. ANO2 is the cilial calcium-activated chloride channel that may mediate olfactory amplification. *PNAS*, USA 106: 11776-81. doi:10.1073/pnas.0903304106.

Stocker, T.F., Qin, D., Plattner, G.K., *et al*. 2013. IPCC: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. 1535.

Strillacci, M.G., Vargas-Murillo, V.E., Román-Ponce, S.I., *et al*. 2018. Looking at genetic structure and selection signatures of the Mexican chicken population using single nucleotide polymorphism markers. *Poultry Science*, 97 (3): 791-802. doi:10.3382/ps/pex374.

Stroud, D.A., Maher, M.J., Lindau, C., *et al*. 2015. COA6 is a mitochondrial complex IV assembly factor critical for biogenesis of mtDNA-encoded COX2. *Hum Mol Genet*, 24 (19): 5404-15. doi:10.1093/hmg/ddv265.

Stucki, S., Orozco-terWengel, P., Forester B.R., *et al*. 2016. High performance computation of landscape genomic models including local indicators of spatial association. *Mol Ecol Res*.

Sun, L., Shi, L., Li, W., *et al*. 2009. *JFK*, a Kelch domain-containing F-box protein, links the *SCF* complex to p53 regulation. *Proc Natl Acad Sci*, USA 106 (25): 10195-200. doi:10.1073/pnas.0901864106.

Sun, C., Lu, J., Yi, G., *et al*. 2015. Promising loci and genes for yolk and ovary weight in chickens revealed by a genome-wide association study. *PLoS ONE,* 10 (9). doi:e0137145. doi:10.1371/journal.pone.0137145.

Sun, J., Chen, T., Zhu, M., *et al*. 2021. Whole-genome sequencing revealed genetic diversity and selection of Guangxi indigenous chickens. *bioRxiv*. doi:10.1101/2021.04.07.438797.

T

Tadelle D., Kijora, C., & Peter, K.J. 2003. Indigenous chicken ecotypes in Ethiopia: Growth and feed utilization potentials. *Int J Poultr Sci,* 2: 144-152.

Tang, S., Zhang, J., Mei, T.T. *et al*. 2019. Association of TM6SF2 rs 58542926 T/C gene polymorphism with hepatocellular carcinoma: a meta-analysis. *BMC Cancer,* 19: 1128. doi:10.1186/s12885-019-6173-4.

Teluguntla, P., Thenkabail, P.S., Xiong, J., *et al*. 2014. Chapter 7: Global Cropland Area Database (GCAD) derived from Remote sensing in support of Food Security in the Twenty-first Century: Current achievements and future possibilities. Vol. II, in Land Resources: Monitoring, Modelling, and Mapping Remote Sensing Handbook, by Prasad S. Thekabail edit.

Teng, Y., Guo, B., Mu, X., *et al*. 2018. KIF26B promotes cell proliferation and migration through the FG2/ERK signalling pathway in breast cancer. *Biomed. Pharmacother*, 108: 766-773. doi:10.1016/j.biopha.2018.09.036.

ter Braak, C.J.F. 1994. Canonical community ordination. Part I: basic theory and linear methods. *Ecoscience,* 1: 127-140.

The UniProt Consortium. 2020. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res*. D1.

Thorn, J.S., Nijman, V., Smith, D., *et al*. 2009. Ecological niche modelling as a technique for assessing threats and setting conservation priorities for Asian slow lorises (Primates: *Nycticebus*). *Div Dist,* 15: 289-298.

Tiffin, P. & Ross-Ibarra, J. 2014. Advances and limits of using population genetics to understand local adaptation. *PeerJ* PrePrints, 2: e488v1. doi:10.7287/peerj.preprints.488v1.

Timson, D.J. 2006. The structural and molecular biology of type III galactosemia. *IUMBM Life,* 58 (2): 83-9. doi:10.1080/15216540600644846.

Topaloglu, A., Reimann, F., Guclu, M. *et al.* 2009. TAC3 and TACR3 mutations in familial hypogonadotropic hypogonadism reveal a key role for Neurokinin B in the central control of reproduction. *Nat Genet,* (41): 354-348. doi:doi.org/10.1038/ng.306.

Truong, L., Zheng, YM., Song, TY., *et al.* 2018. Potential important roles and signalling mechanisms of YPEL4 in pulmonary diseases. *Clin Trans Med*, 7.

U

Ullrich, F., Blin, S., Lazarow, K., *et al*. 2019. Identification of TMEM206 proteins as pore of PAORAC/ASOR acid-sensitive chloride channels. *eLife,* 8: e49187. doi:10.7554/eLife.49187.

UN. 2019. https://population.un.org/wpp/.

V

Van der Auwera, G.A., Carneiro, M., Hartl, C., *et al*. 2013. From FastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr Protoc Bioinformatics,* 43: 11.10.1 - 11.10.33.

van Marle-Köster, E., & Visser, C. 2018. Genetic improvement in South African livestock: can genomics bridge the gap between the developed and developing sectors? *Front Genet*, 9:331. doi: 10.3389/fgene.2018.00331.

Velado-Alonso, E., Morales-Castilla, I. & Gómez-Sal, A. 2020. Recent land use and management changes decouple the adaptation of livestock diversity to the environment. *Sci Rep* 10: 21035. doi:10.1038/s41598-020-77878-2.

Vernikos, G.S. 2020. A review of pangenome tools and recent studies. In: The pangenome: diversity, dynamics and evolution of genomes. Tettelin, H. and Medini, D. editors. Springer.

Voight, B.F., Kudaravalli, S., Wen, X., *et al*. 2006. A map of recent positive selection in the human genome. *PLoS Biol*, 4: e72.

von Papen, M., Gambaryan, S., Schutz, C., *et al*. 2013. Determination of ATP and ADP secretion form human and mouse platelets by an HPLC assay. *Transfus Med Hemother*, 40 (2): 109-116. doi:10.1159/000350294.

W

Wagstaff, P., Kang, H.Y., Mylott, D., *et al*. 1995. Characterization of the avian GLUT1 glucose transporter: differential regulation of *GLUT1* and *GLUT3* in chicken embryo fibroblasts. *Mol Biol Cell,* 6: 1575-1589.

Walugembe, M., Bertolini, F. Dematawewa, C.M.B., *et al*. 2019. Detection of selection signatures among Brazilian, Sri Lankan, and Egyptian chicken populations under different environmental conditions. *Front Genet,* 9.

Wang, J., Liu, Q., Yuan, S., *et al*. 2017. Genetic predisposition to lung cancer: comprehensive literature integration, meta-analysis, and multiple evidence assessment of candidate-gene association studies. *Sci Rep,* 7: 8371. doi:10.1038/s41598-017-07737-0.

Wang, X., Li, Z., Guo, Y., *et al*. 2018. Identification of a novel 43 bp insertion in the heparan sulfate 6-O-sulfotransferase 3 (HS6ST3) gene and its associations with growth and carcass traits in chickens. *Anim Biotech*, doi:10.1080/10495398.2018.1479712.

Warren, D., & Seifert, S.N. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography,* 33: 607-611.

Warren, D.L., Wright, A.N., Seifert, S.N., *et al*. 2014. Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 California vertebrate species of concern. *Diversity and Distribution,* 20: 334-343.

Weigand, H., & Leese, F. 2018. Detecting signatures of positive selection in non-model species using genomic data. *Zool J Linn Soc*, 184: 528-583. doi:10.1093/ zoolinnean/zly007.

Weir, B.S., & Cockerham, C. 1984. Estimating F-statistics for the analysis of population-structure. *Evolution,* 38: 1358-1370.

Wellenreuther, M., Merot, C., Berdan, E., *et al*. 2019. Going beyond SNPs: The role of structural genomic variants in adaptive evolution and species diversification. *Mol Ecol*, 1203-1209. doi:10.1111/mec.15066.

Wepy, J.S., Galligan, J.J., Kingsley, P., *et al*. 2019. Lysophospholipases cooperate to mediate lipid homeostasis and lysophospholipid signalling. *J Lipid Res*, 60 (2): 360-374. doi:10.1194/jlr.M087890.

WFP. 2020. WFP East Africa: Update on the Desert Locust Outbreak. Ethiopia: World Food Programme.

Williams, G. 1966. Adaptation and Natural Selection. Princeton: Princeton University Press.

Williams, K., Liu, M-L., Zhu, Y., *et al*. 2005. Loss of Heparan N-Sulfotransferase in Diabetic Liver. *Diabetes* 54 (4): 1116-1122. doi:10.2337/diabetes. 54.4.1116.

Wilson, R.T. 2010. Poultry production and performance in the Federal Democratic Republic of Ethiopia. *Poultr Sci J,* 66: 441-54.

Wilson, M.A. 2018. Genetic diversity on the sex chromosomes. *Genome Biol Evol*, 10(4): 1064-1078. doi:10.1093/gbe/evy039.

Woldekiros, H., and D’andrea, A. 2016. Early evidence for domestic chickens (*Gallus gallus domesticus*) in the Horn of Africa. *Int J Osteorarch,* 27, 329-341.

Wright, S. 1992. Coefficients of inbreeding and relationship. *Am Nat,* 56: 330-339.

Wu, H., Min, J., Zeng, H., *et al*. 2008. Crystal structure of the methyltransferase domain of human TARBP1. *Proteins,* 72 (1): 19-25. doi:10.1002/prot.22065.

X

Xia, L.C., Bell, J.M., Wood-Bouwens, C., *et al*. 2017. Identification of large rearrangements in cancer genomes with barcode linked reads. *Nucleic Acids Res*, 46: e19.

Xiaotong, Z., Zhang, B., Zhang, Y., *et al*. 2020. Transcriptome analysis of feather follicles reveals candidate genes and pathways associated with pheomelanin pigmentation in chickens. *Sci Rep*.

Xu, A.G., Xu, R.M., Lu, C.Q., *et al*. 2012. Correlation of von Willebrand factor gene polymorphism and coronary heart disease. *Mol Med Rep,* 1107-1110. doi:10.3892/ mmr.2012.1037.

Xu, L., and Zhou, Q. 2020. The female-specific W chromosomes of birds have conserved gene contents but are not feminized. *Genes,* 11, 1126. doi: 10.3390/genes11101126.

Y

Yackulic, C.B., Chandler, R., Zipkin, E.F., *et al*. 2013. Presence-only modelling using MAXENT: when can we trust the inferences? *Methods Ecol Evol,* 4: 236-243. doi:10.1111/2041-210x.12004.

Yamaguchi, H. & Wang, H-G. 2004. *CHOP* is involved in endoplasmic reticulum stress-induced apoptosis by enhancing *DR5* expression in human carcinoma cells. *J. Biol. Chem,* 279 (44): 45495-502. doi:10.1074/jbc.M406933200.

Yang, Y., Zhang, G., Guo, F., *et al*. 2020. Mitochondrial *UQCC3* modulates hypoxia adaptation by orchestrating OXPHOS and glycolysis in hepatocellular carcinoma. *Cell Rep,* 5 (3): 10834. doi:10.1016/j.celrep.2020.108340.

Yao, C., Spurlock, D.M., Armentano, L.E., *et al*. 2013. Random Forest approach for identifying additive and epistatic single nucleotide polymorphism associated with residual feed intake in dairy cattle. *J Dairy Sci,* 96: 6716-6729.

Yeates, N.T.M. 2009. Foetal dwarfism in sheep - an effect of high atmospheric temperature during gestation. *J Agri Sci,* 51 (1): 84-89. doi:10.1017/S00218596000 32846.

Yen, C.F., Wang, H.S., Lee, C.L., *et al*. 2014. Roles of integrin-linked kinase in cell signalling and its perspectives as a therapeutic target. *Gynecol Minim Invasive Ther,* 3 (3): 67-72. doi:10.1016/j.gmit.2014.06.002.

Yoon, W., Hwang, S-H., Lee, S-H., *et al*. 2019. Drosophila *ADCK1* is critical for maintaining mitochondrial structures and functions in the muscle. *PLoS Genet,* 15 (5): e1008184. doi:10.1371/journal.pgen.1008184.

Yuan, J., Sun, C., Dou, T., *et al.* 2015. Identification of promising mutants associated with egg production traits revealed by genome-wide association study. *PLoS One,* 10: e0140615. doi:10.1371/journal.pone.0140615.

Yuan, X., Lan, G., Li, L., *et al*. 2020. Differential gene expression profiling of the goose pineal gland. *Brit Poultr Sci,* 61 (2): 200 -208. doi:10.1080/00071668. 2019.1698014.

Z

Zeng, Y., Low, B.W., & Yeo, D.C.J. 2016. Novel methods to select environmental variables in MaxEnt: A case study using invasive crayfish. *Ecol Modell* 341: 5-13.

Zhang, Q., Ray, K., Acland, G.M., *et al*. 2000. Molecular cloning, characterization and expression of a novel retinal clustering-like protein cDNA. *Gene,* 243 (1-2): 151-160. doi:10.1016/S0378-1119(99)00542-9.

Zhang, K., Huentelman, M.J., Rao, F., *et al*. 2014. Genetic implication of a novel thiamine transporter in human hypertension. *J Am Coll Cardiol*, 65 (12): 1542-55. doi:10.1016/j.jacc.2014.01.007.

Zhang, Q., Gou, W.Y., Wang, X.T., *et al*. 2016a. Genome resequencing identifies unique adaptations of Tibetan chickens to hypoxia and high-dose ultraviolet radiation in high-altitude environments. *Genom Biol Evol,* 8: 765-778.

Zhang, R., Lin, Y., Zhi, L., *et al*. 2016b. Expression profiles and associations of adiponectin and adiponectin receptors with intramuscular fat in Tibetan chicken. *Mol Cell Dev Biol,* 151-157. doi:10.1080/00071668.2016.1268252.

Zhang, M., Peng, W.F., Hu, X.J., *et al*. 2018. Global genomic diversity and conservation priorities for domestic animals are associated with the economies of their region of origin. *Sci Rep,* 8: 1677. doi:10.1038/s41598-018-30061-0.

Zhang, C., Dong, S.S., Xu, J.Y., *et al*. 2019. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics,* 35 (10): 1786-1788.

Zhang, H., Shen, L-Y., Xu, Z-C., *et al*. 2020. Haplotype-based genome-wide association studies for carcass and growth traits in chicken. *Poultr Sci,* 99 (4): 2349 - 2361. doi:10.1016/j.psj.2020.01.009.

Zhao, B., Ye, X., Yu, J., *et al*. 2008. *TEAD* mediates YAP-dependent gene induction and growth control. *Genes Dev,* 22 (14): 1962-71. doi:10.1101/gad.1664408.

Zhong, X.Y., Ding, J.H., Adams, J., *et al*. 2009. Regulation of SR protein phosphorylation and alternative splicing by modulating kinetic interactions of SRPK1 with molecular chaperones. *Genes Dev,* 23 (4): 482-05. doi:10.1101/gad.1752109.

Zhou, Z., Cui, N., Su, S., *et al*. 2018. The molecular basis for host responses to Marek’s disease viruses integrated with different retro-viral long terminal repeat. *Poultr Sci,* 97 (9): 3015-3022. doi:10.3382/ps/pey135.

Zhu B, Li Q, Liu R, *et al*. 2019. Genome-wide association study. *Animals,* 9 (260). doi:10.3390/ani9050260.

Zi-Xuan Z, Shuen-Ei C, Chih-Feng C, *et al*. 2019. Genome-wide association study on the body temperature changes of a broiler-type strain Taiwan country chickens under acute heat stress. *J Ther Biol,* 82: 33-42. doi:109.1016/j.jtherbio.2019.03.007.

Zizioli, D., Tiso, N., Guglielmi, A., *et al*. 2016. Knock-down of pantothenate kinase 2 severely affects the development of the nervous and vascular system in zebrafish, providing new insights into PKAN disease. *Neurobiol Disease,* 85: 35-48.

Zwane, A.A., A. Maiwashe, M.L., Makgahlela, A., *et al*. 2016. Genome-wide identification of breed-informative single-nucleotide polymorphisms in three South African indigenous cattle breeds. *S Afr J Anim Sci,* 46: 302-312.

Supplementary Tables

Chapter 2

Genetic diversity and population structure of indigenous Ethiopian village chicken

Supplementary Table S2.1. Summary results on sequencing and mapping of reads

Population	# reads[a]	MR (%)[b]	MD (X)[c]
Adane	376903935	99.38	35.89
Alfa Midir	342600741	98.02	39.39
Amesha Shinkuri	427316335	99.43	39.02
Arabo	309414324	99.54	31.30
Ashuda	436640467	99.10	39.76
Batambie	436294418	99.37	40.33
Bekele Girissa	383898518	98.78	36.38
Dikuli	418411806	99.17	36.39
Gafera	403298450	99.12	43.45
Gesses	373183155	99.62	35.66
Gijet	439628764	99.53	43.08
Hadush Adi	201579979	98.67	39.68
Hugub	360343162	99.71	35.67
Jarso	368646704	99.79	25.96
Kido	357193268	84.27	37.30
Kumato	404282348	99.75	38.13
Local Horro	387105936	99.83	28.68
Loya	461026931	99.57	42.61
Meseret	433954119	99.46	39.24
Metkilmat	339764180	99.54	33.99
Mihquan	335214296	99.73	34.37
Negasi Amba	362919193	99.02	35.52
Shubi Gemo	421492594	99.36	39.09
Surta	475115042	98.87	41.51
Tzion Teguaz	244270731	99.40	22.04

(a) Average number of reads.
(b) Total number of reads mapped to the GRCga6a reference genome.
(c) Mean sequence depth of the genome sequence coverage.

Supplementary table S2.2. LD pairwise comparison for 6 population groups defined by PCA for different distance (kb) across the chicken autosome genome. (# pw = number of pairwise comparisons. # (r^2 = 0.8-1) = number of pairwise comparison with r2 between 0.8 – 1. % = percentage of pairwise comparisons with r^2 between 0.8 – 1).

		Distance										
		< 2 kb	2 - 4 kb	4 - 6 kb	6 - 8 kb	8 - 10 kb	10 - 12 kb	12 - 14 kb	14 - 16 kb	16 - 18 kb	18 - 20 kb	20 - 25 kb
group 1	# pw	9309328	2654154	569869	160903	60359	27760	15512	9068	5945	4172	1
	# (r^2 = 0.8-1)	3032112	570588	123425	38738	15767	7669	4188	2771	1817	1238	1
	%	32.6	21.5	21.7	24.1	26.1	27.6	27.0	30.6	30.6	29.7	100.0
group 2	# pw	9293198	2058751	369269	99278	37425	17415	9487	5752	3648	2485	5
	# (r^2 = 0.8-1)	3484602	885871	222867	72506	29723	14556	8055	4891	3306	2299	2
	%	37.5	43.0	60.4	73.0	79.4	83.6	84.9	85.0	90.6	92.5	40.0
group 3	# pw	8716725	1167111	158868	39947	14984	6904	4033	2464	1564	1032	0
	# (r^2 = 0.8-1)	2571761	287543	51495	15115	6105	2885	1730	1033	698	457	0
	%	29.5	24.6	32.4	37.8	40.7	41.8	42.9	41.9	44.6	44.3	
group 4	# pw	9168242	1341246	185973	46477	17037	7722	4531	2804	1799	1159	2
	# (r^2 = 0.8-1)	2859848	348560	62934	18378	7378	3543	1993	1292	842	491	0
	%	31.2	26.0	33.8	39.5	43.3	45.9	44.0	46.1	46.8	42.4	
group 5	# pw	8634970	1148292	155732	39768	14929	6757	3758	2290	1514	978	0
	# (r^2 = 0.8-1)	2628924	293893	51965	15609	6278	2863	1609	989	618	402	0
	%	30.4	25.6	33.4	39.3	42.1	42.4	42.8	43.2	40.8	41.1	0
group 6	# pw	8841621	1211699	167163	42406	15762	15762	4052	2326	1552	1000	0
	# (r^2 = 0.8-1)	2681763	307398	56278	16694	6728	3179	1740	1083	703	436	0
	%	30.3	25.4	33.7	39.4	42.7	20.2	42.9	46.6	45.3	43.6	0

(group 1 = Jarso population; group 2 = Hugub population; group 3 = Kumato, Loya, Shubi Gemo and Bekele Girissa populations; group 4 = Alfa Midir, Negasi Amba, Adane, and Arabo populations; group 5 = Metkilimat, Hadush Adi, Mihquan, Gijet, and Meseret populations; group 6 = Gesses, Kido, Gafera, Horro, Tzion Teguaz, Dikuli, Amesha Shinkuri, Batambie, Ashuda and Surta populations).

Chapter 3

Ecological niche modelling for environmental characterisation of indigenous Ethiopian village chickens

Table S3.1 AUC values of environmental variables for individual ENM's ran under different combinations of feature classes and beta-multipliers. (**Model configuration:** number of the evaluated model. **Settings: feature** classes and regularisation multipliers (**RM**) used for the model. -H: hinge, Q: quadratic, P: product, H: hinge, C: categorical, T: threshold-. **Train AUC:** Area Under the Receiving Operator Curve value for train test. **testAUC:** AUC value for test model. **diffAUC:** the difference between the AUC value based on training localities. **OR_{MTP}:** a threshold-dependent metric that indicates the proportion of test localities with suitable values lower than that associated with the lowest-ranking training locality. **OR_{10}:** a threshold-dependent metric that displays the proportion of test localities with suitable values lower than that excluding the 10% of training localities with the lowest predicted suitability. **AICc:** Akaike Information Criterion corrected for small sample sizes. **delta.AICc:** the difference between the AICc of a given model and the AICc of the model with the lowest AICc. **w.AICc:** Akaike weight divided by the sum of the likelihood values of all models included in a run. **nparam:** number of parameters in a Maxent model).

Model configuration	settings	features	RM	train AUC	X̄ testAUC	X̄ testAUC	X̄ diffAUC	S.D. diffAUC	X̄ test. OR_{MTP}	S.D. test OR_{MTP}	X̄ test OR_{10}	S.D. test OR_{10}	AICc	delta. AICc	w.AIC	nparam
126	HQP_3.5	HQP	3.5	0.9587	0.9523	0.0004	0.0056	0.0003	0.008	0.0006	0.128	0.0031	6039.08	0.0000	0.3010	42
127	HQPT_3.5	HQPT	3.5	0.9587	0.9523	0.0004	0.0056	0.0003	0.008	0.0006	0.128	0.0031	6039.08	0.0000	0.3010	42
128	HQPTC_3.5	HQPTC	3.5	0.9587	0.9523	0.0004	0.0056	0.0003	0.008	0.0006	0.128	0.0031	6039.08	0.0000	0.3010	42
110	HQP_3	HQP	3	0.9616	0.9559	0.0003	0.0055	0.0003	0.008	0.0006	0.12	0.0043	6043.59	4.5096	0.0316	53
111	HQPT_3	HQPT	3	0.9616	0.9559	0.0003	0.0055	0.0003	0.008	0.0006	0.12	0.0043	6043.59	4.5096	0.0316	53
112	HQPTC_3	HQPTC	3	0.9616	0.9559	0.0003	0.0055	0.0003	0.008	0.0006	0.12	0.0043	6043.59	4.5096	0.0316	53
43	LQP_1	LQP	1	0.9427	0.9392	0.0014	0.0086	0.0006	0.008	0.0003	0.112	0.0024	6051.11	12.0284	0.0007	33
44	LQPT_1	LQPT	1	0.9427	0.9392	0.0014	0.0086	0.0006	0.008	0.0003	0.112	0.0024	6051.11	12.0284	0.0007	33
45	LQPTC_1	LQPTC	1	0.9427	0.9392	0.0014	0.0086	0.0006	0.008	0.0003	0.112	0.0024	6051.11	12.0284	0.0007	33
11	LQP_0.1	LQP	0.1	0.9468	0.9405	0.0022	0.0098	0.0011	0.004	0.0002	0.108	0.0022	6066.06	26.9742	0.0000	40
12	LQPT_0.1	LQPT	0.1	0.9468	0.9405	0.0022	0.0098	0.0011	0.004	0.0002	0.108	0.0022	6066.06	26.9742	0.0000	40
13	LQPTC_0.1	LQPTC	0.1	0.9468	0.9405	0.0022	0.0098	0.0011	0.004	0.0002	0.108	0.0022	6066.06	26.9742	0.0000	40
27	LQP_0.5	LQP	0.5	0.9451	0.9402	0.0017	0.0090	0.0007	0.008	0.0003	0.108	0.0029	6073.50	34.4203	0.0000	40
28	LQPT_0.5	LQPT	0.5	0.9451	0.9402	0.0017	0.0090	0.0007	0.008	0.0003	0.108	0.0029	6073.50	34.4203	0.0000	40
29	LQPTC_0.5	LQPTC	0.5	0.9451	0.9402	0.0017	0.0090	0.0007	0.008	0.0003	0.108	0.0029	6073.50	34.4203	0.0000	40
142	HQP_4	HQP	4	0.9549	0.9479	0.0004	0.0065	0.0003	0.008	0.0006	0.14	0.0026	6079.52	40.4386	0.0000	43

143	HQPT_4	HQPT	4	0.9549	0.9479	0.0004	0.0065	0.0003	0.008	0.0006	0.14	0.0026	6079.52	40.4386	0.0000	43
144	HQPTC_4	HQPTC	4	0.9549	0.9479	0.0004	0.0065	0.0003	0.008	0.0006	0.14	0.0026	6079.52	40.4386	0.0000	43
94	HQP_2.5	HQP	2.5	0.9642	0.9592	0.0004	0.0055	0.0003	0.004	0.0002	0.128	0.0035	6087.08	48.0005	0.0000	66
95	HQPT_2.5	HQPT	2.5	0.9642	0.9592	0.0004	0.0055	0.0003	0.004	0.0002	0.128	0.0035	6087.08	48.0005	0.0000	66
96	HQPTC_2.5	HQPTC	2.5	0.9642	0.9592	0.0004	0.0055	0.0003	0.004	0.0002	0.128	0.0035	6087.08	48.0005	0.0000	66
59	LQP_1.5	LQP	1.5	0.9406	0.9353	0.0012	0.0089	0.0007	0.012	0.0004	0.108	0.0022	6102.57	63.4844	0.0000	32
60	LQPT_1.5	LQPT	1.5	0.9406	0.9353	0.0012	0.0089	0.0007	0.012	0.0004	0.108	0.0022	6102.57	63.4844	0.0000	32
61	LQPTC_1.5	LQPTC	1.5	0.9406	0.9353	0.0012	0.0089	0.0007	0.012	0.0004	0.108	0.0022	6102.57	63.4844	0.0000	32
158	HQP_4.5	HQP	4.5	0.9516	0.9423	0.0006	0.0077	0.0004	0.008	0.0003	0.136	0.0043	6106.81	67.7212	0.0000	42
159	HQPT_4.5	HQPT	4.5	0.9516	0.9423	0.0006	0.0077	0.0004	0.008	0.0003	0.136	0.0043	6106.81	67.7212	0.0000	42
160	HQPTC_4.5	HQPTC	4.5	0.9516	0.9423	0.0006	0.0077	0.0004	0.008	0.0003	0.136	0.0043	6106.81	67.7212	0.0000	42
174	HQP_5	HQP	5	0.9471	0.9379	0.0007	0.0073	0.0004	0.008	0.0003	0.108	0.0032	6125.50	86.4174	0.0000	39
175	HQPT_5	HQPT	5	0.9471	0.9379	0.0007	0.0073	0.0004	0.008	0.0003	0.108	0.0032	6125.50	86.4174	0.0000	39
176	HQPTC_5	HQPTC	5	0.9471	0.9379	0.0007	0.0073	0.0004	0.008	0.0003	0.108	0.0032	6125.50	86.4174	0.0000	39
75	LQP_2	LQP	2	0.9384	0.9332	0.0010	0.0084	0.0006	0.012	0.0007	0.108	0.0025	6135.97	96.8897	0.0000	30
76	LQPT_2	LQPT	2	0.9384	0.9332	0.0010	0.0084	0.0006	0.012	0.0007	0.108	0.0025	6135.97	96.8897	0.0000	30
77	LQPTC_2	LQPTC	2	0.9384	0.9332	0.0010	0.0084	0.0006	0.012	0.0007	0.108	0.0025	6135.97	96.8897	0.0000	30
91	LQP_2.5	LQP	2.5	0.9352	0.9300	0.0010	0.0084	0.0006	0.008	0.0006	0.112	0.0031	6147.90	108.8160	0.0000	29
92	LQPT_2.5	LQPT	2.5	0.9352	0.9300	0.0010	0.0084	0.0006	0.008	0.0006	0.112	0.0031	6147.90	108.8160	0.0000	29
93	LQPTC_2.5	LQPTC	2.5	0.9352	0.9300	0.0010	0.0084	0.0006	0.008	0.0006	0.112	0.0031	6147.90	108.8160	0.0000	29
190	HQP_5.5	HQP	5.5	0.9421	0.9338	0.0008	0.0073	0.0004	0.004	0.0002	0.108	0.0050	6149.08	109.9936	0.0000	37
191	HQPT_5.5	HQPT	5.5	0.9421	0.9338	0.0008	0.0073	0.0004	0.004	0.0002	0.108	0.0050	6149.08	109.9936	0.0000	37
192	HQPTC_5.5	HQPTC	5.5	0.9421	0.9338	0.0008	0.0073	0.0004	0.004	0.0002	0.108	0.0050	6149.08	109.9936	0.0000	37
107	LQP_3	LQP	3	0.9325	0.9275	0.0008	0.0077	0.0005	0.012	0.0007	0.1	0.0026	6156.08	116.9950	0.0000	25
108	LQPT_3	LQPT	3	0.9325	0.9275	0.0008	0.0077	0.0005	0.012	0.0007	0.1	0.0026	6156.08	116.9950	0.0000	25
109	LQPTC_3	LQPTC	3	0.9325	0.9275	0.0008	0.0077	0.0005	0.012	0.0007	0.1	0.0026	6156.08	116.9950	0.0000	25
206	HQP_6	HQP	6	0.938	0.9296	0.0009	0.0077	0.0004	0.004	0.0002	0.116	0.0055	6171.84	132.7574	0.0000	34

207	HQPT_6	HQPT	6	0.938	0.9296	0.0009	0.0077	0.0004	0.004	0.0002	0.116	0.0055	6171.84	132.7574	0.0000	34
208	HQPTC_6	HQPTC	6	0.938	0.9296	0.0009	0.0077	0.0004	0.004	0.0002	0.116	0.0055	6171.84	132.7574	0.0000	34
123	LQP_3.5	LQP	3.5	0.9299	0.9243	0.0007	0.0077	0.0004	0.012	0.0007	0.108	0.0025	6181.15	142.0691	0.0000	23
124	LQPT_3.5	LQPT	3.5	0.9299	0.9243	0.0007	0.0077	0.0004	0.012	0.0007	0.108	0.0025	6181.15	142.0691	0.0000	23
125	LQPTC_3.5	LQPTC	3.5	0.9299	0.9243	0.0007	0.0077	0.0004	0.012	0.0007	0.108	0.0025	6181.15	142.0691	0.0000	23
139	LQP_4	LQP	4	0.926	0.9204	0.0007	0.0081	0.0004	0.008	0.0006	0.112	0.0024	6205.72	166.6321	0.0000	22
140	LQPT_4	LQPT	4	0.926	0.9204	0.0007	0.0081	0.0004	0.008	0.0006	0.112	0.0024	6205.72	166.6321	0.0000	22
141	LQPTC_4	LQPTC	4	0.926	0.9204	0.0007	0.0081	0.0004	0.008	0.0006	0.112	0.0024	6205.72	166.6321	0.0000	22
105	HQC_3	HQC	3	0.9428	0.9330	0.0013	0.0087	0.0009	0.004	0.0002	0.136	0.0047	6217.80	178.7164	0.0000	64
89	HQC_2.5	HQC	2.5	0.9503	0.9405	0.0012	0.0088	0.0008	0.004	0.0002	0.14	0.0040	6219.10	180.0154	0.0000	79
155	LQP_4.5	LQP	4.5	0.9227	0.9170	0.0007	0.0079	0.0003	0.008	0.0006	0.132	0.0018	6227.21	188.1217	0.0000	22
156	LQPT_4.5	LQPT	4.5	0.9227	0.9170	0.0007	0.0079	0.0003	0.008	0.0006	0.132	0.0018	6227.21	188.1217	0.0000	22
157	LQPTC_4.5	LQPTC	4.5	0.9227	0.9170	0.0007	0.0079	0.0003	0.008	0.0006	0.132	0.0018	6227.21	188.1217	0.0000	22
84	H_2.5	H	2.5	0.9507	0.9394	0.0014	0.0102	0.0011	0.008	0.0003	0.136	0.0043	6241.49	202.4093	0.0000	82
88	HC_2.5	HC	2.5	0.9507	0.9394	0.0014	0.0102	0.0011	0.008	0.0003	0.136	0.0043	6241.49	202.4093	0.0000	82
121	HQC_3.5	HQC	3.5	0.9353	0.9266	0.0011	0.0081	0.0009	0.004	0.0002	0.128	0.0056	6242.02	202.9402	0.0000	53
171	LQP_5	LQP	5	0.9194	0.9137	0.0007	0.0083	0.0003	0.008	0.0006	0.116	0.0019	6245.97	206.8892	0.0000	21
172	LQPT_5	LQPT	5	0.9194	0.9137	0.0007	0.0083	0.0003	0.008	0.0006	0.116	0.0019	6245.97	206.8892	0.0000	21
173	LQPTC_5	LQPTC	5	0.9194	0.9137	0.0007	0.0083	0.0003	0.008	0.0006	0.116	0.0019	6245.97	206.8892	0.0000	21
100	H_3	H	3	0.9435	0.9332	0.0015	0.0099	0.0012	0.004	0.0002	0.128	0.0038	6256.18	217.0997	0.0000	71
104	HC_3	HC	3	0.9435	0.9332	0.0015	0.0099	0.0012	0.004	0.0002	0.128	0.0038	6256.18	217.0997	0.0000	71
116	H_3.5	H	3.5	0.9372	0.9266	0.0017	0.0103	0.0013	0.004	0.0002	0.128	0.0074	6262.88	223.7926	0.0000	57
120	HC_3.5	HC	3.5	0.9372	0.9266	0.0017	0.0103	0.0013	0.004	0.0002	0.128	0.0074	6262.88	223.7926	0.0000	57
153	HQC_4.5	HQC	4.5	0.9262	0.9169	0.0011	0.0082	0.0008	0.004	0.0002	0.12	0.0028	6266.70	227.6164	0.0000	36
137	HQC_4	HQC	4	0.9299	0.9213	0.0011	0.0080	0.0009	0.004	0.0002	0.116	0.0030	6267.12	228.0405	0.0000	47
78	HQP_2	HQP	2	0.9688	0.9623	0.0002	0.0057	0.0002	0.004	0.0002	0.152	0.0024	6268.91	229.8241	0.0000	84
79	HQPT_2	HQPT	2	0.9688	0.9623	0.0002	0.0057	0.0002	0.004	0.0002	0.152	0.0024	6268.91	229.8241	0.0000	84

80	HQPTC_2	HQPTC	2	0.9688	0.9623	0.0002	0.0057	0.0002	0.004	0.0002	0.152	0.0024	6268.91	229.8241	0.0000	84
187	LQP_5.5	LQP	5.5	0.915	0.9093	0.0008	0.0086	0.0003	0.004	0.0002	0.124	0.0034	6272.22	233.1318	0.0000	21
188	LQPT_5.5	LQPT	5.5	0.915	0.9093	0.0008	0.0086	0.0003	0.004	0.0002	0.124	0.0034	6272.22	233.1318	0.0000	21
189	LQPTC_5.5	LQPTC	5.5	0.915	0.9093	0.0008	0.0086	0.0003	0.004	0.0002	0.124	0.0034	6272.22	233.1318	0.0000	21
203	LQP_6	LQP	6	0.9119	0.9063	0.0009	0.0086	0.0004	0.004	0.0002	0.124	0.0034	6284.14	245.0570	0.0000	18
204	LQPT_6	LQPT	6	0.9119	0.9063	0.0009	0.0086	0.0004	0.004	0.0002	0.124	0.0034	6284.14	245.0570	0.0000	18
205	LQPTC_6	LQPTC	6	0.9119	0.9063	0.0009	0.0086	0.0004	0.004	0.0002	0.124	0.0034	6284.14	245.0570	0.0000	18
169	HQC_5	HQC	5	0.9208	0.9117	0.0012	0.0082	0.0008	0.004	0.0002	0.116	0.0030	6287.32	248.2381	0.0000	35
132	H_4	H	4	0.9309	0.9196	0.0017	0.0100	0.0012	0.004	0.0002	0.128	0.0074	6295.48	256.3955	0.0000	52
136	HC_4	HC	4	0.9309	0.9196	0.0017	0.0100	0.0012	0.004	0.0002	0.128	0.0074	6295.48	256.3955	0.0000	52
185	HQC_5.5	HQC	5.5	0.9162	0.9067	0.0013	0.0083	0.0008	0.004	0.0002	0.104	0.0026	6295.91	256.8265	0.0000	31
148	H_4.5	H	4.5	0.9228	0.9141	0.0016	0.0091	0.0013	0.004	0.0002	0.12	0.0082	6308.35	269.2630	0.0000	43
152	HC_4.5	HC	4.5	0.9228	0.9141	0.0016	0.0091	0.0013	0.004	0.0002	0.12	0.0082	6308.35	269.2630	0.0000	43
35	P_1	P	1	0.9145	0.9072	0.0008	0.0087	0.0006	0.012	0.0007	0.128	0.0021	6311.53	272.4458	0.0000	23
164	H_5	H	5	0.918	0.9105	0.0019	0.0087	0.0017	0.004	0.0002	0.128	0.0074	6311.64	272.5577	0.0000	36
168	HC_5	HC	5	0.918	0.9105	0.0019	0.0087	0.0017	0.004	0.0002	0.128	0.0074	6311.64	272.5577	0.0000	36
3	P_0.1	P	0.1	0.9148	0.9088	0.0008	0.0085	0.0007	0.012	0.0007	0.136	0.0022	6313.09	274.0104	0.0000	28
19	P_0.5	P	0.5	0.9153	0.9086	0.0009	0.0086	0.0007	0.008	0.0003	0.128	0.0014	6316.88	277.7956	0.0000	25
201	HQC_6	HQC	6	0.9112	0.9022	0.0013	0.0083	0.0008	0.008	0.0003	0.108	0.0025	6323.12	284.0320	0.0000	33
180	H_5.5	H	5.5	0.9158	0.9079	0.0019	0.0084	0.0017	0.004	0.0002	0.144	0.0097	6323.32	284.2331	0.0000	34
184	HC_5.5	HC	5.5	0.9158	0.9079	0.0019	0.0084	0.0017	0.004	0.0002	0.144	0.0097	6323.32	284.2331	0.0000	34
51	P_1.5	P	1.5	0.912	0.9056	0.0009	0.0089	0.0005	0.008	0.0003	0.136	0.0011	6324.77	285.6856	0.0000	24
73	HQC_2	HQC	2	0.9564	0.9476	0.0009	0.0084	0.0006	0.008	0.0003	0.136	0.0029	6328.80	289.7137	0.0000	93
196	H_6	H	6	0.9132	0.9047	0.0021	0.0084	0.0018	0.004	0.0002	0.108	0.0039	6337.91	298.8219	0.0000	33
200	HC_6	HC	6	0.9132	0.9047	0.0021	0.0084	0.0018	0.004	0.0002	0.108	0.0039	6337.91	298.8219	0.0000	33
67	P_2	P	2	0.9105	0.9028	0.0010	0.0095	0.0005	0.004	0.0002	0.128	0.0014	6342.19	303.1067	0.0000	24
83	P_2.5	P	2.5	0.9073	0.9008	0.0010	0.0092	0.0004	0.004	0.0002	0.128	0.0010	6351.45	312.3667	0.0000	21

99	P_3	P	3	0.9058	0.8993	0.0010	0.0092	0.0004	0.004	0.0002	0.116	0.0009	6352.73	313.6451	0.0000	19
68	H_2	H	2	0.9566	0.9463	0.0011	0.0097	0.0008	0.008	0.0003	0.132	0.0032	6355.60	316.5160	0.0000	97
72	HC_2	HC	2	0.9566	0.9463	0.0011	0.0097	0.0008	0.008	0.0003	0.132	0.0032	6355.60	316.5160	0.0000	97
115	P_3.5	P	3.5	0.9045	0.8983	0.0010	0.0091	0.0004	0.004	0.0002	0.108	0.0018	6356.95	317.8649	0.0000	19
131	P_4	P	4	0.9028	0.8966	0.0011	0.0093	0.0004	0.004	0.0002	0.108	0.0022	6363.32	324.2378	0.0000	19
147	P_4.5	P	4.5	0.9015	0.8952	0.0011	0.0093	0.0004	0.004	0.0002	0.108	0.0029	6364.38	325.3006	0.0000	17
21	LQ_0.5	LQ	0.5	0.8979	0.8924	0.0012	0.0085	0.0007	0.004	0.0002	0.124	0.0055	6364.71	325.6250	0.0000	13
5	LQ_0.1	LQ	0.1	0.8986	0.8930	0.0012	0.0082	0.0007	0.004	0.0002	0.124	0.0055	6364.81	325.7247	0.0000	14
37	LQ_1	LQ	1	0.8969	0.8921	0.0012	0.0080	0.0006	0.004	0.0002	0.116	0.0051	6366.54	327.4572	0.0000	12
163	P_5	P	5	0.9	0.8937	0.0011	0.0092	0.0004	0.004	0.0002	0.108	0.0029	6369.36	330.2741	0.0000	16
53	LQ_1.5	LQ	1.5	0.8953	0.8903	0.0012	0.0082	0.0006	0.004	0.0002	0.124	0.0041	6373.31	334.2268	0.0000	12
179	P_5.5	P	5.5	0.898	0.8921	0.0011	0.0093	0.0004	0.004	0.0002	0.112	0.0031	6377.52	338.4340	0.0000	16
195	P_6	P	6	0.8968	0.8904	0.0012	0.0095	0.0005	0.004	0.0002	0.108	0.0025	6381.58	342.4913	0.0000	15
69	LQ_2	LQ	2	0.8931	0.8877	0.0012	0.0083	0.0007	0.004	0.0002	0.124	0.0030	6385.78	346.6971	0.0000	13
85	LQ_2.5	LQ	2.5	0.8902	0.8852	0.0013	0.0084	0.0007	0.004	0.0002	0.12	0.0028	6394.92	355.8353	0.0000	12
101	LQ_3	LQ	3	0.8866	0.8822	0.0014	0.0083	0.0007	0.004	0.0002	0.116	0.0030	6410.25	371.1644	0.0000	12
117	LQ_3.5	LQ	3.5	0.8837	0.8788	0.0015	0.0084	0.0008	0.004	0.0002	0.132	0.0018	6420.63	381.5451	0.0000	11
133	LQ_4	LQ	4	0.8801	0.8756	0.0016	0.0084	0.0009	0.004	0.0002	0.132	0.0018	6436.44	397.3513	0.0000	11
149	LQ_4.5	LQ	4.5	0.8777	0.8728	0.0017	0.0085	0.0009	0.004	0.0002	0.132	0.0007	6445.72	406.6360	0.0000	10
165	LQ_5	LQ	5	0.8753	0.8704	0.0019	0.0087	0.0010	0.004	0.0002	0.128	0.0017	6455.22	416.1398	0.0000	10
181	LQ_5.5	LQ	5.5	0.8724	0.8678	0.0021	0.0091	0.0010	0.004	0.0002	0.124	0.0005	6466.38	427.2935	0.0000	10
197	LQ_6	LQ	6	0.8703	0.8658	0.0022	0.0093	0.0010	0.004	0.0002	0.108	0.0007	6471.59	432.5109	0.0000	9
98	Q_3	Q	3	0.8533	0.8498	0.0038	0.0116	0.0016	0.004	0.0002	0.096	0.0018	6540.53	501.4437	0.0000	6
103	TQ_3	TQ	3	0.8533	0.8498	0.0038	0.0116	0.0016	0.004	0.0002	0.096	0.0018	6540.53	501.4437	0.0000	6
106	TQ_3	TQ	3	0.8533	0.8498	0.0038	0.0116	0.0016	0.004	0.0002	0.096	0.0018	6540.53	501.4437	0.0000	6
114	Q_3.5	Q	3.5	0.8532	0.8498	0.0039	0.0116	0.0016	0.004	0.0002	0.1	0.0019	6540.84	501.7523	0.0000	6
119	TQ_3.5	TQ	3.5	0.8532	0.8498	0.0039	0.0116	0.0016	0.004	0.0002	0.1	0.0019	6540.84	501.7523	0.0000	6

122	TQ_3.5	TQ	3.5	0.8532	0.8498	0.0039	0.0116	0.0016	0.004	0.0002	0.1	0.0019	6540.84	501.7523	0.0000	6
130	Q_4	Q	4	0.8531	0.8497	0.0039	0.0117	0.0016	0.004	0.0002	0.1	0.0019	6541.19	502.1078	0.0000	6
135	TQ_4	TQ	4	0.8531	0.8497	0.0039	0.0117	0.0016	0.004	0.0002	0.1	0.0019	6541.19	502.1078	0.0000	6
138	TQ_4	TQ	4	0.8531	0.8497	0.0039	0.0117	0.0016	0.004	0.0002	0.1	0.0019	6541.19	502.1078	0.0000	6
34	Q_1	Q	1	0.8538	0.8499	0.0036	0.0116	0.0016	0.004	0.0002	0.088	0.0017	6541.36	502.2738	0.0000	7
39	TQ_1	TQ	1	0.8538	0.8499	0.0036	0.0116	0.0016	0.004	0.0002	0.088	0.0017	6541.36	502.2738	0.0000	7
42	TQ_1	TQ	1	0.8538	0.8499	0.0036	0.0116	0.0016	0.004	0.0002	0.088	0.0017	6541.36	502.2738	0.0000	7
146	Q_4.5	Q	4.5	0.853	0.8496	0.0040	0.0117	0.0016	0.004	0.0002	0.096	0.0015	6541.59	502.5105	0.0000	6
151	TQ_4.5	TQ	4.5	0.853	0.8496	0.0040	0.0117	0.0016	0.004	0.0002	0.096	0.0015	6541.59	502.5105	0.0000	6
154	TQ_4.5	TQ	4.5	0.853	0.8496	0.0040	0.0117	0.0016	0.004	0.0002	0.096	0.0015	6541.59	502.5105	0.0000	6
50	Q_1.5	Q	1.5	0.8537	0.8499	0.0037	0.0116	0.0016	0.004	0.0002	0.088	0.0017	6541.60	502.5135	0.0000	7
55	TQ_1.5	TQ	1.5	0.8537	0.8499	0.0037	0.0116	0.0016	0.004	0.0002	0.088	0.0017	6541.60	502.5135	0.0000	7
58	TQ_1.5	TQ	1.5	0.8537	0.8499	0.0037	0.0116	0.0016	0.004	0.0002	0.088	0.0017	6541.60	502.5135	0.0000	7
66	Q_2	Q	2	0.8536	0.8499	0.0037	0.0116	0.0016	0.004	0.0002	0.088	0.0017	6541.90	502.8132	0.0000	7
71	TQ_2	TQ	2	0.8536	0.8499	0.0037	0.0116	0.0016	0.004	0.0002	0.088	0.0017	6541.90	502.8132	0.0000	7
74	TQ_2	TQ	2	0.8536	0.8499	0.0037	0.0116	0.0016	0.004	0.0002	0.088	0.0017	6541.90	502.8132	0.0000	7
162	Q_5	Q	5	0.8529	0.8495	0.0040	0.0117	0.0016	0.004	0.0002	0.096	0.0015	6542.04	502.9609	0.0000	6
167	TQ_5	TQ	5	0.8529	0.8495	0.0040	0.0117	0.0016	0.004	0.0002	0.096	0.0015	6542.04	502.9609	0.0000	6
170	TQ_5	TQ	5	0.8529	0.8495	0.0040	0.0117	0.0016	0.004	0.0002	0.096	0.0015	6542.04	502.9609	0.0000	6
82	Q_2.5	Q	2.5	0.8534	0.8499	0.0038	0.0116	0.0016	0.004	0.0002	0.096	0.0022	6542.26	503.1722	0.0000	7
87	TQ_2.5	TQ	2.5	0.8534	0.8499	0.0038	0.0116	0.0016	0.004	0.0002	0.096	0.0022	6542.26	503.1722	0.0000	7
90	TQ_2.5	TQ	2.5	0.8534	0.8499	0.0038	0.0116	0.0016	0.004	0.0002	0.096	0.0022	6542.26	503.1722	0.0000	7
178	Q_5.5	Q	5.5	0.8527	0.8494	0.0041	0.0118	0.0016	0.004	0.0002	0.096	0.0015	6542.54	503.4589	0.0000	6
183	TQ_5.5	TQ	5.5	0.8527	0.8494	0.0041	0.0118	0.0016	0.004	0.0002	0.096	0.0015	6542.54	503.4589	0.0000	6
186	TQ_5.5	TQ	5.5	0.8527	0.8494	0.0041	0.0118	0.0016	0.004	0.0002	0.096	0.0015	6542.54	503.4589	0.0000	6
194	Q_6	Q	6	0.8526	0.8492	0.0042	0.0119	0.0017	0.004	0.0002	0.096	0.0015	6543.09	504.0051	0.0000	6
199	TQ_6	TQ	6	0.8526	0.8492	0.0042	0.0119	0.0017	0.004	0.0002	0.096	0.0015	6543.09	504.0051	0.0000	6

202	TQ_6	TQ	6	0.8526	0.8492	0.0042	0.0119	0.0017	0.004	0.0002	0.096	0.0015	6543.09	504.0051	0.0000	6
2	Q_0.1	Q	0.1	0.854	0.8500	0.0036	0.0116	0.0016	0.004	0.0002	0.096	0.0022	6543.14	504.0544	0.0000	8
7	TQ_0.1	TQ	0.1	0.854	0.8500	0.0036	0.0116	0.0016	0.004	0.0002	0.096	0.0022	6543.14	504.0544	0.0000	8
10	TQ_0.1	TQ	0.1	0.854	0.8500	0.0036	0.0116	0.0016	0.004	0.0002	0.096	0.0022	6543.14	504.0544	0.0000	8
18	Q_0.5	Q	0.5	0.8539	0.8500	0.0036	0.0116	0.0016	0.004	0.0002	0.096	0.0022	6543.30	504.2182	0.0000	8
23	TQ_0.5	TQ	0.5	0.8539	0.8500	0.0036	0.0116	0.0016	0.004	0.0002	0.096	0.0022	6543.30	504.2182	0.0000	8
26	TQ_0.5	TQ	0.5	0.8539	0.8500	0.0036	0.0116	0.0016	0.004	0.0002	0.096	0.0022	6543.30	504.2182	0.0000	8
113	L_3.5	L	3.5	0.8489	0.8445	0.0031	0.0112	0.0016	0.004	0.0002	0.12	0.0025	6550.15	511.0634	0.0000	7
118	LC_3.5	LC	3.5	0.8489	0.8445	0.0031	0.0112	0.0016	0.004	0.0002	0.12	0.0025	6550.15	511.0634	0.0000	7
129	L_4	L	4	0.8487	0.8444	0.0032	0.0113	0.0016	0.004	0.0002	0.128	0.0031	6550.64	511.5520	0.0000	7
134	LC_4	LC	4	0.8487	0.8444	0.0032	0.0113	0.0016	0.004	0.0002	0.128	0.0031	6550.64	511.5520	0.0000	7
193	L_6	L	6	0.8481	0.8441	0.0033	0.0115	0.0016	0.004	0.0002	0.096	0.0029	6551.08	511.9912	0.0000	6
198	LC_6	LC	6	0.8481	0.8441	0.0033	0.0115	0.0016	0.004	0.0002	0.096	0.0029	6551.08	511.9912	0.0000	6
145	L_4.5	L	4.5	0.8486	0.8443	0.0032	0.0113	0.0016	0.004	0.0002	0.116	0.0037	6551.20	512.1121	0.0000	7
150	LC_4.5	LC	4.5	0.8486	0.8443	0.0032	0.0113	0.0016	0.004	0.0002	0.116	0.0037	6551.20	512.1121	0.0000	7
33	L_1	L	1	0.8496	0.8453	0.0030	0.0107	0.0016	0.004	0.0002	0.112	0.0021	6551.26	512.1747	0.0000	8
38	LC_1	LC	1	0.8496	0.8453	0.0030	0.0107	0.0016	0.004	0.0002	0.112	0.0021	6551.26	512.1747	0.0000	8
49	L_1.5	L	1.5	0.8495	0.8451	0.0030	0.0108	0.0016	0.004	0.0002	0.108	0.0018	6551.27	512.1892	0.0000	8
54	LC_1.5	LC	1.5	0.8495	0.8451	0.0030	0.0108	0.0016	0.004	0.0002	0.108	0.0018	6551.27	512.1892	0.0000	8
17	L_0.5	L	0.5	0.8498	0.8455	0.0030	0.0105	0.0015	0.004	0.0002	0.108	0.0025	6551.31	512.2308	0.0000	8
22	LC_0.5	LC	0.5	0.8498	0.8455	0.0030	0.0105	0.0015	0.004	0.0002	0.108	0.0025	6551.31	512.2308	0.0000	8
65	L_2	L	2	0.8494	0.8450	0.0030	0.0109	0.0016	0.004	0.0002	0.104	0.0018	6551.39	512.3100	0.0000	8
70	LC_2	LC	2	0.8494	0.8450	0.0030	0.0109	0.0016	0.004	0.0002	0.104	0.0018	6551.39	512.3100	0.0000	8
1	L_0.1	L	0.1	0.8499	0.8456	0.0030	0.0104	0.0015	0.004	0.0002	0.112	0.0028	6551.45	512.3626	0.0000	8
6	LC_0.1	LC	0.1	0.8499	0.8456	0.0030	0.0104	0.0015	0.004	0.0002	0.112	0.0028	6551.45	512.3626	0.0000	8
81	L_2.5	L	2.5	0.8493	0.8449	0.0031	0.0110	0.0016	0.004	0.0002	0.1	0.0019	6551.60	512.5150	0.0000	8
86	LC_2.5	LC	2.5	0.8493	0.8449	0.0031	0.0110	0.0016	0.004	0.0002	0.1	0.0019	6551.60	512.5150	0.0000	8

161	L_5	L	5	0.8484	0.8442	0.0032	0.0114	0.0016	0.004	0.0002	0.1	0.0033	6551.84	512.7571	0.0000	7
166	LC_5	LC	5	0.8484	0.8442	0.0032	0.0114	0.0016	0.004	0.0002	0.1	0.0033	6551.84	512.7571	0.0000	7
97	L_3	L	3	0.8491	0.8447	0.0031	0.0111	0.0016	0.004	0.0002	0.108	0.0014	6551.89	512.8081	0.0000	8
102	LC_3	LC	3	0.8491	0.8447	0.0031	0.0111	0.0016	0.004	0.0002	0.108	0.0014	6551.89	512.8081	0.0000	8
177	L_5.5	L	5.5	0.8482	0.8442	0.0032	0.0115	0.0016	0.004	0.0002	0.1	0.0033	6552.56	513.4719	0.0000	7
182	LC_5.5	LC	5.5	0.8482	0.8442	0.0032	0.0115	0.0016	0.004	0.0002	0.1	0.0033	6552.56	513.4719	0.0000	7
62	HQP_1.5	HQP	1.5	0.9744	0.9665	0.0002	0.0061	0.0003	0.008	0.0003	0.14	0.0033	6677.11	638.0266	0.0000	103
63	HQPT_1.5	HQPT	1.5	0.9744	0.9665	0.0002	0.0061	0.0003	0.008	0.0003	0.14	0.0033	6677.11	638.0266	0.0000	103
64	HQPTC_1.5	HQPTC	1.5	0.9744	0.9665	0.0002	0.0061	0.0003	0.008	0.0003	0.14	0.0033	6677.11	638.0266	0.0000	103
46	HQP_1	HQP	1	0.9765	0.9706	0.0004	0.0061	0.0003	0.004	0.0002	0.128	0.0074	6799.24	760.1562	0.0000	111
47	HQPT_1	HQPT	1	0.9765	0.9706	0.0004	0.0061	0.0003	0.004	0.0002	0.128	0.0074	6799.24	760.1562	0.0000	111
48	HQPTC_1	HQPTC	1	0.9765	0.9706	0.0004	0.0061	0.0003	0.004	0.0002	0.128	0.0074	6799.24	760.1562	0.0000	111
57	HQC_1.5	HQC	1.5	0.9626	0.9533	0.0008	0.0087	0.0006	0.004	0.0002	0.132	0.0046	6803.77	764.6811	0.0000	113
52	H_1.5	H	1.5	0.9644	0.9544	0.0008	0.0088	0.0006	0.004	0.0002	0.14	0.0047	6841.65	802.5637	0.0000	117
56	HC_1.5	HC	1.5	0.9644	0.9544	0.0008	0.0088	0.0006	0.004	0.0002	0.14	0.0047	6841.65	802.5637	0.0000	117
36	H_1	H	1	0.9693	0.9623	0.0010	0.0083	0.0007	0	0.0000	0.14	0.0051	7001.26	962.1772	0.0000	129
40	HC_1	HC	1	0.9693	0.9623	0.0010	0.0083	0.0007	0	0.0000	0.14	0.0051	7001.26	962.1772	0.0000	129
41	HQC_1	HQC	1	0.9684	0.9614	0.0011	0.0087	0.0008	0	0.0000	0.152	0.0060	7075.89	1036.8024	0.0000	138
30	HQP_0.5	HQP	0.5	0.9757	0.9677	0.0010	0.0090	0.0008	0.008	0.0003	0.124	0.0030	7216.96	1177.8755	0.0000	148
31	HQPT_0.5	HQPT	0.5	0.9757	0.9677	0.0010	0.0090	0.0008	0.008	0.0003	0.124	0.0030	7216.96	1177.8755	0.0000	148
32	HQPTC_0.5	HQPTC	0.5	0.9757	0.9677	0.0010	0.0090	0.0008	0.008	0.0003	0.124	0.0030	7216.96	1177.8755	0.0000	148
20	H_0.5	H	0.5	0.9701	0.9601	0.0014	0.0104	0.0013	0.008	0.0003	0.14	0.0029	7244.74	1205.6511	0.0000	141
24	HC_0.5	HC	0.5	0.9701	0.9601	0.0014	0.0104	0.0013	0.008	0.0003	0.14	0.0029	7244.74	1205.6511	0.0000	141
25	HQC_0.5	HQC	0.5	0.9696	0.9606	0.0013	0.0101	0.0012	0.012	0.0004	0.14	0.0058	7497.19	1458.1048	0.0000	155
9	HQC_0.1	HQC	0.1	0.9605	0.6956	0.2614	0.0191	0.0054	0	0.0000	0.036	0.0044	9626.05	3586.9652	0.0000	206
14	HQP_0.1	HQP	0.1	0.9677	0.7521	0.3240	0.0203	0.0058	0.004	0.0002	0.044	0.0041	10057.97	4018.8861	0.0000	214
15	HQPT_0.1	HQPT	0.1	0.9677	0.7521	0.3240	0.0203	0.0058	0.004	0.0002	0.044	0.0041	10057.97	4018.8861	0.0000	214

16	HQPTC_0.1	HQPTC	0.1	0.9677	0.7521	0.3240	0.0203	0.0058	0.004	0.0002	0.044	0.0041	10057.97	4018.8861	0.0000	214
4	H_0.1	H	0.1	0.9287	0.6754	0.2259	0.0185	0.0069	0	0.0000	0.016	0.0011	NA	NA	NA	179
8	HC_0.1	HC	0.1	0.9287	0.6754	0.2259	0.0185	0.0069	0	0.0000	0.016	0.0011	NA	NA	NA	179

Table S3.2 AUC values of environmental variables for individual population models

Population	**Training AUC**	**Test AUC**	**AUC Standard Deviation**	**AUC with only Cult_T**	**AUC with only Soil_Organic Carbon**	**AUC with only bio15**	**AUC with only bio16**	**AUC with only bio17**	**AUC with only bio6**	**Entropy**
Adane	0.9887	0.9931	0.0007	0.9268	0.9862	0.5	0.8162	0.9585	0.9577	7.4064
AlfaMidir	0.9994	0.999	0.0002	0.958	0.9592	0.6191	0.8831	0.9855	0.9952	4.5331
Am_shi	0.9986	0.998	0.0003	0.8124	0.8762	0.5588	0.9915	0.8939	0.9612	5.0265
Arabo	0.9838	0.9849	0.0021	0.9318	0.8278	0.9225	0.9894	0.9964	0.7011	7.3832
Ashuda	0.9982	0.9983	0.0003	0.9815	0.9366	0.6337	0.9646	0.9791	0.9149	5.1076
Batambie	0.9899	0.9891	0.001	0.8727	0.9062	0.4944	0.9755	0.9748	0.9004	6.3433
BekeleGirissa	0.9959	0.9949	0.0005	0.9854	0.6556	0.9296	0.9958	0.9866	0.7841	7.1866
Dikuli	0.9992	0.9989	0.0005	0.9889	0.9435	0.6142	0.9613	0.9601	0.9183	4.8128
Gafera	0.9988	0.9994	0.0002	0.8386	0.8756	0.7604	0.996	0.9187	0.9683	4.4546
Gesses	0.9995	0.9996	0.0001	0.995	0.8481	0.6142	0.9985	0.7954	0.6927	3.8396
Gijet	0.9995	0.9996	0.0001	0.9237	0.6979	0.9979	0.9906	0.9673	0.8711	4.2835
HadushAdi	0.9689	0.9722	0.0018	0.8618	0.8315	0.8145	0.9754	0.8	0.7369	7.925
Hugub	0.9945	0.9893	0.0035	0.9849	0.775	0.7424	0.7583	0.9664	0.5	7.6898
Jarso	0.9948	0.9948	0.0005	0.861	0.871	0.8084	0.9771	0.9786	0.7348	6.8453
Kido	0.9972	0.9984	0.0003	0.8706	0.6619	0.5992	0.9968	0.9673	0.685	5.2794
Kumato	0.9869	0.9865	0.0008	0.9814	0.9762	0.9383	0.6641	0.9453	0.7641	7.5567
LocalHorro	0.9919	0.9916	0.0036	0.9331	0.9053	0.5117	0.946	0.7897	0.8882	6.9746
Loya	0.9984	0.9936	0.0022	0.9837	0.8806	0.9208	0.9774	0.9008	0.9025	6.1841
Meseret	0.9995	0.9995	0.0002	0.936	0.7707	0.9935	0.9947	0.8231	0.92	4.049
Metkilimat	0.997	0.9964	0.001	0.8738	0.7316	0.9829	0.9868	0.8231	0.7141	6.5701
Mihquan	0.9754	0.9722	0.0034	0.8297	0.8338	0.7749	0.959	0.9559	0.6722	7.496
NegasiAmba	0.996	0.9967	0.0009	0.8416	0.6946	0.6385	0.862	0.9772	0.9893	5.6731
ShubiGemo	0.9981	0.9979	0.0003	0.9769	0.716	0.8211	0.9895	0.9949	0.7788	6.4751
Surta	0.9988	0.9989	0.0002	0.8082	0.9176	0.5862	0.9946	0.9237	0.962	4.7008
TsionTeguaz	0.9958	0.9955	0.0006	0.8317	0.6973	0.8117	0.8699	0.9908	0.6784	6.2425

Chapter 4

Whole-genome sequences reveal the drivers of environmental adaptation in Ethiopian village chicken

Table S4.1. Environmental data for six selected agro-climatic variables (selected by ENM approach) for 25 Ethiopian village chicken populations. (Values presented are average and deviation standard calculated across 10 geographic points for each population)

Population	**minTemp (°C)**	**precSeasonality (mm/m^2)**	**precWQ (mm/m^2)**	**precDQ (mm/m^2)**	**SoilOrgC (g/Kg)**	**LandUse (%)**
Adane	5.3 ±1.6	90.1 ± 3.9	653.7 ± 20.39	73.8 ± 10.89	148.5 ± 6.1	37.9 ± 0.33
AlfaMidir	0.8 ± 0.18	104.6 ± 0.52	788.9 ± 3.81	49 ± 0.67	143.1 ± 8.08	40.6 ± 0.99
AmeshaShinkuri	5.09 ± 0.22	86.6 ± 1.51	1060 ± 14.45	85 ± 7.79	128.3 ± 3.53	32.92 ± 0.71
Arabo	11.02 ± 0.11	84.20 ± 0.42	516 ± 7.42	73 ± 0.67	121.8 ± 6.68	38.16 ± 0.13
Ashuda	7.9 ± 1.29	106.7 ± 0.67	975 ± 1.63	23.90 ± 1.29	139.3 ± 7.53	13.07 ± 0.86
Batambie	7.33 ± 0.12	92.9 ± 0.88	1000.8 ± 1.99	46.5 ± 2.07	131.4 ± 3.34	29.26 ± 2.03
BekeleGirissa	9.8 ± 0.00	70 ± 0.67	430.1 ± 4.01	39.1 ± 0.32	108 ± 3.33	14.96 ± 0.79
Dikuli	6.87 ± 1.7	103.9 ± 0.74	965.6 ± 2.41	25.6 ± 1.58	139.6 ± 4.33	10.17 ± 0.58
Gafera	5.04 ± 0.15	85.5 ± 1.08	1067.2 ± 9.38	90.1 ± 5.09	128.8 ± 3.22	34.36 ± 0.38
Gesses	11.17 ± 0.6	104.1 ± 0.88	1111.2 ± 8.56	12.70 ± 3.53	125.6 ± 5.25	0.26 ± 0.02
Gijet	8.11 ± 0.68	143.6 ± 1.84	502.3 ± 4.42	9.0 ± 1.05	111.6 ± 3.6	37.81 ± 0.43
HadushAdi	10.74 ± 0.65	108.6 ± 4.27	443.6 ± 13.06	41.0 ± 7.07	108.0 ± 3.4	35.02 ± 0.71
Hugub	13.53 ± 0.24	72.9 ± 1.79	278.5 ± 19.57	38.4 ± 3.78	97.9 ± 6.38	25.04 ± 1.39
Jarso	10.38 ± 0.72	64.9 ± 1.2	349.6 ± 16.87	38.1 ± 1.45	129.4 ± 3.89	34.95 ± 0.32
Kido	11.71 ± 0.27	103.10 ± 0.57	1068.20 ± 8.51	10.6 ± 1.71	117.90 ± 10.99	2.31 ± 1.86
Kumato	9.86 ± 0.98	49.60 ± 1.51	466.70 ± 46.97	108.00 ± 6.93	78.10 ± 6.94	26.03 ± 0.73
Horro	7.64 ± 0.74	92.0 ± 1.83	924.0 ± 45.17	38.60 ± 10.32	131.40 ± 5.74	38.31 ± 0.82
Loya	7.49 ± 0.13	50.90 ± 0.57	547.30 ± 7.57	125.90 ± 1.37	65.30 ± 15.71	23.99 ± 0.58
Meseret	6.72 ± 0.14	138.7 ± 1.49	419.6 ± 1.43	9.90 ± 0.32	116.0 ± 2.0	38.51 ± 1.62
Metkilimat	10.55 ± 0.74	136.9 ± 2.96	480.7 ± 8.43	19.6 ± 4.33	111.0 ± 2.91	29.19 ± 2.05
Mihquan	11.8 ± 0.26	99.80 ± 2.86	412.0 ± 10.23	54.30 ± 4.88	110.10 ± 5.15	33.89 ± 0.33
NegasiAmba	2.85 ± 0.21	106.90	751.3 ± 4.47	43.2 ± 0.42	111.7 ± 2.06	34.45 ± 0.32
ShubiGemo	9.90 ± 0.00	72.8 ± 0.79	427.1 ± 3.93	36.8 ± 0.42	110.0 ± 2.11	9.72 ± 0.76
Surta	5.09 ± 0.20	85.5 ± 0.85	1064.9 ± 7.68	94.1 ± 4.61	133.7 ± 7.32	35.17 ± 2.88
TzionTeguaz	11.29 ± 0.19	121.6 ± 0.70	767.5 ± 1.65	11.3 ± 0.82	105.7 ± 6.72	34.22 ± 2.17

Table S4.2a. Candidate selection signature regions and overlapping genes detected from analyses of the minimum temperature of the coldest month (minTemp). (**SSR_ID:** selection signature regions ID as chromosome:start_stop position. **DD:** directionality of selection)

SSR_ID	**size (kb)**	**detection method**	**DD**	**no. of SNPs per *XP-EHH* window (range)**	**absolute value of mean *XP-EHH*_std (range)**	**no. of SNPs per *Fst* window (range)**	**weighted *Fst* per window (range)**	***ZFst* per window (range)**	**Overlapping genes**
chr1:430000_450000	20	*xp-ehh*	low	98 - 98	3.34 - 3.34	384 - 384	0.08 - 0.08	0.8 - 0.8	
chr1:460000_510000	50	*xp-ehh*	low	98 - 137	3.25 - 3.42	437 - 509	0.15 - 0.2	2.16 - 3.21	*ENSGALG00000033919*
chr1:8920000_8990000	70	*xp-ehh*	high	57 - 119	3.08 - 3.18	106 - 201	0.07 - 0.18	0.41 - 2.74	*SEMA3D*
chr1:27520000_27540000	20	*fst*	low	139 - 139	1.19 - 1.19	205 - 205	0.3 - 0.3	5.38 - 5.38	*IMMP2L*
chr1:37160000_37190000	30	*fst*	low	138 - 147	1.86 - 2.03	184 - 195	0.37 - 0.38	6.78 - 6.99	*ENSGALG00000052790 (LncRNA)*
chr1:37610000_37640000	30	*xp-ehh*	high	157 - 167	3.12 - 3.18	317 - 358	0.07 - 0.07	0.46 - 0.5	
chr1:49840000_50020000	180	*xp-ehh*	low	100 - 259	3.22 - 3.89	197 - 395	0.12 - 0.23	1.48 - 3.93	*ENSGALG00000046744 (LncRNA), gga-mir-1683 (MiRNA), MRTFA, RBX1, SLC25A17, ST13, XPNPEP3*
chr1:50030000_50060000	30	*xp-ehh*	low	139 - 164	3.23 - 3.37	242 - 264	0.16 - 0.18	2.41 - 2.82	*MRTFA*
chr1:56010000_56060000	50	*xp-ehh*	low	113 - 178	3.01 - 3.49	241 - 319	0.19 - 0.25	2.99 - 4.32	*HIPK2*
chr1:56180000_56250000	70	*xp-ehh*	low	218 - 335	3.13 - 3.32	336 - 502	0.12 - 0.22	1.51 - 3.64	*ENSGALG00000012796*
chr1:77820000_77840000	20	*xp-ehh*	high	278 - 278	3.11 - 3.11	373 - 373	0.12 - 0.12	1.56 - 1.56	*COPS7A, MLF2*
chr1:97180000_97220000	40	*xp-ehh*	low	71 - 92	3.02 - 3.14	142 - 224	0.15 - 0.24	2.13 - 4.06	*ENSGALG00000051825 (LncRNA)*
chr1:109000000_109120000	120	*xp-ehh*	low	96 - 158	3.3 - 4.13	231 - 293	0.15 - 0.24	2.16 - 4.09	*ENSGALG00000050080 (LncRNA), ERG*
chr1:109130000_109150000	20	*xp-ehh*	low	116 - 116	3.06 - 3.06	290 - 290	0.18 - 0.18	2.9 - 2.9	*ENSGALG00000050080 (LncRNA), ETS2*
chr1:109430000_109540000	110	*xp-ehh*	low	106 - 260	3.08 - 3.46	237 - 386	0.14 - 0.23	1.96 - 3.91	*ENSGALG00000049718 (LncRNA), ENSGALG00000049863 (LncRNA), IGSF5, PCP4*
chr1:109620000_109640000	20	*xp-ehh*	low	155 - 155	3.01 - 3.01	271 - 271	0.18 - 0.18	2.82 - 2.82	*DSCAM*
chr1:109760000_109800000	40	*xp-ehh*	low	145 - 178	3 - 3.08	262 - 274	0.18 - 0.2	2.78 - 3.32	*DSCAM, ENSGALG00000042372 (LncRNA), ENSGALG00000052898 (LncRNA)*

chr1:109810000_109830000	20	*xp-ehh*	low	158 - 158	3.02 - 3.02	262 - 262	0.23 - 0.23	4 - 4	*DSCAM, ENSGALG00000042372 (LncRNA), ENSGALG00000052898 (LncRNA)*
chr1:111010000_111030000	20	*fst*	low	262 - 262	2.58 - 2.58	387 - 387	0.29 - 0.29	5.09 - 5.09	*ENSGALG00000016196*
chr1:126080000_126100000	20	*xp-ehh*	high	280 - 280	3.1 - 3.1	414 - 414	0.06 - 0.06	0.2 - 0.2	*MID1*
chr1:126370000_126390000	20	*xp-ehh*	high	235 - 235	3.05 - 3.05	332 - 332	0.08 - 0.08	0.63 - 0.63	*SHROOM2*
chr1:130370000_130430000	60	*xp-ehh*	high	163 - 239	3.01 - 3.26	289 - 407	0.05 - 0.08	0.16 - 0.76	*ASMTL, ENSGALG00000049423 (LncRNA), P2RY8*
chr1:145470000_145500000	30	*xp-ehh*	low	212 - 243	3.11 - 3.19	330 - 336	0.22 - 0.24	3.73 - 4.19	*PCCA*
chr1:166170000_166190000	20	*fst*	low	101 - 101	0.33 - 0.33	151 - 151	0.31 - 0.31	5.49 - 5.49	*ENSGALG00000027756 (ncRNA)*
chr1:186620000_186650000	30	*xp-ehh*	low	244 - 254	3.01 - 3.4	415 - 417	0.13 - 0.14	1.74 - 1.96	
chr1:190110000_190130000	20	*fst*	low	154 - 154	0.39 - 0.39	303 - 303	0.34 - 0.34	6.28 - 6.28	
chr1:195800000_195850000	50	*fst*	low	131 - 210	1.77 - 2.16	270 - 366	0.3 - 0.32	5.44 - 5.88	*ENSGALG00000030652, PGM2L1*
chr1:196790000_196860000	70	*xp-ehh*	low	178 - 240	3.1 - 3.21	309 - 357	0.11 - 0.2	1.36 - 3.16	*GDPD5, KLHL35*
chr1:196910000_196940000	30	*xp-ehh*	low	149 - 177	3.15 - 3.21	269 - 314	0.19 - 0.2	3.1 - 3.24	*MAP6, MOGAT2*
chr1:197220000_197290000	70	*xp-ehh / fst*	low	104 - 194	2.3 - 3.02	149 - 394	0.27 - 0.41	4.66 - 7.61	*ENSGALG00000031533, ENSGALG00000052767, ENSGALG00000053527, ENSGALG00000053623, ENSGALG00000054014, HBBA, HBBR, HBE, HBE1, OR51E2, OR51M1, R52R1*
chr1:197370000_197420000	50	*xp-ehh*	low	97 - 176	3.26 - 3.88	315 - 413	0.24 - 0.26	4.07 - 4.48	*APBB1, CCKBR, ENSGALG00000022704, FAM160A2, ENSGALG00000021748 (SnoRNA), ENSGALG00000021766 (SnoRNA), RPS11, SMPD1, TRIM3*
chr1:197430000_197470000	40	*xp-ehh*	low	114 - 137	3.01 - 4.24	309 - 351	0.17 - 0.25	2.57 - 4.33	*ARFIP2, ENSGALG00000017336, ENSGALG00000042572, ENSGALG00000048769, ILK, TRIM3*
chr2:1210000_1240000	30	*xp-ehh*	low	200 - 232	3.02 - 3.18	475 - 480	0.07 - 0.09	0.48 - 0.97	*ENSGALG00000053636, GHRHR*
chr2:4610000_4650000	40	*xp-ehh*	low	175 - 236	3.24 - 3.56	298 - 374	0.17 - 0.26	2.71 - 4.51	*CTDSPL, gga-mir-26a*
chr2:5050000_5070000	20	*xp-ehh*	low	247 - 247	3.11 - 3.11	367 - 367	0.14 - 0.14	2.09 - 2.09	*XIRP1*
chr2:42090000_42120000	30	*xp-ehh*	low	294 - 304	3 - 3.01	404 - 411	0.21 - 0.21	3.5 - 3.58	*ENSGALG00000052910 (LncRNA)*
chr2:42160000_42190000	30	*xp-ehh*	low	270 - 286	3.08 - 3.15	392 - 421	0.15 - 0.16	2.22 - 2.34	*ACPP*

chr2:43430000_43550000	120	*xp-ehh*	low	149 - 206	3.04 - 3.27	324 - 457	0.14 - 0.25	1.9 - 4.28	*ENSGALG00000036110 (LncRNA), ENSGALG00000048965 (LncRNA), ENSGALG00000053882*
chr2:43920000_43960000	40	*xp-ehh*	low	206 - 262	3.16 - 3.31	390 - 479	0.13 - 0.18	1.87 - 2.77	*EIF1B, ENSGALG00000041372, ENSGALG00000043381, ENSGALG00000054003, gga-mir-7457*
chr2:49530000_49560000	30	*fst*	low	97 - 123	0.96 - 0.99	166 - 196	0.31 - 0.31	5.53 - 5.63	
chr2:49570000_49590000	20	*fst*	low	55 - 55	1.02 - 1.02	109 - 109	0.31 - 0.31	5.61 - 5.61	
chr2:58650000_58690000	40	*xp-ehh*	high	143 - 172	3.27 - 3.55	277 - 294	0.14 - 0.18	2 - 2.76	*ENSGALG00000048667 (LncRNA)*
chr2:60780000_60820000	40	*fst*	low	6 - 55	0.17 - 0.75	143 - 148	0.3 - 0.39	5.43 - 7.27	
chr2:68080000_68110000	30	*fst*	low	195 - 242	1.22 - 1.42	330 - 393	0.33 - 0.34	6.07 - 6.32	*PHLPP1*
chr2:77820000_77850000	30	*fst*	low	23 - 49	0.98 - 1.03	120 - 153	0.35 - 0.37	6.49 - 6.85	*ENSGALG00000050420, gga-mir-1613*
chr2:127490000_127530000	40	*xp-ehh* and/or *fst*	low	158 - 166	3.01 - 3.32	252 - 260	0.21 - 0.32	3.43 - 5.74	
chr2:130550000_130570000	20	*fst*	low	60 - 60	0.68 - 0.68	86 - 86	0.3 - 0.3	5.29 - 5.29	
chr3:36710000_36740000	30	*fst*	low	47 - 90	0.7 - 1.66	120 - 162	0.3 - 0.32	5.37 - 5.75	*CHRM1*
chr3:36750000_36780000	30	*fst*	high	60 - 79	0.49 - 0.69	159 - 159	0.31 - 0.31	5.53 - 5.66	
chr3:43730000_43760000	30	*fst*	low	121 - 140	1.8 - 1.96	274 - 304	0.29 - 0.29	5.11 - 5.16	*ENSGALG00000051296 (LncRNA), ENSGALG00000053691 (LncRNA)*
chr3:60300000_60320000	20	*xp-ehh*	low	127 - 127	3.06 - 3.06	249 - 249	0.11 - 0.11	1.46 - 1.46	*ENSGALG00000051906 (LncRNA), NKAIN2*
chr3:68260000_68290000	30	*xp-ehh*	high	164 - 175	3.11 - 3.14	278 - 279	0.11 - 0.15	1.45 - 2.26	*ATG5*
chr3:94720000_94760000	40	*xp-ehh*	low	144 - 170	3.16 - 3.19	221 - 254	0.14 - 0.18	1.9 - 2.89	
chr3:96690000_96710000	20	*xp-ehh*	high	239 - 239	3.14 - 3.14	414 - 414	0.03 - 0.03	-0.23 - -0.23	*ENSGALG00000048509 (LncRNA)*
chr3:107830000_107860000	30	*fst*	high	93 - 140	0.26 - 0.27	208 - 236	0.29 - 0.3	5.12 - 5.4	*ENSGALG00000050753, ENSGALG00000053865 (LncRNA), ENSGALG00000027879 (ncRNA)*
chr3:108130000_108320000	190	*xp-ehh*	low	67 - 206	3.01 - 3.7	150 - 362	0.13 - 0.26	1.76 - 4.52	*ENSGALG00000047589, ENSGALG00000047869 (LncRNA), TFAP2B, TFAP2D*
chr4:8010000_8040000	30	*xp-ehh*	high	140 - 154	3.13 - 3.22	308 - 322	0.12 - 0.12	1.47 - 1.58	
chr4:14070000_14110000	40	*xp-ehh*	low	276 - 303	3.44 - 4.01	384 - 394	0.14 - 0.2	2 - 3.33	*ATG4A, ENSGALG00000048179*
chr4:25290000_25320000	30	*xp-ehh*	high	223 - 239	3.08 - 3.16	345 - 362	0.02 - 0.04	-0.52 - -0.05	

chr4:43410000_43510000	100	*fst*	low	91 - 157	0.32 - 1.08	111 - 204	0.28 - 0.43	5.06 - 8.07	*ENSGALG00000046699 (LncRNA),* *ENSGALG00000048030 (LncRNA),* *ENSGALG00000050160 (LncRNA),* *ENSGALG00000052572 (LncRNA)*
chr4:45270000_45300000	30	*xp-ehh*	high	208 - 212	3.34 - 3.53	296 - 317	0.05 - 0.07	0.12 - 0.53	*ANTXR2, ENSGALG00000051314 (LncRNA)*
chr4:52710000_52730000	20	*fst*	low	206 - 206	1.9 - 1.9	347 - 347	0.31 - 0.31	5.55 - 5.55	*ENSGALG00000052662 (LncRNA)*
chr4:63130000_63150000	20	*xp-ehh*	low	186 - 186	3.15 - 3.15	322 - 322	0.22 - 0.22	3.76 - 3.76	*CNOT7, MTMR7, VPS37A*
chr4:63860000_64090000	230	*xp-ehh*	low	100 - 222	3.07 - 3.99	138 - 315	0.06 - 0.22	0.4 - 3.71	*gga-mir-383, SGCZ*
chr4:65660000_65710000	50	*xp-ehh*	low	207 - 274	3.09 - 3.15	344 - 476	0.12 - 0.13	1.67 - 1.89	*LNX1*
chr4:89590000_89610000	20	*fst*	low	177 - 177	0.48 - 0.48	293 - 293	0.3 - 0.3	5.45 - 5.45	*ADAM33, ENSGALG00000044620*
chr5:16420000_16440000	20	*fst*	low	183 - 183	1.99 - 1.99	289 - 289	0.29 - 0.29	5.18 - 5.18	
chr5:16900000_16930000	30	*xp-ehh*	low	225 - 275	3.01 - 3.03	326 - 432	0.24 - 0.26	4.17 - 4.44	*ENSGALG00000007127,* *ENSGALG00000020485*
chr5:16940000_16980000	40	*fst*	low	206 - 231	1.83 - 2.77	339 - 409	0.29 - 0.32	5.12 - 5.77	*FADS2, RAB3IL1*
chr5:17180000_17210000	30	*fst*	low	218 - 239	2.08 - 2.32	397 - 419	0.29 - 0.32	5.24 - 5.86	*CTNND1, ENSGALG00000007287, LPXN*
chr5:17230000_17290000	60	*xp-ehh* and/or *fst*	low	119 - 193	3.04 - 3.45	212 - 345	0.26 - 0.32	4.58 - 5.74	*CLP1, ENSGALG00000007381, RTN4RL2, P2RX3,* *UBE2L6, MED19, PGR2/3, SLC43A3, SSRP1, TIMM10, YPEL4, ZDHHC5*
chr5:17530000_17580000	50	*xp-ehh*	low	121 - 152	3.17 - 3.88	237 - 304	0.23 - 0.26	3.82 - 4.52	*ENSGALG00000051988 (LncRNA),* *ENSGALG00000027034 (ncRNA)*
chr5:34720000_34740000	20	*fst*	low	189 - 189	1.75 - 1.75	266 - 266	0.29 - 0.29	5.25 - 5.25	*NUBPL*
chr5:58940000_59000000	60	*xp-ehh*	high	180 - 262	3.36 - 4.02	256 - 416	0.02 - 0.08	-0.58 - 0.73	*FANCM, FKBP3, PRPF39,* *ENSGALG00000025674 (SnoRNA), TOGARAM1*
chr6:25320000_25360000	40	*xp-ehh*	low	250 - 329	3.03 - 3.08	369 - 424	0.22 - 0.24	3.64 - 4.05	*ENSGALG00000050593 (LncRNA)*
chr6:27270000_27380000	110	*xp-ehh* and/or *fst*	low	163 - 258	3.05 - 3.98	231 - 369	0.15 - 0.3	2.3 - 5.39	*ENSGALG00000049592 (LncRNA),* *ENSGALG00000051688 (LncRNA),* *ENSGALG00000052216 (LncRNA)*
chr6:35940000_35990000	50	*xp-ehh*	high	228 - 323	3.2 - 3.99	382 - 467	0.02 - 0.08	-0.58 - 0.64	
chr6:36170000_36220000	50	*xp-ehh*	high	35 - 77	3.03 - 3.1	77 - 122	0.01 - 0.06	-0.68 - 0.34	*INPP5A, NKX6-2*
chr7:3300000_3330000	30	*xp-ehh*	high	344 - 371	3.03 - 3.12	537 - 548	0.03 - 0.04	-0.3 - -0.12	*ERBB4*
chr7:6750000_6780000	30	*fst*	low	16 - 24	0.37 - 0.37	25 - 37	0.38 - 0.42	7.08 - 7.9	*COL6A1*

chr7:9030000_9050000	20	*fst*	low	81 - 81	0.48 - 0.48	138 - 138	0.31 - 0.31	5.66 - 5.66	
chr8:22820000_22850000	30	*fst*	low	106 - 120	1.31 - 1.32	139 - 175	0.3 - 0.32	5.49 - 5.71	
chr8:23750000_23820000	70	*xp-ehh*	high	141 - 210	3.15 - 3.55	257 - 385	0.06 - 0.14	0.25 - 1.97	*BEND5*
chr8:23870000_23890000	20	*xp-ehh*	high	226 - 226	3.01 - 3.01	397 - 397	0.06 - 0.06	0.35 - 0.35	*BEND5*
chr8:24330000_24350000	20	*fst*	low	132 - 132	1.59 - 1.59	273 - 273	0.32 - 0.32	5.8 - 5.8	*RNF11*
chr10:17810000_17830000	20	*fst*	low	196 - 196	2.02 - 2.02	319 - 319	0.29 - 0.29	5.27 - 5.27	
chr10:18040000_18070000	30	*xp-ehh*	low	173 - 183	3.26 - 3.5	272 - 283	0.09 - 0.1	0.89 - 1.17	*CHSY1, ENSGALG00000007248, PCSK6, SELENOS, SNRPA1*
chr10:18140000_18180000	40	*xp-ehh*	low	192 - 245	3.01 - 3.24	292 - 392	0.16 - 0.19	2.36 - 3.15	*ANKDD1A, CLPX, SPG21*
chr12:3100000_3130000	30	*xp-ehh*	low	245 - 260	3.49 - 3.63	416 - 431	0.2 - 0.23	3.35 - 3.8	*RAD54L2, RBM6*
chr13:17070000_17100000	30	*xp-ehh*	low	215 - 224	3.06 - 3.16	337 - 371	0.12 - 0.14	1.53 - 1.91	*FSTL4*
chr13:17120000_17140000	20	*xp-ehh*	low	244 - 244	3 - 3	418 - 418	0.09 - 0.09	0.86 - 0.86	*FSTL4*
chr14:1620000_1640000	20	*xp-ehh*	low	120 - 120	3.01 - 3.01	257 - 257	0.24 - 0.24	4.12 - 4.12	*NPTX2*
chr14:3930000_3980000	50	*fst*	low	79 - 163	0.81 - 1.19	151 - 235	0.29 - 0.43	5.25 - 8.16	*SDK1*
chr14:4880000_4910000	30	*xp-ehh*	low	190 - 200	3 - 3.4	344 - 355	0.17 - 0.21	2.73 - 3.43	*MAP2K3*
chr14:5120000_5180000	60	*xp-ehh*	low	168 - 251	3.2 - 3.56	322 - 372	0.19 - 0.21	3.09 - 3.42	*COPS3, FLCN, MED9, NT5M, PLD6*
chr14:7440000_7460000	20	*fst*	low	143 - 143	2.48 - 2.48	290 - 290	0.31 - 0.31	5.5 - 5.5	
chr17:3850000_3870000	20	*xp-ehh*	low	145 - 145	3.03 - 3.03	332 - 332	0.13 - 0.13	1.85 - 1.85	*ASTN2, ENSGALG00000025230 (MiRNA)*
chr18:3690000_3770000	80	*xp-ehh*	low	168 - 237	3.01 - 3.21	271 - 451	0.12 - 0.17	1.52 - 2.68	*CD7, ENSGALG00000054039 (LncRNA), SEPT9*
chr18:5090000_5190000	100	*xp-ehh* and/or *fst*	low	95 - 227	2.29 - 4.4	157 - 415	0.27 - 0.36	4.69 - 6.64	*CA10, MBTD1, UTP18*
chr18:5680000_5710000	30	*xp-ehh*	low	169 - 176	3.25 - 3.34	319 - 352	0.22 - 0.28	3.75 - 4.98	
chr18:6570000_6590000	20	*xp-ehh*	low	155 - 155	3.11 - 3.11	292 - 292	0.26 - 0.26	4.64 - 4.64	*UTP6*
chr18:10410000_10510000	100	*xp-ehh*	low	52 - 179	3.05 - 3.67	127 - 345	0.09 - 0.18	1.02 - 2.95	*CD300LG, ENSGALG00000019755, GRIN2C, NAT9, RAB37, SLC9A3R1, TMEM104, XYLT2*
chr18:10540000_10570000	30	*xp-ehh*	low	65 - 75	3.08 - 3.15	143 - 167	0.16 - 0.18	2.43 - 2.79	*HID1, OTOP2, OTOP3, USH1G*
chr19:7450000_7470000	20	*xp-ehh*	low	203 - 203	3.01 - 3.01	322 - 322	0.15 - 0.15	2.26 - 2.26	*CLTC, PTRH2, VMP1*
chr20:4510000_4550000	40	*xp-ehh*	low	89 - 230	3.23 - 3.31	241 - 384	0.16 - 0.19	2.42 - 3	*ENSGALG00000046818 (LncRNA)*
chr20:4570000_4600000	30	*fst*	low	110 - 139	2.46 - 2.77	217 - 333	0.31 - 0.34	5.59 - 6.32	*ENSGALG00000052811 (LncRNA)*
chr21:1990000_2010000	20	*fst*	low	154 - 154	0.24 - 0.24	218 - 218	0.28 - 0.28	5.06 - 5.06	*ENSGALG00000021676, NADK, SLC35E2B*
chr21:3340000_3380000	40	*xp-ehh*	low	101 - 121	3.12 - 3.27	285 - 302	0.16 - 0.18	2.44 - 2.88	*SPSB1*

chr21:5230000_5250000	20	*xp-ehh*	low	203 - 203	3.22 - 3.22	339 - 339	0.19 - 0.19	3.04 - 3.04	*LRRC38*
chr23:440000_460000	20	*fst*	low	169 - 169	0.96 - 0.96	303 - 303	0.3 - 0.3	5.39 - 5.39	
chr23:2180000_2200000	20	*fst*	low	99 - 99	2.29 - 2.29	251 - 251	0.29 - 0.29	5.2 - 5.2	*AHDC1*
chr23:2980000_3010000	30	*xp-ehh*	low	116 - 140	3.25 - 3.44	251 - 272	0.17 - 0.24	2.58 - 4.13	*ENSGALG00000049526 (LncRNA), PTPRU*
chr26:4160000_4180000	20	*xp-ehh*	low	97 - 97	3.08 - 3.08	207 - 207	0.24 - 0.24	4.17 - 4.17	*TCP11*
chr27:5690000_5730000	40	*xp-ehh*	low	162 - 206	3.08 - 3.79	322 - 365	0.19 - 0.23	3 - 3.99	*FAM117A, KAT7*

Table S4.2b. Candidate selection signature regions and overlapping genes detected from analyses of the precipitation seasonality (precSeasonality). (**SSR_ID:** selection signature regions ID as chromosome:start_stop position. **DD:** directionality of selection)

SSR_ID	**size (kb)**	**detection method**	**DD**	**no. of SNPs per *XP-EHH* window (range)**	**absolute value of mean *XP-EHH*_std (range)**	**no. of SNPs per *Fst* window (range)**	**weighted *Fst* per window (range)**	**Z*Fst* per window (range)**	**Overlapping genes**
chr1:450000_490000	40	*xp-ehh*	low	108 -144	3.06 - 3.29	461 - 526	0.06 - 0.11	0.76 - 2.18	*ENSGALG00000033919*
chr1:12910000_12940000	30	*fst*	low	177 - 191	2.76 - 2.77	294 - 366	0.22 - 0.27	5.18 - 6.61	*ENSGALG00000008309, FAM185A, FGL2*
chr1:13950000_13970000	20	*fst*	high	152 - 152	0.56 - 0.56	225 - 225	0.21 - 0.21	5.03 - 5.03	*KMT2E, SRPK2*
chr1:16040000_16060000	20	*xp-ehh*	low	157 - 157	3.04 - 3.04	326 - 326	0.07 - 0.07	1.26 - 1.26	*ENSGALG00000054287*
chr1:19670000_19690000	20	*fst*	high	156 - 156	0.28 - 0.28	191 - 191	0.23 - 0.23	5.45 - 5.45	
chr1:24350000_24370000	20	*fst*	high	185 - 185	1.5 - 1.5	274 - 274	0.22 - 0.22	5.29 - 5.29	
chr1:27410000_27490000	80	*xp-ehh*	high	113 - 253	3 - 3.55	197 - 412	0.02 - 0.1	-0.19 - 1.84	*IMMP2L*
chr1:27560000_27660000	100	*xp-ehh*	high	115 - 185	3.31 - 4	177 - 327	0.06 - 0.14	0.83 - 3.15	*IMMP2L, LRRN3*
chr1:52400000_52450000	50	*fst*	low	237 - 290	2.04 - 2.35	383 - 450	0.23 - 0.3	5.46 - 7.39	*ENSGALG00000036131 (LncRNA)*
chr1:81890000_81930000	40	*xp-ehh*	low	152 - 196	3.12 - 3.66	306 - 352	0.12 - 0.16	2.62 - 3.63	
chr1:105230000_105250000	20	*xp-ehh*	high	144 - 144	3.01 - 3.01	401 - 401	0.06 - 0.06	0.9 - 0.9	*ENSGALG00000027781 (ncRNA)*
chr1:134820000_134840000	20	*xp-ehh*	low	202 - 202	3.16 - 3.16	385 - 385	0.13 - 0.13	2.8 - 2.8	*RFX8*
chr1:137260000_137280000	20	*xp-ehh*	high	211 - 211	3.07 - 3.07	368 - 368	0.02 - 0.02	-0.18 - -0.18	*ENSGALG00000049876 (LncRNA), ENSGALG00000052777 (LncRNA)*

chr1:147500000_147530000	30	*xp-ehh*	low	251 - 281	3.24 - 3.44	411 - 440	0.1 - 0.11	2.03 - 2.28	*CLDN10*
chr1:157150000_157170000	20	*xp-ehh*	high	179 - 179	3.06 - 3.06	342 - 342	0.1 - 0.1	1.96 - 1.96	*KLF12*
chr1:164020000_164060000	40	*fst*	high	80 - 103	1.03 - 1.32	153 - 178	0.26 - 0.28	6.46 - 6.84	*ENSGALG00000050284 (LncRNA)*
chr1:167460000_167510000	50	*xp-ehh*	high	225 - 275	3.14 - 3.89	392 - 470	0.03 - 0.05	-0.07 - 0.61	*ENSGALG00000017789 (5S_rRNA), VWA8*
chr1:178310000_178340000	30	*xp-ehh*	low	171 - 207	3.06 - 3.23	296 - 318	0.07 - 0.11	1.2 - 2.13	*ATP8A2*
chr1:185880000_185920000	40	*xp-ehh*	high	149 - 196	3.16 - 3.33	254 - 331	0.01 - 0.03	-0.64 - -0.07	*ENSGALG00000047983 (LncRNA), ENSGALG00000054022 (LncRNA)*
chr2:9330000_9360000	30	*xp-ehh*	high	164 - 195	3.16 - 3.3	255 - 271	0.05 - 0.05	0.46 - 0.6	*PTPRN2*
chr2:9730000_9750000	20	*xp-ehh*	high	223 - 223	3.02 - 3.02	345 - 345	0.1 - 0.1	1.92 - 1.92	*VIPR2*
chr2:28140000_28200000	60	*fst*	high	43 - 83	0.72 - 0.84	122 - 200	0.23 - 0.25	5.65 - 6.04	*ENSGALG00000054958 (LncRNA)*
chr2:54120000_54140000	20	*xp-ehh*	low	126 - 126	3.02 - 3.02	186 - 186	0.04 - 0.04	0.24 - 0.24	*CNTNAP2*
chr2:55230000_55300000	70	*xp-ehh* and/or *fst*	low	141 - 238	2.47 - 3.32	246 - 363	0.13 - 0.32	2.78 - 7.87	*ENSGALG00000054801 (LncRNA), TNS3*
chr2:56080000_56100000	20	*xp-ehh*	high	205 - 205	3.08 - 3.08	302 - 302	0.1 - 0.1	1.94 - 1.94	*RBFA*
chr2:63890000_63930000	40	*fst*	high	94 - 195	2.26 - 2.62	239 - 301	0.23 - 0.24	5.48 - 5.8	*ENSGALG00000049311 (LncRNA)*
chr2:68700000_68740000	40	*xp-ehh*	high	122 - 245	3 - 3.24	225 - 368	0.05 - 0.11	0.62 - 2.15	*ENSGALG00000053182 (LncRNA)*
chr2:69270000_69290000	20	*xp-ehh*	high	87 - 87	3.05 - 3.05	174 - 174	0.14 - 0.14	3.17 - 3.17	*CDH6*
chr2:92020000_92050000	30	*xp-ehh*	low	220 - 282	3.2 - 3.21	348 - 403	0.1 - 0.15	2.07 - 3.33	
chr2:94060000_94080000	20	*fst*	high	67 - 67	0.04 - 0.04	119 - 119	0.22 - 0.22	5.25 - 5.25	*CCDC102B*
chr2:94690000_94740000	50	*fst*	high	49 - 124	0.81 - 1.2	117 - 214	0.21 - 0.23	5.06 - 5.45	
chr2:95600000_95660000	60	*fst*	low	63 - 91	1.19 - 2.1	111 - 147	0.21 - 0.28	5.09 - 6.97	
chr2:95750000_95770000	20	*fst*	low	21 - 21	1.31 - 1.31	45 - 45	0.36 - 0.36	9.05 - 9.05	
chr2:95830000_95850000	20	*fst*	low	86 - 86	0.16 - 0.16	150 - 150	0.22 - 0.22	5.32 - 5.32	
chr2:95860000_95890000	30	*fst*	low	110 - 111	0.58 - 0.6	163 - 163	0.23 - 0.24	5.46 - 5.71	*ENSGALG00000050654 (LncRNA)*
chr2:95910000_95930000	20	*fst*	low	96 - 96	0.97 - 0.97	163 - 163	0.28 - 0.28	6.9 - 6.9	*ENSGALG00000050654 (LncRNA)*
chr2:96920000_96950000	30	*xp-ehh*	low	280 - 331	3.11 - 3.29	379 - 418	0.08 - 0.09	1.3 - 1.65	*CEP192*
chr2:102480000_102500000	20	*xp-ehh*	low	310 - 310	3 - 3	501 - 501	0.04 - 0.04	0.19 - 0.19	*MIB1*
chr2:103460000_103550000	90	*xp-ehh*	low	91 - 241	3.02 - 3.73	205 - 374	0.08 - 0.16	1.33 - 3.66	*ENSGALG00000049694 (LncRNA)*
chr2:104260000_104320000	60	*xp-ehh*	low	190 - 245	3.37 - 3.81	304 - 394	0.06 - 0.09	0.86 - 1.68	*KCTD1*
chr2:114000000_114100000	100	*xp-ehh* and/or *fst*	low	87 - 152	3.02 - 3.65	170 - 239	0.12 - 0.23	2.41 - 5.49	*ENSGALG00000053888 (LncRNA)*
chr2:146340000_146390000	50	*xp-ehh*	low	169 - 294	3.25 - 3.53	301 - 371	0.07 - 0.14	1.24 - 3.07	
chr2:146690000_146710000	20	*fst*	low	121 - 121	0.31 - 0.31	195 - 195	0.21 - 0.21	5.07 - 5.07	

chr2:146890000_146920000	30	*fst*	low	102 - 122	0.54 - 0.63	169 - 213	0.28 - 0.29	6.95 - 7.05	*TSNARE1*
chr2:148130000_148170000	40	*fst*	high	208 - 234	2.23 - 2.32	337 - 360	0.24 - 0.26	5.82 - 6.39	*ENSGALG00000028660 (MiRNA),* *ENSGALG00000047330 (LncRNA)*
chr2:148190000_148210000	20	*xp-ehh*	high	245 - 245	3.17 - 3.17	439 - 439	0.16 - 0.16	3.68 - 3.68	*ENSGALG00000048947 (LncRNA)*
chr2:148240000_148260000	20	*fst*	high	184 - 184	1.41 - 1.41	399 - 399	0.22 - 0.22	5.15 - 5.15	*ENSGALG00000048947 (LncRNA)*
chr3:1160000_1210000	50	*fst*	low	7 - 42	1.85 - 2.58	63 - 92	0.25 - 0.38	5.93 - 9.63	*ENSGALG00000047366 (LncRNA)*
chr3:2860000_2880000	20	*fst*	low	77 - 77	0.68 - 0.68	141 - 141	0.24 - 0.24	5.84 - 5.84	
chr3:5540000_5570000	30	*xp-ehh*	low	157 - 165	3.23 - 3.41	266 - 315	0.15 - 0.17	3.37 - 3.84	*KIF16B*
chr3:18250000_18310000	60	*fst*	low	120 - 197	0.16 - 0.43	223 - 284	0.21 - 0.24	5.01 - 5.79	*ENSGALG00000052742 (LncRNA)*
chr3:20080000_20110000	30	*fst*	high	80 - 156	1.46 - 1.97	152 - 226	0.23 - 0.29	5.55 - 7.06	*ESRRG*
chr3:20220000_20240000	20	*fst*	high	115 - 115	0.18 - 0.18	179 - 179	0.21 - 0.21	5.06 - 5.06	*ESRRG*
chr3:21940000_21970000	30	*fst*	high	101 - 102	1.1 - 1.51	186 - 298	0.33 - 0.43	8.19 - 10.95	*ATF3, gga-mir-1649, TMEM206*
chr3:32600000_32620000	20	*fst*	low	162 - 162	0.56 - 0.56	249 - 249	0.22 - 0.22	5.12 - 5.12	*ENSGALG00000035137 (LncRNA)*
chr3:50820000_50860000	40	*fst*	high	75 - 82	0.75 - 2.34	229 - 266	0.24 - 0.29	5.83 - 7.09	*ENSGALG00000047502 (LncRNA)*
chr3:50890000_50920000	30	*fst*	high	45 - 45	1.43 - 1.87	237 - 238	0.22 - 0.24	5.25 - 5.68	*ENSGALG00000049389 (LncRNA)*
chr3:70960000_70980000	20	*xp-ehh*	high	140 - 140	3.07 - 3.07	272 - 272	0.09 - 0.09	1.74 - 1.74	*SIM1*
chr3:71640000_71660000	20	*fst*	high	141 - 141	2.14 - 2.14	274 - 274	0.22 - 0.22	5.23 - 5.23	*ENSGALG00000047943 (LncRNA),* *ENSGALG00000052916 (LncRNA)*
chr3:76730000_76760000	30	*fst*	high	215 - 226	1.99 - 2.98	305 - 346	0.26 - 0.27	6.38 - 6.66	
chr3:78860000_78890000	30	*xp-ehh*	high	218 - 253	3.02 - 3.05	280 - 343	0.08 - 0.1	1.52 - 1.91	
chr3:80440000_80480000	40	*xp-ehh*	high	159 - 188	3.13 - 3.37	317 - 343	0.14 - 0.17	2.93 - 3.78	*IMPG1, MYO6*
chr3:85880000_85900000	20	*fst*	low	81 - 81	2.06 - 2.06	171 - 171	0.23 - 0.23	5.63 - 5.63	*ENSGALG00000047770 (miRNA)*
chr3:85990000_86100000	110	*fst*	low	42 - 127	0.85 - 2.15	128 - 272	0.22 - 0.3	5.38 - 7.41	*KHDRBS2*
chr3:86110000_86170000	60	*fst*	low	34 - 127	0.17 - 0.37	99 - 230	0.22 - 0.28	5.32 - 6.93	*KHDRBS2*
chr3:86680000_86770000	90	*fst*	low	214 - 278	0.83 - 1.86	376 - 479	0.22 - 0.32	5.34 - 7.94	*BAG2, PRIM2, RAB23*
chr3:96570000_96590000	20	*xp-ehh*	high	247 - 247	3.12 - 3.12	387 - 387	0.07 - 0.07	1.17 - 1.17	
chr3:99080000_99100000	20	*xp-ehh*	high	141 - 141	3.25 - 3.25	203 - 203	0.11 - 0.11	2.2 - 2.2	
chr3:100050000_100070000	20	*xp-ehh*	high	170 - 170	3.11 - 3.11	246 - 246	0.1 - 0.1	2.08 - 2.08	*VSNL1*
chr3:100240000_100270000	30	*xp-ehh*	high	175 - 195	3.15 - 3.18	300 - 309	0.07 - 0.08	1.02 - 1.47	*KCNS3*
chr3:100370000_100390000	20	*xp-ehh*	high	156 - 156	3.07 - 3.07	311 - 311	0.13 - 0.13	2.78 - 2.78	
chr3:108260000_108280000	20	*xp-ehh*	high	190 - 190	3.08 - 3.08	347 - 347	0.06 - 0.06	0.95 - 0.95	*TFAP2D*
chr4:650000_670000	20	*fst*	high	147 - 147	1.71 - 1.71	211 - 211	0.24 - 0.24	5.84 - 5.84	
chr4:2730000_2750000	20	*xp-ehh* *fst*	high	250 - 250	3.19 - 3.19	455 - 455	0.21 - 0.21	5.09 - 5.09	*HTR2C*

chr4:11090000_11110000	20	*xp-ehh*	high	217 - 217	3.08 - 3.08	422 - 422	0.05 - 0.05	0.66 - 0.66	
chr4:26940000_27010000	70	*xp-ehh*	low	64 - 101	3.05 - 3.12	160 - 207	0.06 - 0.07	0.74 - 1.07	
chr4:28270000_28360000	90	*xp-ehh*	low	21 - 77	3.02 - 3.39	59 - 141	0.06 - 0.15	0.78 - 3.23	*ENSGALG00000047082 (LncRNA)*
chr4:28380000_28450000	70	*xp-ehh*	low	22 - 58	3.08 - 3.23	122 - 176	0.06 - 0.08	0.8 - 1.48	
chr4:30270000_30290000	20	*fst*	high	226 - 226	2.48 - 2.48	290 - 290	0.24 - 0.24	5.79 - 5.79	*INPP4B*
chr4:41990000_42020000	30	*fst*	high	56 - 59	0.85 - 0.9	67 - 92	0.22 - 0.26	5.23 - 6.25	*ENSGALG00000047221 (LncRNA), ENSGALG00000049124*
chr4:60250000_60290000	40	*xp-ehh* and/or *fst*	low	157 - 201	3.47 - 3.75	229 - 294	0.14 - 0.22	3.03 - 5.14	*ENSGALG00000049121 (LncRNA)*
chr4:60950000_60980000	30	*xp-ehh*	low	256 - 285	3 - 3.16	384 - 420	0.12 - 0.12	2.47 - 2.5	*BDH2, CISD2, SLC9B2*
chr4:61240000_61270000	30	*xp-ehh*	low	235 - 284	3.12 - 3.28	379 - 423	0.08 - 0.08	1.33 - 1.37	
chr4:74180000_74240000	60	*fst*	high	106 - 145	1.25 - 2.66	164 - 234	0.29 - 0.3	7.05 - 7.53	*ENSGALG00000046053 (LncRNA), ENSGALG00000048521 (LncRNA), ENSGALG00000050078 (LncRNA)*
chr4:74250000_74360000	110	*xp-ehh* and/or *fst*	high	67 - 125	1.88 - 3.2	137 - 229	0.08 - 0.28	1.5 - 6.81	*ENSGALG00000046053 (LncRNA), ENSGALG00000051573 (LncRNA)*
chr4:74400000_74510000	110	*fst*	high	107 - 207	2.98 - 4.13	180 - 342	0.14 - 0.24	3.06 - 5.8	*ADGRA3, ENSGALG00000047968, ENSGALG00000048082*
chr4:74710000_74740000	30	*fst*	high	178 - 252	1.71 - 2.19	256 - 305	0.21 - 0.22	5.1 - 5.12	
chr4:75070000_75090000	20	*xp-ehh*	high	118 - 118	3.01 - 3.01	284 - 284	0.09 - 0.09	1.55 - 1.55	*SLIT2*
chr4:75150000_75170000	20	*xp-ehh*	high	126 - 126	3.01 - 3.01	248 - 248	0.08 - 0.08	1.43 - 1.43	*SLIT2*
chr4:75660000_75780000	120	*xp-ehh*	high	114 - 174	3.05 - 3.85	155 - 280	0.02 - 0.17	-0.14 - 3.77	*ENSGALG00000047307 (LncRNA), ENSGALG00000052764 (LncRNA), NCAPG*
chr4:77500000_77530000	30	*xp-ehh*	high	85 - 123	3.03 - 3.12	194 - 228	0.04 - 0.1	0.28 - 1.99	*ENSGALG00000052609 (MiRNA), RAB28*
chr4:81570000_81590000	20	*xp-ehh*	high	153 - 153	3.17 - 3.17	309 - 309	0.12 - 0.12	2.45 - 2.45	*ADRA2C*
chr4:89960000_89980000	20	*xp-ehh*	high	235 - 235	3.01 - 3.01	380 - 380	0.14 - 0.14	3.13 - 3.13	*EXOC6B*
chr4:90310000_90340000	30	*xp-ehh*	high	143 - 171	3.02 - 3.14	263 - 289	0.05 - 0.07	0.5 - 1.01	
chr5:2310000_2330000	20	*fst*	high	45 - 45	1.7 - 1.7	109 - 109	0.25 - 0.25	6.03 - 6.03	*PRMT3*
chr5:12750000_12790000	40	*xp-ehh*	low	289 - 306	3.15 - 3.28	422 - 452	0.05 - 0.1	0.68 - 1.93	*LDHA, TSG101, UEVLD*
chr5:12800000_12820000	20	*xp-ehh*	low	223 - 223	3.02 - 3.02	393 - 393	0.16 - 0.16	3.51 - 3.51	*SPTY2D1*
chr5:21780000_21800000	20	*fst*	high	142 - 142	2.24 - 2.24	282 - 282	0.21 - 0.21	5.09 - 5.09	
chr5:28910000_28970000	60	*xp-ehh*	low	167 - 340	3.04 - 3.34	329 - 490	0.1 - 0.15	2.06 - 3.37	*ENSGALG00000051208 (LncRNA), RAD51B, TMEM229B*

chr5:35390000_35410000	20	*fst*	high	103 - 103	1.2 - 1.2	158 - 158	0.21 - 0.21	5.03 - 5.03	*NPAS3, ENSGALG00000025613 (ncRNA)*
chr5:40460000_40570000	110	*fst*	high	5 - 103	0.34 - 2.01	133 - 203	0.22 - 0.36	5.16 - 9.04	*NRXN3*
chr5:43730000_43750000	20	*xp-ehh*	low	251 - 251	3.44 - 3.44	424 - 424	0.11 - 0.11	2.15 - 2.15	*FOXN3*
chr5:48710000_49000000	290	*xp-ehh*	high	25 - 205	3.11 - 3.89	153 - 397	0.03 - 0.13	-0.03 - 2.87	*BEGAIN, ENSGALG00000050163 (miRNA), ENSGALG00000052029 (LncRNA), SLC25A29, SLC25A47, WARS, DR25, YY1*
chr5:49070000_49190000	120	*xp-ehh*	high	30 - 218	3.02 - 3.58	92 - 281	0.04 - 0.09	0.21 - 1.79	*DLK1, ENSGALG00000047452 (LncRNA), ENSGALG00000047870 (LncRNA), ENSGALG00000053427 (LncRNA)*
chr5:49240000_49300000	60	*xp-ehh*	high	53 - 133	3.03 - 3.25	176 - 347	0.08 - 0.08	1.3 - 1.52	*ENSGALG00000050843 (LncRNA)*
chr5:58640000_58710000	70	*xp-ehh*	high	94 - 168	3.06 - 3.51	187 - 296	0.06 - 0.09	0.85 - 1.58	*ENSGALG00000050220 (LncRNA)*
chr6:1620000_1650000	30	*xp-ehh*	low	239 - 244	3.24 - 3.4	370 - 402	0.05 - 0.07	0.57 - 1.1	*CCSER2, ENSGALG00000031866*
chr6:12470000_12500000	30	*xp-ehh*	low	184 - 219	3.04 - 3.05	313 - 386	0.14 - 0.16	3.08 - 3.48	*ENSGALG00000053115 (LncRNA), UNC5B*
chr6:12510000_12620000	110	*xp-ehh*	low	85 - 258	3.16 - 3.28	193 - 396	0.07 - 0.2	1.16 - 4.58	*CDH23, SLC29A3*
chr6:19430000_19450000	20	*xp-ehh*	low	165 - 165	3.28 - 3.28	273 - 273	0.01 - 0.01	-0.54 - -0.54	*ENSGALG00000047645 (LncRNA), ENSGALG00000049384 (LncRNA), ENSGALG00000054411 (LncRNA), ENSGALG00000054709 (LncRNA)*
chr6:25040000_25070000	30	*xp-ehh*	high	167 - 230	3.04 - 3.09	321 - 427	0.12 - 0.17	2.43 - 3.85	*SLK*
chr7:11540000_11570000	30	*xp-ehh*	high	294 - 307	3.31 - 3.41	403 - 408	0.09 - 0.09	1.64 - 1.8	*ALS2*
chr7:25930000_25950000	20	*xp-ehh*	low	170 - 170	3.13 - 3.13	308 - 308	0.13 - 0.13	2.78 - 2.78	*GLI2*
chr7:33100000_33180000	80	*xp-ehh*	low	113 - 204	3.11 - 4.65	230 - 286	0.03 - 0.14	-0.05 - 3.12	*GTDC1*
chr7:33200000_33240000	40	*xp-ehh*	low	154 - 187	3.39 - 4.05	239 - 329	0.07 - 0.14	1.25 - 2.93	*GTDC1*
chr7:33280000_33310000	30	*xp-ehh*	low	148 - 163	3.19 - 3.6	323 - 335	0.11 - 0.13	2.26 - 2.79	*ENSGALG00000053032 (LncRNA), ZEB2*
chr7:33820000_33840000	20	*fst*	high	193 - 193	0.74 - 0.74	274 - 274	0.22 - 0.22	5.19 - 5.19	*ENSGALG00000041782 (LncRNA), ENSGALG00000047423 (LncRNA)*
chr7:35150000_35170000	20	*xp-ehh*	low	198 - 198	3.12 - 3.12	430 - 430	0.13 - 0.13	2.64 - 2.64	*RIF1*
chr7:35360000_35400000	40	*xp-ehh* and/or *fst*	low	237 - 240	3.07 - 3.69	434 - 467	0.2 - 0.23	4.83 - 5.64	*ENSGALG00000035466*
chr7:36100000_36120000	20	*fst*	high	167 - 167	0.48 - 0.48	337 - 337	0.23 - 0.23	5.39 - 5.39	*GPD2*

chr8:12720000_12740000	20	*fst*	low	72 - 72	0.57 - 0.57	91 - 91	0.21 - 0.21	5.03 - 5.03	*PLPPR5*
chr8:20440000_20470000	30	*xp-ehh*	high	151 - 183	3.01 - 3.01	257 - 307	0.08 - 0.08	1.49 - 1.53	*PTPRF*
chr9:11400000_11430000	30	*fst*	low	261 - 310	1.97 - 2.16	338 - 395	0.21 - 0.22	5.05 - 5.31	*ENSGALG00000052830 (LncRNA)*
chr9:11440000_11470000	30	*xp-ehh*	low	168 - 168	3.01 - 3.04	256 - 256	0.09 - 0.1	1.72 - 2.08	*ENSGALG00000052830 (LncRNA)*
chr9:11600000_11620000	20	*fst*	low	118 - 118	1.24 - 1.24	145 - 145	0.22 - 0.22	5.11 - 5.11	
chr9:12140000_12160000	20	*xp-ehh*	low	177 - 177	3.08 - 3.08	254 - 254	0.04 - 0.04	0.41 - 0.41	*ENSGALG00000049451 (LncRNA)*
chr9:12510000_12570000	60	*fst*	high	81 - 161	0.33 - 0.92	163 - 261	0.22 - 0.32	5.16 - 7.9	*AGTR1*
chr11:1090000_1150000	60	*xp-ehh*	high	152 - 249	3.14 - 4.14	280 - 423	0.03 - 0.11	0.07 - 2.1	*CARMIL2, CTCFL, RIPOR1*
chr11:1200000_1220000	20	*xp-ehh*	high	237 - 237	3.13 - 3.13	418 - 418	0.1 - 0.1	1.82 - 1.82	*ENSGALG00000002091, ENSGALG00000035945*
chr11:1660000_1680000	20	*fst*	high	128 - 128	2.35 - 2.35	288 - 288	0.23 - 0.23	5.38 - 5.38	*VAC14*
chr11:11940000_11960000	20	*xp-ehh*	high	216 - 216	3.13 - 3.13	341 - 341	0.03 - 0.03	0.04 - 0.04	*ENSGALG00000053476 (LncRNA)*
chr11:14210000_14260000	50	*xp-ehh*	high	152 - 181	3.04 - 3.52	282 - 312	0.06 - 0.09	0.77 - 1.62	*WWOX*
chr12:12730000_12750000	20	*xp-ehh*	high	172 - 172	3.07 - 3.07	272 - 272	0.08 - 0.08	1.41 - 1.41	*FHIT*
chr12:18090000_18120000	30	*xp-ehh*	high	134 - 159	3.03 - 3.17	241 - 317	0.08 - 0.1	1.39 - 2	*ENSGALG00000046751 (LncRNA), ENSGALG00000050333 (LncRNA)*
chr13:5180000_5210000	30	*xp-ehh*	low	210 - 234	3.2 - 3.36	325 - 336	0.07 - 0.1	1.26 - 1.89	*SLIT3*
chr13:5740000_5790000	50	*xp-ehh*	low	145 - 228	3.03 - 3.11	285 - 407	0.09 - 0.13	1.74 - 2.88	*TENM2*
chr13:7310000_7340000	30	*fst*	high	71 - 83	1.63 - 1.8	286 - 323	0.21 - 0.23	5.08 - 5.55	*ENSGALG00000054885 (LncRNA)*
chr14:4850000_4870000	20	*xp-ehh*	low	214 - 214	3.39 - 3.39	349 - 349	0.07 - 0.07	1.16 - 1.16	*ENSGALG00000031051 (LncRNA)*
chr15:9090000_9120000	30	*xp-ehh*	low	165 - 201	3.11 - 3.12	228 - 255	0.08 - 0.1	1.42 - 1.88	*ENSGALG00000030684 (LncRNA), YWHAH*
chr17:9480000_9510000	30	*fst*	high	68 - 79	0.2 - 0.24	81 - 91	0.24 - 0.27	5.7 - 6.65	*DENND1A, ENSGALG00000026417 (ncRNA)*
chr18:9880000_9920000	40	*xp-ehh*	low	119 - 172	3.17 - 3.6	206 - 337	0.08 - 0.09	1.38 - 1.65	*GCGR, HGS, MCRIP1, MRPL12, 4HB, PPP1R27, SLC25A10*
chr19:1750000_1790000	40	*xp-ehh*	high	139 - 179	3.04 - 3.29	281 - 301	0.07 - 0.13	1.17 - 2.8	*AUTS2*
chr20:1090000_1110000	20	*fst*	low	224 - 224	1.14 - 1.14	297 - 297	0.22 - 0.22	5.27 - 5.27	*CPNE1, RBM12*
chr20:6560000_6580000	20	*fst*	low	152 - 152	1.5 - 1.5	230 - 230	0.21 - 0.21	5.05 - 5.05	*KCNB1, TGM3*
chr22:2610000_2630000	20	*fst*	low	8 - 8	1.98 - 1.98	36 - 36	0.22 - 0.22	5.27 - 5.27	*NSD3*
chr23:4350000_4400000	50	*xp-ehh*	high	113 - 169	3.27 - 3.75	219 - 340	0.11 - 0.19	2.16 - 4.48	*AGO1, AGO4, C1orf216, CLSPN*
chr24:1140000_1180000	40	*fst*	low	281 - 361	0.7 - 0.88	371 - 471	0.22 - 0.26	5.12 - 6.4	*KCNJ5*
chr24:1210000_1240000	30	*fst*	low	339 - 366	0.44 - 0.46	483 - 497	0.24 - 0.24	5.75 - 5.83	

chr25:240000_260000	20	*xp-ehh*	high	113 - 113	3.26 - 3.26	268 - 268	0.08 - 0.08	1.46 - 1.46	*GABPB2, SCNM1, SEMA6C, TMOD4, VPS72*
chr27:3760000_3780000	20	*xp-ehh*	low	275 - 275	3.09 - 3.09	419 - 419	0.06 - 0.06	0.77 - 0.77	*ENSGALG00000000824, ENSGALG00000035554 (LncRNA), FMNL1*
chr28:4540000_4580000	40	*xp-ehh*	low	131 - 189	3.05 - 3.17	150 - 224	0.12 - 0.17	2.43 - 3.89	*ENSGALG00000050334 (LncRNA), TPM4*
chr31:920000_950000	30	*fst*	low	7 - 8	1.21 - 1.27	12 - 15	0.24 - 0.26	5.77 - 6.38	*ENSGALG00000044223, ENSGALG00000044521, ENSGALG00000045132, ENSGALG00000050726 (LncRNA)*

Table S4.2c. Candidate selection signature regions and overlapping genes detected from analyses of the precipitation in the wettest quarter of a year (precWQ). (**SSR_ID:** selection signature regions ID as chromosome:start_stop position. **DD:** directionality of selection)

SSR_ID	**size (kb)**	**detection method**	**DD**	**no. of SNPs per *XP-EHH* window (range)**	**absolute value of mean *XP-EHH*_std (range)**	**no. of SNPs per *Fst* window (range)**	**weighted *Fst* per window (range)**	**Z*Fst* per window (range)**	**Overlapping genes**
chr1:20430000_20450000	20	*fst*	low	108 - 108	0.51 - 0.51	344 - 344	0.4 - 0.4	5.08 - 5.08	*SELENOO*
chr1:38010000_38050000	40	*xp-ehh*	low	225 - 263	3.01 - 3.77	421 - 540	0.17 - 0.22	1.46 - 2.34	*NAP1L1, PHLDA1*
chr1:45020000_45040000	20	*xp-ehh*	low	129 - 129	3 - 3	270 - 270	0.2 - 0.2	1.92 - 1.92	
chr1:49330000_49400000	70	*xp-ehh*	high	148 - 180	3.04 - 3.36	248 - 360	0.24 - 0.29	2.57 - 3.35	*TCF20*
chr1:55780000_55800000	20	*xp-ehh*	high	93 - 93	3.05 - 3.05	217 - 217	0.18 - 0.18	1.68 - 1.68	*TBXAS1*
chr1:83090000_83110000	20	*fst*	low	96 - 96	2 - 2	235 - 235	0.42 - 0.42	5.43 - 5.43	*ENSGALG00000054748 (LncRNA)*
chr1:99010000_99040000	30	*xp-ehh*	high	48 - 51	3.24 - 3.28	112 - 176	0.28 - 0.34	3.13 - 4.11	
chr1:99050000_99150000	100	*xp-ehh*	high	18 - 135	3.04 - 3.25	77 - 202	0.12 - 0.32	0.72 - 3.89	*ENSGALG00000052669 (LncRNA)*
chr1:99190000_99220000	30	*xp-ehh*	high	34 - 54	3.01 - 3.01	185 - 201	0.17 - 0.19	1.44 - 1.83	
chr1:110570000_110590000	20	*fst*	high	31 - 31	0.4 - 0.4	53 - 53	0.4 - 0.4	5.13 - 5.13	*ENSGALG00000048672*
chr1:133250000_133270000	20	*xp-ehh*	low	276 - 276	3.1 - 3.1	482 - 482	0.09 - 0.09	0.25 - 0.25	*CNGA3*
chr1:150050000_150100000	50	*fst*	low	26 - 42	1.12 - 1.33	88 - 104	0.4 - 0.44	5.02 - 5.61	*ENSGALG00000052317*

chr1:150130000_150150000	20	*fst*	low	13 - 13	0.69 - 0.69	123 - 123	0.41 - 0.41	5.21 - 5.21	
chr1:150460000_150490000	30	*fst*	low	96 - 112	1.17 - 1.19	139 - 158	0.47 - 0.5	6.22 - 6.56	
chr1:177010000_177050000	40	*fst*	low	160 - 187	0.95 - 1.15	294 - 324	0.42 - 0.47	5.44 - 6.15	
chr1:186050000_186070000	20	*fst*	high	86 - 86	0.84 - 0.84	232 - 232	0.4 - 0.4	5.01 - 5.01	
chr2:8150000_8170000	20	*xp-ehh*	high	151 - 151	3.09 - 3.09	342 - 342	0.22 - 0.22	2.31 - 2.31	*SHH*
chr2:21970000_21990000	20	*fst*	high	163 - 163	0.53 - 0.53	325 - 325	0.4 - 0.4	5.06 - 5.06	*CDK14*
chr2:37120000_37140000	20	*xp-ehh*	low	172 - 172	3.37 - 3.37	369 - 369	0.19 - 0.19	1.78 - 1.78	*NKIRAS1, RPL15, UBE2E1*
chr2:41870000_41900000	30	*xp-ehh*	high	206 - 241	3.04 - 3.05	474 - 513	0.15 - 0.16	1.18 - 1.35	*CPNE4*
chr2:42570000_42630000	60	*xp-ehh*	high	194 - 234	3.03 - 3.19	329 - 371	0.11 - 0.15	0.59 - 1.2	*BFSP2, CCR5, CDV3, TOPBP1*
chr2:57870000_57910000	40	*xp-ehh*	low	205 - 259	3.11 - 3.33	416 - 451	0.07 - 0.13	-0.08 - 0.88	*DCDC2*
chr2:58640000_58770000	130	*xp-ehh*	low	76 - 159	3.17 - 3.6	208 - 331	0.05 - 0.17	-0.39 - 1.56	*ENSGALG00000048667 (LncRNA)*
chr2:61180000_61210000	30	*xp-ehh*	low	185 - 211	3.07 - 3.1	300 - 339	0.18 - 0.2	1.67 - 1.93	*JARID2*
chr2:61380000_61400000	20	*fst*	low	152 - 152	1.1 - 1.1	437 - 437	0.4 - 0.4	5.02 - 5.02	
chr2:65510000_65540000	30	*xp-ehh*	low	207 - 229	3.09 - 3.2	363 - 401	0.1 - 0.14	0.45 - 1.03	*ENSGALG00000047910 (LncRNA), LYRM4*
chr2:65560000_65580000	20	*xp-ehh*	low	130 - 130	3.2 - 3.2	488 - 488	0.23 - 0.23	2.38 - 2.38	*ENSGALG00000053424*
chr2:65630000_65650000	20	*xp-ehh*	low	165 - 165	3.02 - 3.02	490 - 490	0.14 - 0.14	1.04 - 1.04	*ENSGALG00000012808*
chr2:66210000_66230000	20	*xp-ehh*	low	176 - 176	3.03 - 3.03	328 - 328	0.1 - 0.1	0.43 - 0.43	*SLC22A23*
chr2:68080000_68110000	30	*fst*	high	137 - 181	1.33 - 1.59	332 - 406	0.42 - 0.42	5.34 - 5.36	*PHLPP1*
chr2:83520000_83540000	20	*xp-ehh*	low	109 - 109	3.09 - 3.09	403 - 403	0.39 - 0.39	4.94 - 4.94	*ENSGALG00000052392 (LncRNA), RPRD1A*
chr2:127110000_127130000	20	*fst*	high	132 - 132	0.94 - 0.94	345 - 345	0.41 - 0.41	5.21 - 5.21	
chr2:135380000_135440000	60	*xp-ehh*	low	63 - 123	3.02 - 3.31	204 - 319	0.19 - 0.31	1.8 - 3.66	*EIF3H*
chr2:135760000_135790000	30	*xp-ehh*	low	189 - 203	3.03 - 3.07	393 - 411	0.14 - 0.19	0.97 - 1.81	*MED30*
chr2:135870000_135900000	30	*xp-ehh*	low	136 - 195	3.1 - 3.16	283 - 351	0.11 - 0.16	0.54 - 1.34	*EXT1*
chr2:136080000_136100000	20	*xp-ehh*	low	159 - 159	3.35 - 3.35	310 - 310	0.2 - 0.2	2.01 - 2.01	*SAMD12*
chr2:139590000_139620000	30	*fst*	high	98 - 109	1.41 - 1.76	157 - 194	0.42 - 0.52	5.38 - 6.99	
chr3:21650000_21700000	50	*xp-ehh*	high	145 - 166	3.01 - 3.23	264 - 291	0.16 - 0.2	1.37 - 2.01	*ENSGALG00000047938 (LncRNA), ENSGALG00000052926 (LncRNA)*
chr3:28650000_28690000	40	*xp-ehh*	high	128 - 180	3.07 - 3.2	282 - 360	0.11 - 0.21	0.61 - 2.16	
chr3:71850000_71870000	20	*fst*	high	110 - 110	1.9 - 1.9	186 - 186	0.43 - 0.43	5.46 - 5.46	
chr3:87810000_87870000	60	*xp-ehh*	low	109 - 162	3.08 - 3.38	256 - 443	0.24 - 0.29	2.62 - 3.3	*ENSGALG00000049936 (LncRNA), TINAG*
chr4:1090000_1130000	40	*xp-ehh*	low	152 - 191	3.23 - 3.49	282 - 351	0.2 - 0.24	1.93 - 2.55	*FAM155B*

chr4:4570000_4590000	20	*xp-ehh*	high	236 - 236	3.18 - 3.18	445 - 445	0.25 - 0.25	2.77 - 2.77	*ENSGALG00000049919 (LncRNA)*
chr4:15360000_15380000	20	*xp-ehh*	low	95 - 95	3.08 - 3.08	223 - 223	0.27 - 0.27	3.11 - 3.11	*TENM1*
chr4:24500000_24530000	30	*xp-ehh*	low	192 - 204	3.06 - 3.1	448 - 488	0.08 - 0.12	0.14 - 0.7	*SPOCK3*
chr4:24870000_24910000	40	*xp-ehh*	low	216 - 248	3.06 - 3.38	402 - 441	0.13 - 0.16	0.9 - 1.36	*ENSGALG00000049163 (LncRNA)*
chr4:25200000_25220000	20	*xp-ehh*	low	219 - 219	3.01 - 3.01	464 - 464	0.25 - 0.25	2.81 - 2.81	
chr4:25450000_25550000	100	*xp-ehh*	low	182 - 242	3.16 - 3.83	336 - 454	0.2 - 0.27	1.95 - 3.09	*CBR4, SH3RF1*
chr4:29770000_29790000	20	*xp-ehh*	low	252 - 252	3.16 - 3.16	409 - 409	0.15 - 0.15	1.14 - 1.14	*RNF150*
chr4:29820000_29870000	50	*xp-ehh*	low	224 - 237	3.03 - 3.21	391 - 463	0.12 - 0.18	0.79 - 1.64	*RNF150*
chr4:31960000_31980000	20	*xp-ehh*	low	227 - 227	3.03 - 3.03	481 - 481	0.21 - 0.21	2.19 - 2.19	*PRMT9, TMEM184C*
chr4:32220000_32240000	20	*xp-ehh*	low	258 - 258	3.03 - 3.03	354 - 354	0.1 - 0.1	0.44 - 0.44	*NR3C2*
chr4:32340000_32360000	20	*xp-ehh*	low	144 - 144	3 - 3	245 - 245	0.07 - 0.07	-0.11 - -0.11	*ENSGALG00000030689 (LncRNA)*
chr4:37920000_37940000	20	*xp-ehh*	low	244 - 244	3.01 - 3.01	416 - 416	0.05 - 0.05	-0.31 - -0.31	*HADH, LEF1*
chr4:39410000_39440000	30	*fst*	high	91 - 121	0.63 - 1.01	266 - 289	0.41 - 0.45	5.14 - 5.87	*ACSL1, ENSGALG00000050111 (LncRNA)*
chr4:59770000_59800000	30	*xp-ehh*	high	182 - 190	3.02 - 3.13	327 - 338	0.14 - 0.17	1.06 - 1.54	*RAP1GDS1, TSPAN5*
chr4:65470000_65500000	30	*xp-ehh*	low	194 - 203	3.1 - 3.11	427 - 451	0.12 - 0.12	0.76 - 0.76	*CHIC2*
chr4:89580000_89600000	20	*fst*	low	158 - 158	0.83 - 0.83	384 - 384	0.41 - 0.41	5.17 - 5.17	*ADAM33, ENSGALG00000044620*
chr4:89960000_89990000	30	*xp-ehh*	high	96 - 146	3.05 - 3.18	346 - 362	0.14 - 0.22	1.01 - 2.21	*EXOC6B*
chr5:29730000_29750000	20	*fst*	high	221 - 221	1.25 - 1.25	441 - 441	0.4 - 0.4	5.14 - 5.14	*GPR176*
chr5:39190000_39350000	160	*xp-ehh*	low	66 - 199	3.03 - 3.51	136 - 352	0.02 - 0.33	-0.8 - 3.91	*CIPC, ENSGALG00000048945 (LncRNA), ENSGALG00000049515 (LncRNA), ENSGALG00000050873 (LncRNA), gga-mir-6645 (miRNA), IRF2BPL, NGB, NSGALG00000024525 (5S_rRNA), TMEM63C, ZDHHC22*
chr5:43320000_43360000	40	*xp-ehh*	low	177 - 208	3.12 - 3.22	332 - 412	0.14 - 0.19	0.97 - 1.78	*EML5, PTPN21, ZC3H14*
chr5:43580000_43610000	30	*xp-ehh*	low	157 - 205	3.05 - 3.18	285 - 359	0.03 - 0.08	-0.62 - 0.1	*ENSGALG00000047714 (LncRNA), FOXN3*
chr5:55470000_55500000	30	*xp-ehh*	low	182 - 183	3.1 - 3.24	366 - 404	0.05 - 0.07	-0.41 - -0.08	*ACTR10, ENSGALG00000012085, PSMA3*
chr7:8590000_8610000	20	*xp-ehh*	low	35 - 35	3.04 - 3.04	74 - 74	0.32 - 0.32	3.81 - 3.81	
chr7:28430000_28460000	30	*xp-ehh*	low	141 - 194	3.02 - 3.14	354 - 385	0.1 - 0.12	0.36 - 0.78	*C1QL2*

chr7:30450000_30480000	30	*xp-ehh*	low	143 - 162	3.09 - 3.11	323 - 364	0.14 - 0.18	1.01 - 1.64	*ACMSD, ENSGALG00000012207, ENSGALG00000046969, ENSGALG00000054339*
chr7:34590000_34610000	20	*fst*	low	88 - 88	0.12 - 0.12	302 - 302	0.41 - 0.41	5.16 - 5.16	
chr7:34830000_34860000	30	*xp-ehh*	low	140 - 201	3.07 - 3.12	378 - 424	0.17 - 0.2	1.46 - 1.92	*KIF5C*
chr8:6440000_6460000	20	*xp-ehh*	low	178 - 178	3.01 - 3.01	416 - 416	0.1 - 0.1	0.43 - 0.43	*TOR1AIP1, TOR1AIP2*
chr8:10350000_10380000	30	*fst*	low	47 - 49	1.96 - 1.96	62 - 73	0.4 - 0.41	5.01 - 5.23	*ENSGALG00000043258 (LncRNA)*
chr8:10430000_10670000	240	*fst*	low	4 - 48	1.17 - 2.23	30 - 71	0.41 - 0.49	5.21 - 6.44	*ENSGALG00000047029, PLA2G4A*
chr8:10680000_10700000	20	*fst*	low	9 - 9	1.45 - 1.45	56 - 56	0.4 - 0.4	5.01 - 5.01	*PLA2G4A*
chr8:22190000_22210000	20	*xp-ehh*	low	234 - 234	3.09 - 3.09	526 - 526	0.09 - 0.09	0.31 - 0.31	*ATPAF1, EFCAB14, MOB3C*
chr8:23360000_23420000	60	*xp-ehh*	low	111 - 173	3.05 - 3.48	234 - 378	0.07 - 0.15	-0.11 - 1.27	*BEND5*
chr8:27970000_27990000	20	*xp-ehh*	low	161 - 161	3.2 - 3.2	352 - 352	0.15 - 0.15	1.25 - 1.25	*ENSGALG00000047521 (LncRNA), ENSGALG00000053679 (LncRNA)*
chr8:28310000_28340000	30	*xp-ehh*	low	209 - 210	3.08 - 3.29	357 - 367	0.16 - 0.16	1.32 - 1.34	*ROR1*
chr9:1240000_1260000	20	*xp-ehh*	low	144 - 144	3.02 - 3.02	266 - 266	0.09 - 0.09	0.33 - 0.33	*NEU2, NGEF*
chr9:8040000_8070000	30	*fst*	low	104 - 140	1.55 - 1.86	270 - 272	0.41 - 0.42	5.19 - 5.34	*Pax3*
chr9:20250000_20300000	50	*xp-ehh*	low	157 - 191	3.22 - 3.36	278 - 307	0.14 - 0.17	1.1 - 1.58	*MECOM*
chr10:13630000_13730000	100	*xp-ehh*	high	134 - 256	3.08 - 3.41	282 - 515	0.13 - 0.28	0.82 - 3.18	*NTRK3*
chr10:17820000_17840000	20	*xp-ehh*	high	86 - 86	3.1 - 3.1	211 - 211	0.2 - 0.2	1.9 - 1.9	*ENSGALG00000053260 (LncRNA)*
chr10:20300000_20350000	50	*xp-ehh*	high	186 - 226	3.03 - 3.42	343 - 413	0.08 - 0.15	0.13 - 1.15	*BLM, FRMD5, WDR76*
chr11:4580000_4600000	20	*fst*	low	85 - 85	0.18 - 0.18	201 - 201	0.4 - 0.4	5.07 - 5.07	
chr11:16300000_16320000	20	*fst*	high	151 - 151	1.57 - 1.57	250 - 250	0.4 - 0.4	5.13 - 5.13	*CDH13*
chr13:1280000_1310000	30	*xp-ehh*	low	51 - 93	3.04 - 3.1	128 - 191	0.24 - 0.27	2.64 - 3.11	*ENSGALG00000052992*
chr13:2570000_2590000	20	*fst*	low	32 - 32	1.08 - 1.08	104 - 104	0.4 - 0.4	5.13 - 5.13	*PCDHGA2*
chr13:2940000_2980000	40	*fst*	low	69 - 92	1.31 - 1.46	229 - 267	0.45 - 0.54	5.9 - 7.25	*ENSGALG00000042617 (LncRNA), ENSGALG00000051725 (LncRNA), ENSGALG00000054361 (LncRNA)*
chr14:4400000_4440000	40	*xp-ehh*	high	194 - 266	3.16 - 3.66	371 - 429	0.13 - 0.19	0.84 - 1.89	*AP5Z1, ENSGALG00000004504, gga-mir-12224, gga-mir-2129, MMD2*
chr15:4930000_4950000	20	*xp-ehh*	high	162 - 162	3 - 3	288 - 288	0.22 - 0.22	2.31 - 2.31	*NCOR2*
chr15:7930000_8000000	70	*xp-ehh*	high	160 - 242	3.03 - 3.47	294 - 439	0.12 - 0.23	0.77 - 2.46	*CCDC117, CHEK2, SLC2A11L5, gga-mir-6650, HSCB, TTC28*
chr15:10640000_10670000	30	*xp-ehh*	high	176 - 200	3.08 - 3.3	314 - 355	0.17 - 0.18	1.43 - 1.66	*RTN4R*

chr15:10680000_10700000	20	*xp-ehh*	high	167 - 167	3.05 - 3.05	274 - 274	0.12 - 0.12	0.69 - 0.69	*ENSGALG00000033639 (LncRNA), RTN4R*
chr15:10710000_10750000	40	*xp-ehh*	high	183 - 203	3.08 - 3.41	269 - 317	0.14 - 0.2	1.01 - 1.9	*ENSGALG00000047400 (LncRNA)*
chr17:2810000_2840000	30	*xp-ehh*	low	95 - 124	3.14 - 3.31	329 - 396	0.16 - 0.19	1.35 - 1.87	*AKNA, gga-mir-7442, STXBP1, WHRN*
chr17:7790000_7820000	30	*xp-ehh*	low	192 - 193	3.32 - 3.32	348 - 372	0.21 - 0.29	2.19 - 3.36	*COL5A1*
chr19:8420000_8470000	50	*xp-ehh*	low	147 - 157	3.18 - 3.4	246 - 273	0.08 - 0.21	0.06 - 2.06	
chr19:8780000_8820000	40	*xp-ehh/fst*	low	60 - 71	2.65 - 3	218 - 221	0.19 - 0.41	1.88 - 5.17	*DDX52, HEATR6, HNF1B*
chr20:9120000_9160000	40	*xp-ehh*	low	163 - 228	3.07 - 3.51	349 - 452	0.11 - 0.16	0.62 - 1.36	*ENSGALG00000049648 (LncRNA), MYT1*
chr22:1060000_1090000	30	*fst*	high	7 - 8	1.14 - 1.33	34 - 35	0.42 - 0.47	5.34 - 6.12	*NEFL*
chr22:1150000_1180000	30	*fst*	high	8 - 15	0.33 - 0.37	35 - 53	0.46 - 0.49	6.01 - 6.42	
chr22:5290000_5320000	30	*xp-ehh*	high	69 - 90	3.04 - 3.09	215 - 247	0.03 - 0.1	-0.61 - 0.39	*CASP14, ENSGALG00000047954, TGOLN2, TCF7L1*
chr23:5550000_5610000	60	*fst*	low	65 - 86	0.97 - 1.82	200 - 224	0.4 - 0.5	5.08 - 6.61	*BMP8A, MACF1*
chr24:1310000_1330000	20	*xp-ehh*	high	271 - 271	3.12 - 3.12	439 - 439	0.16 - 0.16	1.35 - 1.35	
chr24:2930000_2960000	30	*xp-ehh*	low	191 - 199	3.02 - 3.11	339 - 359	0.15 - 0.19	1.17 - 1.86	*GRAMD1B*
chr26:2140000_2180000	40	*xp-ehh*	low	148 - 185	3.17 - 3.46	335 - 348	0.06 - 0.08	-0.19 - 0.17	*CDK18, MFSD4A*
chr27:4810000_4850000	40	*xp-ehh*	high	156 - 215	3.08 - 3.19	278 - 389	0.09 - 0.14	0.31 - 1.06	*MRC2, TLK2*

Table S4.2d. Candidate selection signature regions and overlapping genes detected from analyses of the precipitation in the driest quarter of a year (precDQ). (**SSR_ID:** selection signature regions ID as chromosome:start_stop position. **DD:** directionality of selection)

SSR_ID	**size (kb)**	**detection method**	**DD**	**no. of SNPs per *XP-EHH* window (range)**	**absolute value of mean *XP-EHH*_std (range)**	**no. of SNPs per *Fst* window (range)**	**weighted *Fst* per window (range)**	**Z*Fst* per window (range)**	**Overlapping genes**
chr1:9000000_9020000	20	*fst*	low	66 - 66	1.85 - 1.85	135 - 135	0.23 - 0.23	5.18 - 5.18	
chr1:13780000_13800000	20	*fst*	low	127 - 127	0.59 - 0.59	209 - 209	0.22 - 0.22	5.03 - 5.03	
chr1:15060000_15080000	20	*xp-ehh*	low	274 - 274	3.29 - 3.29	389 - 389	0.14 - 0.14	2.83 - 2.83	*LAMB1, LAMB4*
chr1:33810000_33830000	20	*xp-ehh*	high	269 - 269	3.05 - 3.05	402 - 402	0.1 - 0.1	1.78 - 1.78	*RASSF3*

chr1:52790000_52820000	30	*fst*	high	94 - 117	0.71 - 1.01	246 - 271	0.22 - 0.24	5.05 - 5.43	*LARGE1*
chr1:60010000_60030000	20	*xp-ehh*	low	238 - 238	3.31 - 3.31	335 - 335	0.04 - 0.04	0.27 - 0.27	*ENSGALG00000048398 (LncRNA), IQSEC3*
chr1:65410000_65440000	30	*xp-ehh*	low	145 - 175	3.16 - 3.24	222 - 273	0.06 - 0.07	0.81 - 1.16	*BCAT1, ENSGALG00000048674*
chr1:78060000_78080000	20	*fst*	high	148 - 148	1.65 - 1.65	248 - 248	0.22 - 0.22	5.05 - 5.05	*C1R, CLSTN3, RBP5*
chr1:81240000_81260000	20	*fst*	low	215 - 215	1.18 - 1.18	369 - 369	0.24 - 0.24	5.42 - 5.42	*CASQ2*
chr1:88290000_88330000	40	*xp-ehh*	low	144 - 182	3.37 - 3.93	365 - 405	0.12 - 0.16	2.46 - 3.48	*BBX*
chr1:97320000_97340000	20	*xp-ehh*	low	114 - 114	3.3 - 3.3	292 - 292	0.15 - 0.15	3.21 - 3.21	*ENSGALG00000054585 (LncRNA)*
chr1:107790000_107810000	20	*fst*	low	61 - 61	1.41 - 1.41	205 - 205	0.22 - 0.22	5.04 - 5.04	*ENSGALG00000050248 (LncRNA)*
chr1:109700000_109730000	30	*xp-ehh*	low	181 - 191	3.15 - 3.18	314 - 366	0.09 - 0.14	1.74 - 2.98	*DSCAM*
chr1:109740000_109790000	50	*xp-ehh*	low	132 - 171	3.02 - 3.08	265 - 302	0.08 - 0.15	1.29 - 3.12	*DSCAM, ENSGALG00000042372 (LncRNA), ENSGALG00000052898(ncRNA)*
chr1:110360000_110390000	30	*xp-ehh / fst*	low	240 - 270	2.98 - 3.21	382 - 395	0.18 - 0.22	4.11 - 5.12	*ENSGALG00000051940 (LncRNA), RIPK4*
chr1:111190000_111210000	20	*fst*	low	161 - 161	1.99 - 1.99	288 - 288	0.25 - 0.25	5.74 - 5.74	
chr1:112350000_112370000	20	*xp-ehh*	low	188 - 188	3.02 - 3.02	435 - 435	0.06 - 0.06	0.8 - 0.8	*FUNDC1*
chr1:113640000_113670000	30	*xp-ehh*	low	117 - 179	3.2 - 3.76	217 - 310	0.06 - 0.11	0.78 - 2.16	*ENSGALG00000046706 (LncRNA)*
chr1:113700000_113720000	20	*xp-ehh*	low	134 - 134	3.22 - 3.22	287 - 287	0.13 - 0.13	2.7 - 2.7	*BCOR*
chr1:128020000_128040000	20	*fst*	low	82 - 82	2.31 - 2.31	113 - 113	0.22 - 0.22	5.01 - 5.01	
chr1:128240000_128260000	20	*xp-ehh*	low	53 - 53	3.04 - 3.04	88 - 88	0.09 - 0.09	1.67 - 1.67	*NLGN4Y*
chr1:130660000_130690000	30	*xp-ehh*	low	150 - 156	3.3 - 3.36	248 - 275	0.04 - 0.06	0.45 - 0.82	
chr1:134820000_134840000	20	*xp-ehh*	high	187 - 187	3.27 - 3.27	385 - 385	0.11 - 0.11	2.06 - 2.06	*RFX8*
chr1:153410000_153460000	50	*fst*	high	9 - 39	1.68 - 2.02	79 - 113	0.26 - 0.34	5.94 - 8.22	
chr1:155020000_155060000	40	*fst*	low	90 - 120	0.2 - 0.35	253 - 267	0.23 - 0.24	5.2 - 5.51	*ENSGALG00000051545 (LncRNA)*
chr1:159570000_159590000	20	*fst*	low	100 - 100	1.17 - 1.17	172 - 172	0.22 - 0.22	5.1 - 5.1	
chr1:159600000_159640000	40	*fst*	low	109 - 150	1.18 - 1.58	224 - 294	0.24 - 0.29	5.54 - 6.74	
chr1:178310000_178340000	30	*xp-ehh*	high	191 - 217	3.02 - 3.23	305 - 305	0.07 - 0.08	1.17 - 1.27	*ATP8A2*
chr2:2960000_2990000	30	*fst*	low	183 - 217	1.9 - 1.95	255 - 351	0.23 - 0.24	5.38 - 5.58	
chr2:5460000_5480000	20	*fst*	low	185 - 185	0.08 - 0.08	292 - 292	0.31 - 0.31	7.24 - 7.24	*ENSGALG00000006072*
chr2:10260000_10280000	20	*xp-ehh*	low	213 - 213	3.09 - 3.09	404 - 404	0.05 - 0.05	0.61 - 0.61	*DIP2C*
chr2:11620000_11650000	30	*fst*	high	77 - 113	0.85 - 1.07	141 - 181	0.24 - 0.24	5.42 - 5.45	
chr2:12200000_12220000	20	*fst*	low	33 - 33	0.28 - 0.28	106 - 106	0.23 - 0.23	5.27 - 5.27	
chr2:26940000_26970000	30	*fst*	low	150 - 160	0.48 - 0.53	203 - 210	0.23 - 0.23	5.15 - 5.38	*ENSGALG00000050874 (LncRNA)*

chr2:28010000_28040000	30	*fst*	low	101 - 133	1.4 - 1.51	263 - 319	0.23 - 0.24	5.19 - 5.42	*AGMO, MEOX2*
chr2:28060000_28080000	20	*fst*	low	68 - 68	1.13 - 1.13	175 - 175	0.23 - 0.23	5.34 - 5.34	*MEOX2*
chr2:28120000_28210000	90	*fst*	low	42 - 82	0.69 - 1.73	108 - 194	0.23 - 0.27	5.4 - 6.35	*ENSGALG00000054958 (LncRNA), ENSGALG00000027900 (5S_rRNA)*
chr2:46360000_46380000	20	*xp-ehh*	high	166 - 166	3.09 - 3.09	348 - 348	0.05 - 0.05	0.55 - 0.55	*ELMO1*
chr2:46670000_46690000	20	*fst*	high	189 - 189	2.39 - 2.39	422 - 422	0.24 - 0.24	5.48 - 5.48	*ANLN, AOAH*
chr2:55230000_55290000	60	*xp-ehh*	high	138 - 198	3.32 - 3.82	251 - 367	0.11 - 0.17	2.09 - 3.62	*ENSGALG00000054801 (LncRNA), TNS3*
chr2:95670000_95800000	130	*fst*	high	33 - 138	2.13 - 2.67	58 - 197	0.24 - 0.3	5.63 - 7.1	
chr2:103790000_103820000	30	*fst*	low	140 - 156	1.13 - 1.4	178 - 210	0.25 - 0.26	5.84 - 6.11	*gga-mir-1597, ZNF521*
chr2:113980000_114020000	40	*xp-ehh*	high	125 - 167	3.14 - 3.3	215 - 267	0.05 - 0.14	0.67 - 2.82	
chr2:114940000_114960000	20	*xp-ehh*	high	185 - 185	3.05 - 3.05	271 - 271	0.06 - 0.06	0.94 - 0.94	*DNAJC5B*
chr2:123760000_123780000	20	*xp-ehh*	low	59 - 59	3.1 - 3.1	123 - 123	0.01 - 0.01	-0.46 - -0.46	
chr2:124060000_124150000	90	*xp-ehh*	low	80 - 205	3.01 - 3.47	282 - 442	0.06 - 0.14	0.91 - 2.83	*ENSGALG00000049271 (LncRNA)*
chr2:124380000_124410000	30	*xp-ehh*	low	257 - 273	3.09 - 3.13	405 - 413	0.02 - 0.03	-0.1 - 0.17	
chr2:145640000_145670000	30	*xp-ehh*	high	160 - 222	3.43 - 3.46	322 - 353	0.12 - 0.13	2.41 - 2.71	*AGO2, ENSGALG00000052658 (LncRNA), gga-mir-6572*
chr3:1160000_1210000	50	*fst*	high	5 - 44	1.49 - 2.14	55 - 90	0.23 - 0.39	5.19 - 9.42	*ENSGALG00000047366 (LncRNA)*
chr3:5550000_5570000	20	*xp-ehh*	high	157 - 157	3.02 - 3.02	317 - 317	0.1 - 0.1	1.76 - 1.76	*KIF16B*
chr3:8430000_8450000	20	*fst*	low	233 - 233	0.7 - 0.7	399 - 399	0.22 - 0.22	5.05 - 5.05	*GALNT14*
chr3:19210000_19230000	20	*xp-ehh*	high	155 - 155	3.19 - 3.19	264 - 264	0.12 - 0.12	2.44 - 2.44	*LYPLAL1*
chr3:20090000_20110000	20	*fst*	low	83 - 83	1.58 - 1.58	152 - 152	0.23 - 0.23	5.31 - 5.31	*ESRRG*
chr3:39090000_39110000	20	*xp-ehh*	low	180 - 180	3.01 - 3.01	285 - 285	0.07 - 0.07	1.03 - 1.03	*ENSGALG00000048291, ENSGALG00000050186, NTPCR*
chr3:49700000_49720000	20	*fst*	high	168 - 168	0.63 - 0.63	293 - 293	0.27 - 0.27	6.24 - 6.24	*FBXO5, MTRF1L*
chr3:50830000_50860000	30	*fst*	low	79 - 84	0.65 - 1.05	255 - 271	0.23 - 0.24	5.27 - 5.58	*ENSGALG00000047502 (LncRNA)*
chr3:70090000_70160000	70	*fst*	low	42 - 57	0.98 - 1.61	74 - 124	0.23 - 0.29	5.22 - 6.9	*ENSGALG00000051525, GRIK2*
chr3:70280000_70300000	20	*fst*	low	143 - 143	1.34 - 1.34	186 - 186	0.23 - 0.23	5.19 - 5.19	*GRIK2*
chr3:70960000_71090000	130	*xp-ehh*	low	83 - 190	3.28 - 4.05	129 - 273	0.11 - 0.21	2.12 - 4.81	*ENSGALG00000053011 (LncRNA), SIM1*
chr3:71640000_71670000	30	*fst*	low	169 - 199	1.21 - 2.1	246 - 255	0.24 - 0.24	5.43 - 5.58	*ENSGALG00000047943 (LncRNA), ENSGALG00000052916 (LncRNA)*
chr3:71840000_72120000	280	*xp-ehh* and/or *fst*	low	21 - 179	3.01 - 3.75	96 - 315	0.05 - 0.24	0.62 - 5.67	*ENSGALG00000034564 (LncRNA), ENSGALG00000049307 (LncRNA), gga-mir-12262*

chr3:72130000_72150000	20	*xp-ehh*	low	81 - 81	3 - 3	147 - 147	0.06 - 0.06	0.95 - 0.95	
chr3:72220000_72340000	120	*xp-ehh*	low	69 - 148	3.01 - 3.63	108 - 248	0.04 - 0.22	0.37 - 4.92	*ENSGALG00000036204 (LncRNA)*
chr3:72360000_72540000	180	*xp-ehh* and/or *fst*	low	76 - 254	3.01 - 4.08	178 - 411	0.04 - 0.22	0.37 - 5.03	*ENSGALG00000036204 (LncRNA), KLHL32, MMS22L*
chr3:72840000_72880000	40	*fst*	low	119 - 165	2.79 - 2.86	210 - 276	0.24 - 0.32	5.61 - 7.71	
chr3:73000000_73050000	50	*fst*	low	66 - 138	1.92 - 2.24	143 - 241	0.25 - 0.33	5.85 - 7.82	*FUT9*
chr3:73120000_73250000	130	*xp-ehh* and/or *fst*	low	76 - 291	1.85 - 3.19	159 - 497	0.22 - 0.43	4.95 - 10.42	*MANEA, ENSGALG00000025686 (SnRNA)*
chr3:73300000_73390000	90	*fst*	low	36 - 100	2.84 - 2.95	132 - 174	0.22 - 0.31	5.03 - 7.42	*ENSGALG00000046901 (LncRNA), ENSGALG00000050194*
chr3:73530000_73840000	310	*xp-ehh* and/or *fst*	low	39 - 91	1.72 - 3.21	91 - 217	0.15 - 0.34	3.13 - 8.16	*ENSGALG00000026577 (MiRNA)*
chr3:73860000_73950000	90	*xp-ehh* and/or *fst*	low	42 - 138	2.92 - 3.4	87 - 238	0.12 - 0.32	2.37 - 7.56	*EPHA7*
chr3:73960000_73980000	20	*fst*	low	71 - 71	2 - 2	124 - 124	0.22 - 0.22	5 - 5	*EPHA7, gga-mir-1677*
chr3:74410000_74430000	20	*fst*	low	88 - 88	1.03 - 1.03	170 - 170	0.23 - 0.23	5.21 - 5.21	
chr3:74450000_74470000	20	*fst*	low	72 - 72	1.33 - 1.33	137 - 137	0.24 - 0.24	5.42 - 5.42	*ENSGALG00000047305 (MiRNA)*
chr3:74490000_74510000	20	*fst*	low	62 - 62	1.72 - 1.72	123 - 123	0.23 - 0.23	5.21 - 5.21	
chr3:78890000_78930000	40	*xp-ehh*	low	108 - 123	3.03 - 3.14	184 - 202	0.1 - 0.14	1.88 - 2.95	
chr3:86030000_86060000	30	*fst*	high	55 - 59	0.73 - 0.81	112 - 116	0.22 - 0.23	5.08 - 5.39	
chr3:86680000_86780000	100	*fst*	high	231 - 263	0.64 - 1.87	319 - 464	0.25 - 0.34	5.7 - 8.09	*BAG2, PRIM2, RAB23, ZNF451*
chr3:86780000_86800000	20	*fst*	low	263 - 263	0.74 - 0.74	395 - 395	0.28 - 0.28	6.58 - 6.58	*ZNF451*
chr3:87510000_87540000	30	*fst*	low	91 - 172	0.96 - 1.44	196 - 327	0.25 - 0.3	5.79 - 7.03	*HMGCLL1*
chr3:99060000_99080000	20	*fst*	low	216 - 216	2.35 - 2.35	336 - 336	0.23 - 0.23	5.16 - 5.16	
chr3:104340000_104370000	30	*xp-ehh*	low	169 - 252	3.35 - 3.37	392 - 462	0.09 - 0.11	1.62 - 2.27	*NUMBL, IFT172, IFT172*
chr3:106430000_106510000	80	*xp-ehh* and/or *fst*	low	157 - 238	2.12 - 3.25	300 - 399	0.14 - 0.29	2.99 - 6.81	*MSRA*
chr4:1350000_1420000	70	*xp-ehh*	low	162 - 308	3.26 - 3.95	359 - 491	0.06 - 0.1	0.86 - 1.98	*CYSLTR1, ENSGALG00000004127, ENSGALG00000053473 (LncRNA), LPAR4*
chr4:9970000_9990000	20	*fst*	high	121 - 121	0.58 - 0.58	249 - 249	0.24 - 0.24	5.59 - 5.59	*ENSGALG00000052590 (LncRNA)*
chr4:11090000_11110000	20	*xp-ehh*	low	181 - 181	3.07 - 3.07	420 - 420	0.07 - 0.07	1.15 - 1.15	
chr4:11220000_11250000	30	*xp-ehh*	low	184 - 221	3.06 - 3.14	283 - 353	0.07 - 0.1	1.05 - 1.82	*ENSGALG00000052786*
chr4:14070000_14090000	20	*xp-ehh*	low	238 - 238	3.28 - 3.28	404 - 404	0.07 - 0.07	1.01 - 1.01	
chr4:22570000_22600000	30	*fst*	low	147 - 204	0.14 - 0.14	217 - 245	0.24 - 0.28	5.43 - 6.48	

chr4:23350000_23430000	80	*xp-ehh* / *fst*	high	111 - 205	2.13 - 3.14	266 - 384	0.13 - 0.32	2.76 - 7.71	*ENSGALG00000046688 (LncRNA), NPY1R, NPY5R, TMA16*
chr4:27820000_27880000	60	*fst*	high	74 - 97	1.88 - 2.04	108 - 156	0.22 - 0.31	5.11 - 7.46	*ENSGALG00000051876 (LncRNA)*
chr4:27890000_27910000	20	*fst*	high	84 - 84	2.05 - 2.05	123 - 123	0.24 - 0.24	5.43 - 5.43	
chr4:28120000_28140000	20	*fst*	high	86 - 86	2.14 - 2.14	120 - 120	0.23 - 0.23	5.21 - 5.21	
chr4:28210000_28230000	20	*fst*	high	40 - 40	2.2 - 2.2	89 - 89	0.23 - 0.23	5.26 - 5.26	*ENSGALG00000050821 (LncRNA), ENSGALG00000053850 (LncRNA)*
chr4:28270000_28300000	30	*fst*	high	31 - 51	2.09 - 2.11	58 - 80	0.22 - 0.23	5.11 - 5.38	*ENSGALG00000047082 (LncRNA)*
chr4:28350000_28380000	30	*fst*	high	68 - 70	2.13 - 2.14	102 - 122	0.23 - 0.24	5.17 - 5.64	
chr4:28490000_28520000	30	*xp-ehh*	high	43 - 55	3.02 - 3.11	148 - 150	0.15 - 0.16	3.3 - 3.43	*ENSGALG00000048582 (LncRNA), ENSGALG00000054220 (LncRNA)*
chr4:28770000_28870000	100	*xp-ehh*	high	138 - 208	3.15 - 3.49	269 - 369	0.07 - 0.14	1.02 - 2.95	
chr4:29620000_29640000	20	*xp-ehh*	high	222 - 222	3.38 - 3.38	366 - 366	0.11 - 0.11	2.13 - 2.13	*CLGN, ENSGALG00000048914 (LncRNA), SCOC*
chr4:35180000_35200000	20	*fst*	low	241 - 241	2.46 - 2.46	371 - 371	0.23 - 0.23	5.24 - 5.24	*FAM13A*
chr4:43270000_43290000	20	*xp-ehh*	low	195 - 195	3.02 - 3.02	301 - 301	0.09 - 0.09	1.71 - 1.71	*GALNT7*
chr4:48300000_48320000	20	*xp-ehh*	high	191 - 191	3.16 - 3.16	302 - 302	0.14 - 0.14	2.93 - 2.93	*ADGRL3*
chr4:62850000_62870000	20	*xp-ehh*	high	213 - 213	3.22 - 3.22	442 - 442	0.02 - 0.02	-0.2 - -0.2	*FGL1*
chr4:70510000_70530000	20	*fst*	low	137 - 137	0.86 - 0.86	235 - 235	0.26 - 0.26	5.96 - 5.96	
chr4:74170000_74300000	130	*xp-ehh* and/or *fst*	low	76 - 165	1.32 - 3.06	148 - 234	0.22 - 0.31	5.08 - 7.46	*ENSGALG00000046053 (LncRNA), ENSGALG00000048521 (LncRNA), ENSGALG00000050078 (LncRNA), ENSGALG00000051573 (LncRNA)*
chr4:74430000_74540000	110	*xp-ehh*	low	118 - 216	3.02 - 3.42	213 - 351	0.1 - 0.22	1.79 - 4.92	*ADGRA3, ENSGALG00000048082, ENSGALG00000054975 (LncRNA)*
chr4:74550000_74620000	70	*xp-ehh*	low	105 - 175	3.06 - 4.05	232 - 263	0.1 - 0.14	1.93 - 2.83	*ENSGALG00000051973 (LncRNA), ENSGALG00000054975 (LncRNA)*
chr4:74650000_74670000	20	*fst*	low	158 - 158	1.96 - 1.96	250 - 250	0.26 - 0.26	6.04 - 6.04	
chr4:89950000_89980000	30	*xp-ehh*	low	227 - 229	3.06 - 3.24	375 - 380	0.15 - 0.17	3.17 - 3.59	*EXOC6B*
chr5:45810000_45830000	20	*fst*	high	182 - 182	0.08 - 0.08	219 - 219	0.23 - 0.23	5.41 - 5.41	
chr6:2890000_2910000	20	*xp-ehh*	high	183 - 183	3 - 3	275 - 275	0.02 - 0.02	-0.19 - -0.19	
chr6:12510000_12530000	20	*xp-ehh*	high	253 - 253	3.04 - 3.04	375 - 375	0.08 - 0.08	1.27 - 1.27	*CDH23, SLC29A3*
chr6:21700000_21720000	20	*fst*	low	304 - 304	2.17 - 2.17	538 - 538	0.22 - 0.22	5.01 - 5.01	*IDE, KIF11*
chr6:21780000_21810000	30	*fst*	low	221 - 243	1.56 - 1.63	367 - 448	0.31 - 0.32	7.32 - 7.73	*MARCH5*
chr6:21870000_21890000	20	*fst*	low	106 - 106	1.05 - 1.05	288 - 288	0.27 - 0.27	6.35 - 6.35	*CPEB3*

chr6:25040000_25070000	30	*xp-ehh / fst*	low	140 - 169	3.06 - 3.25	311 - 420	0.24 - 0.26	5.45 - 6.13	*SLK*
chr7:5670000_5690000	20	*xp-ehh*	low	203 - 203	3.07 - 3.07	346 - 346	0.09 - 0.09	1.6 - 1.6	
chr7:10350000_10370000	20	*xp-ehh*	low	160 - 160	3.02 - 3.02	270 - 270	0.05 - 0.05	0.65 - 0.65	*PLCL1*
chr7:12960000_12990000	30	*xp-ehh*	high	246 - 268	3.02 - 3.24	359 - 364	0.02 - 0.03	-0.29 - 0.15	*ENSGALG00000046972 (LncRNA), PARD3B*
chr7:14390000_14410000	20	*fst*	low	87 - 87	2.06 - 2.06	281 - 281	0.22 - 0.22	5.09 - 5.09	*ENSGALG00000051682 (LncRNA)*
chr7:29150000_29180000	30	*xp-ehh*	high	212 - 227	3.01 - 3.17	419 - 437	0.05 - 0.06	0.67 - 0.93	*ENSGALG00000050982 (LncRNA), ENSGALG00000054456 (LncRNA)*
chr7:31030000_31090000	60	*xp-ehh*	low	113 - 186	3.09 - 3.34	310 - 358	0.13 - 0.18	2.59 - 3.86	*THSD7B*
chr7:33130000_33150000	20	*xp-ehh*	high	127 - 127	3.05 - 3.05	247 - 247	0.06 - 0.06	0.95 - 0.95	*GTDC1*
chr7:33280000_33310000	30	*xp-ehh*	high	144 - 158	3.44 - 3.51	309 - 317	0.14 - 0.18	2.81 - 3.88	*ENSGALG00000053032 (LncRNA), ZEB2*
chr7:34370000_34410000	40	*xp-ehh*	low	168 - 188	3.04 - 3.24	268 - 321	0.13 - 0.19	2.72 - 4.17	*ENSGALG00000053295 (LncRNA)*
chr7:35150000_35170000	20	*xp-ehh*	high	188 - 188	3.28 - 3.28	428 - 428	0.12 - 0.12	2.35 - 2.35	*RIF1*
chr7:35350000_35400000	50	*xp-ehh / fst*	high	228 - 263	1.48 - 3.49	412 - 452	0.2 - 0.25	4.49 - 5.69	*CACNB4*
chr7:36100000_36120000	20	*fst*	low	158 - 158	0.64 - 0.64	325 - 325	0.23 - 0.23	5.33 - 5.33	*GPD2*
chr8:15370000_15440000	70	*xp-ehh*	low	147 - 251	3.21 - 3.67	257 - 376	0.08 - 0.2	1.37 - 4.4	*LRRC8D, ZNF326*
chr8:15640000_15660000	20	*fst*	high	223 - 223	0.26 - 0.26	342 - 342	0.25 - 0.25	5.93 - 5.93	*ENSGALG00000052940 (LncRNA)*
chr8:20280000_20300000	20	*xp-ehh*	low	184 - 184	3.06 - 3.06	260 - 260	0.17 - 0.17	3.62 - 3.62	*PTPRF*
chr8:20440000_20480000	40	*xp-ehh*	low	129 - 200	3.41 - 3.78	233 - 305	0.11 - 0.17	2.26 - 3.82	*PTPRF*
chr8:26030000_26060000	30	*xp-ehh*	low	249 - 259	3.01 - 3.07	381 - 403	0.15 - 0.18	3.1 - 3.97	*ENSGALG00000049636 (LncRNA), PLPP3*
chr9:6290000_6310000	20	*fst*	high	162 - 162	0.14 - 0.14	290 - 290	0.24 - 0.24	5.49 - 5.49	
chr9:12140000_12170000	30	*xp-ehh*	high	132 - 175	3.08 - 3.1	237 - 246	0.02 - 0.02	-0.22 - -0.13	*ENSGALG00000049451 (LncRNA)*
chr9:12490000_12570000	80	*fst*	low	34 - 103	0.2 - 0.79	163 - 265	0.23 - 0.36	5.28 - 8.61	*AGTR1*
chr10:20210000_20240000	30	*xp-ehh*	low	85 - 113	3.07 - 3.17	181 - 222	0.12 - 0.12	2.28 - 2.31	*CTDSPL2*
chr11:1090000_1140000	50	*xp-ehh*	low	159 - 268	3.07 - 3.38	293 - 415	0.07 - 0.11	1.15 - 2.21	*CARMIL2, CTCFL, RIPOR1*
chr11:1660000_1680000	20	*fst*	low	133 - 133	1.76 - 1.76	292 - 292	0.23 - 0.23	5.24 - 5.24	*VAC14*
chr11:12500000_12520000	20	*fst*	low	157 - 157	1.64 - 1.64	253 - 253	0.23 - 0.23	5.34 - 5.34	
chr11:15360000_15380000	20	*xp-ehh*	low	291 - 291	3.49 - 3.49	447 - 447	0.09 - 0.09	1.62 - 1.62	*ATMIN, CENPN, CMC2*
chr11:18110000_18130000	20	*xp-ehh / fst*	low	163 - 163	3.32 - 3.32	268 - 268	0.24 - 0.24	5.58 - 5.58	*BANP*
chr12:17700000_17730000	30	*xp-ehh*	low	171 - 212	3.36 - 3.41	317 - 407	0.09 - 0.1	1.68 - 1.98	*CHL1*

chr12:17880000_17910000	30	*xp-ehh*	low	133 - 222	3.77 - 3.79	302 - 362	0.07 - 0.09	1.16 - 1.5	*ENSGALG00000052938*
chr12:18080000_18110000	30	*xp-ehh*	low	129 - 131	3.08 - 3.18	296 - 313	0.1 - 0.13	1.91 - 2.55	*ENSGALG00000046751 (LncRNA), ENSGALG00000050333 (LncRNA)*
chr13:1310000_1330000	20	*fst*	low	115 - 115	0.48 - 0.48	176 - 176	0.23 - 0.23	5.16 - 5.16	
chr13:5190000_5210000	20	*xp-ehh*	high	204 - 204	3 - 3	334 - 334	0.11 - 0.11	2.21 - 2.21	*SLIT3*
chr13:6520000_6570000	50	*xp-ehh*	high	112 - 153	3.01 - 3.39	213 - 253	0.08 - 0.11	1.33 - 2.19	
chr13:7320000_7340000	20	*fst*	low	103 - 103	1.87 - 1.87	342 - 342	0.23 - 0.23	5.35 - 5.35	*ENSGALG00000054885 (LncRNA)*
chr13:17870000_17890000	20	*xp-ehh*	low	158 - 158	3.09 - 3.09	300 - 300	0.19 - 0.19	4.14 - 4.14	*GNPDA1, NDFIP1, RNF14*
chr14:9130000_9170000	40	*xp-ehh*	low	193 - 214	3.14 - 3.18	304 - 338	0.03 - 0.06	-0.01 - 0.82	*ENSGALG00000007070, ENSGALG00000054272, IGSF6, METTL9*
chr14:9930000_9960000	30	*xp-ehh*	low	237 - 271	3.07 - 3.27	389 - 418	0.14 - 0.17	2.94 - 3.66	*ENSGALG00000047816 (LncRNA), ENSGALG00000050138 (LncRNA)*
chr15:6790000_6810000	20	*xp-ehh*	high	198 - 198	3.31 - 3.31	332 - 332	0.05 - 0.05	0.58 - 0.58	*FICD, SART3*
chr15:10780000_10800000	20	*xp-ehh*	low	142 - 142	3.03 - 3.03	236 - 236	0.09 - 0.09	1.74 - 1.74	
chr17:8280000_8300000	20	*xp-ehh*	low	181 - 181	3.1 - 3.1	374 - 374	0.03 - 0.03	0.11 - 0.11	*ENSGALG00000050011 (LncRNA), ENSGALG00000053422*
chr17:8460000_8480000	20	*xp-ehh*	low	172 - 172	3.24 - 3.24	292 - 292	0.08 - 0.08	1.36 - 1.36	*CAMSAP1, gga-mir-1465, UBAC1*
chr17:9440000_9520000	80	*fst*	low	42 - 72	0.38 - 0.88	62 - 91	0.24 - 0.36	5.51 - 8.6	*DENND1A, ENSGALG00000049802 (LncRNA), ENSGALG00000026417 (ncRNA), ENSGALG00000028552 (ncRNA)*
chr18:9860000_9900000	40	*xp-ehh*	high	120 - 128	3.2 - 3.61	203 - 214	0.07 - 0.14	1.2 - 3.04	*AFMID, ARL16, ENSGALG00000019795 (LncRNA), ENSGALG00000038671, GCGR, HGS, MRPL12, SLC25A10, SYNGR2, TK1, TMC6*
chr19:940000_960000	20	*fst*	high	202 - 202	0.72 - 0.72	314 - 314	0.23 - 0.23	5.38 - 5.38	*ELN*
chr19:1810000_1850000	40	*xp-ehh*	low	81 - 135	3.09 - 3.53	227 - 271	0.18 - 0.21	3.85 - 4.82	*AUTS2*
chr19:7450000_7480000	30	*xp-ehh*	high	249 - 267	3.13 - 3.48	325 - 365	0.06 - 0.07	0.96 - 1.2	*CLTC, PTRH2, VMP1*
chr20:1090000_1110000	20	*fst*	high	228 - 228	1.05 - 1.05	297 - 297	0.23 - 0.23	5.31 - 5.31	*CPNE1, RBM12*
chr21:530000_570000	40	*xp-ehh*	high	131 - 201	3.25 - 3.4	253 - 318	0.19 - 0.21	4.13 - 4.7	*DNAJC11, ENSGALG00000033569, KLHL21, PHF13*
chr22:2620000_2650000	30	*xp-ehh*	high	26 - 40	3.48 - 3.57	73 - 113	0.05 - 0.11	0.6 - 2.02	*FGFR1, LETM2, NSD3*
chr23:4350000_4400000	50	*xp-ehh*	low	104 - 213	3.3 - 3.63	220 - 343	0.12 - 0.15	2.32 - 3.3	*AGO1, AGO4, C1orf216, CLSPN*

chr26:3940000_3960000	20	*fst*	low	226 - 226	1.56 - 1.56	339 - 339	0.22 - 0.22	5.03 - 5.03	*ENSGALG00000002170,* *ENSGALG00000054224*
chr27:4280000_4330000	50	*xp-ehh*	high	222 - 240	3.04 - 3.29	337 - 380	0.05 - 0.09	0.64 - 1.66	
chr28:1890000_1910000	20	*xp-ehh*	high	157 - 157	3.11 - 3.11	305 - 305	0.13 - 0.13	2.64 - 2.64	*ADAMTS10*
chr33:5260000_5280000	20	*xp-ehh*	high	180 - 180	3.16 - 3.16	298 - 298	0.03 - 0.03	0.02 - 0.02	*ENSGALG00000043689, KRT5*

Table S4.2e. Candidate selection signature regions and overlapping genes detected from analyses of the Soil organic carbon content at 0.00 depth (SoilOrgC). (**SSR_ID:** selection signature regions ID as chromosome:start_stop position. **DD:** directionality of selection)

SSR_ID	**size (kb)**	**detection method**	**DD**	**no. of SNPs per *XP-EHH* window (range)**	**absolute value of mean *XP-EHH*_std (range)**	**no. of SNPs per *Fst* window (range)**	**weighted *Fst* per window (range)**	**Z*Fst* per window (range)**	**Overlapping genes**
chr1:5410000_5440000	30	*fst*	high	102 - 126	1.03 - 1.38	163 - 192	0.24 - 0.26	5.11 - 5.72	
chr1:26090000_26120000	30	*xp-ehh*	high	170 - 200	3.29 - 3.39	234 - 303	0.18 - 0.19	3.69 - 3.97	*ENSGALG00000053114, FOXP2*
chr1:26680000_26730000	50	*xp-ehh*	high	125 - 185	3.06 - 3.15	255 - 344	0.11 - 0.15	1.84 - 2.98	
chr1:44930000_44960000	30	*fst*	high	109 - 143	1.65 - 1.95	223 - 246	0.25 - 0.28	5.32 - 6.07	*CRADD*
chr1:55980000_56050000	70	*xp-ehh*	high	159 - 231	3.17 - 3.78	239 - 339	0.09 - 0.12	1.51 - 2.08	*HIPK2, TBXAS1*
chr1:56150000_56170000	20	*xp-ehh*	high	240 - 240	3 - 3	505 - 505	0.11 - 0.11	1.79 - 1.79	
chr1:60370000_60400000	30	*xp-ehh*	high	156 - 203	3.22 - 3.27	327 - 345	0.12 - 0.13	2.11 - 2.42	*NINJ2*
chr1:60810000_60870000	60	*xp-ehh*	high	151 - 208	3.01 - 3.33	271 - 419	0.16 - 0.19	3.13 - 3.88	*ENSGALG00000049382 (LncRNA),* *ERC1, WNT5B*
chr1:64890000_64910000	20	*fst*	low	156 - 156	1.38 - 1.38	281 - 281	0.24 - 0.24	5.17 - 5.17	*PDE3A*
chr1:78040000_78060000	20	*fst*	high	102 - 102	0.33 - 0.33	233 - 233	0.24 - 0.24	5.06 - 5.06	*C1R, C1S*
chr1:78050000_78080000	30	*fst*	low	102 - 106	0.62 - 0.73	218 - 227	0.26 - 0.26	5.54 - 5.56	*C1R, C1S, CLSTN3, RBP5*
chr1:93700000_93740000	40	*xp-ehh*	high	210 - 278	3.25 - 3.32	397 - 496	0.09 - 0.16	1.33 - 3.15	*ENSGALG00000052608 (LncRNA),* *POGLUT1, TIMMDC1*
chr1:96710000_96740000	30	*fst*	high	40 - 51	0.82 - 0.82	105 - 125	0.26 - 0.27	5.7 - 5.94	*ENSGALG00000053960 (LncRNA)*
chr1:105380000_105420000	40	*xp-ehh*	high	177 - 191	3.1 - 3.51	303 - 332	0.13 - 0.17	2.37 - 3.37	*ENSGALG00000015822,* *ENSGALG00000036613 (LncRNA)*
chr1:108010000_108030000	20	*fst*	high	183 - 183	0.5 - 0.5	353 - 353	0.24 - 0.24	5.11 - 5.11	*CHAF1B*

chr1:110490000_110530000	40	*fst*	high	191 - 232	1.54 - 2.26	274 - 322	0.25 - 0.28	5.46 - 6.19	*ENSGALG00000052523*
chr1:111070000_111110000	40	*xp-ehh*	high	143 - 219	3.16 - 3.93	397 - 546	0.1 - 0.2	1.72 - 4.27	*CRYAA*
chr1:115950000_115970000	20	*fst*	low	265 - 265	1.12 - 1.12	397 - 397	0.26 - 0.26	5.72 - 5.72	
chr1:117620000_117640000	20	*fst*	high	92 - 92	1.26 - 1.26	153 - 153	0.25 - 0.25	5.28 - 5.28	*IL1RAPL1*
chr1:117650000_117700000	50	*fst*	high	80 - 165	0.48 - 0.69	164 - 235	0.25 - 0.28	5.44 - 6.12	*IL1RAPL1*
chr1:117750000_117780000	30	*fst*	high	126 - 137	0.53 - 0.66	182 - 197	0.25 - 0.26	5.5 - 5.64	*IL1RAPL1*
chr1:118420000_118440000	20	*fst*	high	84 - 84	2.5 - 2.5	147 - 147	0.24 - 0.24	5.11 - 5.11	
chr1:146480000_146500000	20	*fst*	high	139 - 139	1.84 - 1.84	349 - 349	0.27 - 0.27	5.86 - 5.86	*FARP1*
chr1:167320000_167350000	30	*xp-ehh*	high	227 - 265	3.16 - 3.22	371 - 377	0.14 - 0.15	2.67 - 2.82	*VWA8*
chr1:167480000_167540000	60	*xp-ehh*	high	163 - 281	3.1 - 3.18	373 - 437	0.05 - 0.12	0.42 - 2.19	*ENSGALG00000017789 (5S_rRNA), VWA8*
chr1:168920000_168990000	70	*xp-ehh* / *fst*	high	216 - 264	2.42 - 3.77	357 - 425	0.11 - 0.26	1.79 - 5.72	*GPALPP1, GTF2F2, NUFIP1, ENSGALG00000025645 (SnRNA)*
chr1:171050000_171090000	40	*fst*	high	119 - 125	0.81 - 1.18	185 - 218	0.28 - 0.28	6.03 - 6.21	*ENSGALG00000048894 (LncRNA)*
chr1:171730000_171760000	30	*xp-ehh*	high	222 - 239	3.14 - 3.19	408 - 423	0.15 - 0.19	2.89 - 3.82	*CKAP2, VPS36*
chr1:173800000_173820000	20	*fst*	high	196 - 196	1.13 - 1.13	343 - 343	0.24 - 0.24	5.03 - 5.03	*ENSGALG00000051912 (LncRNA)*
chr1:174170000_174240000	70	*fst*	high	135 - 232	0.31 - 1.4	245 - 422	0.26 - 0.45	5.69 - 10.36	*DCLK1*
chr1:191500000_191540000	40	*xp-ehh*	low	236 - 267	3.02 - 3.35	520 - 578	0.06 - 0.13	0.56 - 2.51	*ENSGALG00000014066, ENSGALG00000049171*
chr1:193630000_193650000	20	*xp-ehh*	low	131 - 131	3.43 - 3.43	329 - 329	0.12 - 0.12	2.21 - 2.21	*TENM4*
chr1:197230000_197310000	80	*xp-ehh* and/or *fst*	high	105 - 332	1.87 - 3.25	153 - 521	0.18 - 0.3	3.64 - 6.61	*ENSGALG00000052767, ENSGALG00000053527, ENSGALG00000053623, ENSGALG00000054014, HBBA, HBBR, HBE, HBE1, OR51M1,OR52R1*
chr1:197480000_197580000	100	*xp-ehh*	high	73 - 198	3.04 - 3.76	145 - 338	0.09 - 0.16	1.36 - 3.11	*ENSGALG00000017334, ENSGALG00000033819, gga-mir-1600, LCMT2, TPP1*
chr2:5480000_5500000	20	*fst*	high	185 - 185	0.51 - 0.51	323 - 323	0.26 - 0.26	5.57 - 5.57	*ENSGALG00000006072*
chr2:5870000_5900000	30	*fst*	high	200 - 202	1.8 - 2.23	356 - 360	0.24 - 0.25	5.13 - 5.45	*ACVR2B, gga-mir-6600*
chr2:6110000_6130000	20	*fst*	high	197 - 197	1.03 - 1.03	368 - 368	0.24 - 0.24	5.21 - 5.21	*ENSGALG00000053252 (LncRNA)*
chr2:9780000_9800000	20	*fst*	high	293 - 293	2 - 2	460 - 460	0.24 - 0.24	5.06 - 5.06	*ENSGALG00000048545 (LncRNA), ENSGALG00000053524 (LncRNA)*
chr2:17220000_17240000	20	*fst*	high	241 - 241	1.68 - 1.68	474 - 474	0.23 - 0.23	5.02 - 5.02	*KIAA1217*
chr2:56220000_56250000	30	*xp-ehh*	high	190 - 270	3.15 - 3.44	333 - 416	0.11 - 0.16	2.03 - 3.14	*PQLC1*

chr2:68510000_68530000	20	*fst*	high	269 - 269	0.53 - 0.53	429 - 429	0.24 - 0.24	5.05 - 5.05	
chr2:89020000_89040000	20	*xp-ehh*	low	165 - 165	3.22 - 3.22	233 - 233	0.04 - 0.04	0.27 - 0.27	*AHRR*
chr2:91690000_91710000	20	*xp-ehh*	low	115 - 115	3.01 - 3.01	199 - 199	0.11 - 0.11	1.84 - 1.84	*ENSGALG00000033168, ZNF407*
chr2:104780000_104820000	40	*xp-ehh*	low	146 - 196	3.06 - 3.23	204 - 253	0.06 - 0.1	0.65 - 1.67	*CDH2*
chr2:114100000_114150000	50	*xp-ehh*	low	91 - 128	3.1 - 3.38	169 - 193	0.05 - 0.11	0.52 - 1.93	*ENSGALG00000048169 (LncRNA)*
chr2:124370000_124420000	50	*xp-ehh*	high	194 - 214	3.21 - 3.58	337 - 406	0.1 - 0.14	1.62 - 2.66	
chr2:124450000_124500000	50	*xp-ehh*	high	123 - 199	3.01 - 3.23	381 - 465	0.12 - 0.14	2.08 - 2.76	*ENSGALG00000053987 (LncRNA)*
chr2:126300000_126330000	30	*xp-ehh*	high	179 - 187	3 - 3.17	345 - 353	0.1 - 0.11	1.66 - 1.84	*ENSGALG00000054155 (LncRNA), NDUFAF6, TP53INP1*
chr2:126380000_126450000	70	*xp-ehh*	high	205 - 273	3.19 - 3.58	305 - 381	0.07 - 0.15	0.96 - 3.03	*C8orf37*
chr2:128500000_128530000	30	*xp-ehh*	high	165 - 169	3.03 - 3.2	330 - 334	0.17 - 0.17	3.3 - 3.49	*PABPC1*
chr2:128540000_128560000	20	*xp-ehh*	high	276 - 276	3.04 - 3.04	390 - 390	0.08 - 0.08	1.15 - 1.15	*YWHAZ*
chr2:128600000_128680000	80	*xp-ehh*	high	160 - 246	3.08 - 3.25	292 - 413	0.11 - 0.15	1.89 - 3.03	*ENSGALG00000049820 (LncRNA), ZNF706*
chr2:137920000_137940000	20	*xp-ehh*	low	89 - 89	3.18 - 3.18	299 - 299	0.15 - 0.15	2.92 - 2.92	*TBC1D31*
chr2:147030000_147090000	60	*xp-ehh*	low	35 - 86	3.35 - 3.63	63 - 218	0.05 - 0.09	0.51 - 1.34	*TSNARE1*
chr2:147220000_147260000	40	*fst*	low	82 - 166	0.77 - 1.03	95 - 219	0.24 - 0.31	5.05 - 6.77	*TSNARE1*
chr3:28290000_28330000	40	*fst*	high	119 - 215	1.36 - 1.93	247 - 326	0.25 - 0.28	5.28 - 6.27	
chr3:28340000_28360000	20	*fst*	high	118 - 118	2.14 - 2.14	185 - 185	0.24 - 0.24	5.23 - 5.23	*LRFN2*
chr3:32790000_32870000	80	*xp-ehh*	low	127 - 166	3.05 - 3.38	232 - 329	0.1 - 0.16	1.61 - 3.16	*ENSGALG00000049188 (LncRNA), ENSGALG00000050560 (LncRNA)*
chr3:35160000_35180000	20	*fst*	high	164 - 164	0.34 - 0.34	270 - 270	0.31 - 0.31	6.89 - 6.89	*AKT3*
chr3:48750000_48770000	20	*xp-ehh*	low	265 - 265	3.01 - 3.01	378 - 378	0.01 - 0.01	-0.47 - -0.47	*MTHFD1L*
chr3:59420000_59440000	20	*fst*	high	91 - 91	1.37 - 1.37	201 - 201	0.26 - 0.26	5.58 - 5.58	
chr3:66570000_66630000	60	*xp-ehh*	high	223 - 289	3.02 - 3.9	338 - 502	0.1 - 0.16	1.66 - 3.12	*CDC40, METTL24*
chr3:67200000_67300000	100	*xp-ehh* and/or *fst*	high	224 - 286	3.02 - 3.54	335 - 388	0.16 - 0.24	3.23 - 5.09	*ARMC2, FOXO3*
chr3:86680000_86720000	40	*xp-ehh*	low	251 - 284	3.33 - 3.58	409 - 492	0.09 - 0.18	1.41 - 3.6	*PRIM2*
chr3:87480000_87500000	20	*fst*	high	225 - 225	0.56 - 0.56	329 - 329	0.24 - 0.24	5.07 - 5.07	*HMGCLL1*
chr3:87490000_87510000	20	*fst*	low	182 - 182	0.58 - 0.58	275 - 275	0.25 - 0.25	5.5 - 5.5	*HMGCLL1*
chr3:91850000_91870000	20	*xp-ehh*	high	220 - 220	3.29 - 3.29	311 - 311	0.15 - 0.15	2.79 - 2.79	*ENSGALG00000050914 (LncRNA), ENSGALG00000025589 (SnRNA)*
chr3:106440000_106520000	80	*fst*	high	198 - 266	2.11 - 2.71	319 - 394	0.25 - 0.33	5.43 - 7.4	*MSRA*
chr3:106820000_106850000	30	*xp-ehh*	high	150 - 157	3.01 - 3.13	279 - 283	0.14 - 0.14	2.57 - 2.58	*PINX1*

chr3:107760000_107820000	60	*fst*	high	103 - 185	1.5 - 2.18	259 - 354	0.24 - 0.3	5.07 - 6.57	*ENSGALG00000050553,* *ENSGALG00000050753*
chr3:108140000_108180000	40	*xp-ehh*	high	140 - 217	3.04 - 3.23	228 - 303	0.12 - 0.15	2.28 - 2.95	*ENSGALG00000047869 (LncRNA)*
chr3:108200000_108250000	50	*xp-ehh*	high	86 - 196	3.03 - 3.58	212 - 344	0.13 - 0.19	2.47 - 3.8	*ENSGALG00000047589, TFAP2B*
chr4:6480000_6500000	20	*fst*	high	111 - 111	0.71 - 0.71	192 - 192	0.28 - 0.28	6.23 - 6.23	*ENSGALG00000047994 (LncRNA)*
chr4:13110000_13130000	20	*xp-ehh*	high	253 - 253	3.26 - 3.26	395 - 395	0.12 - 0.12	2.17 - 2.17	*TRPC5*
chr4:14070000_14100000	30	*xp-ehh*	high	220 - 224	3.09 - 3.11	396 - 397	0.12 - 0.13	2.27 - 2.37	*ENSGALG00000048179*
chr4:20300000_20340000	40	*fst*	low	54 - 77	0.16 - 0.21	67 - 113	0.26 - 0.32	5.57 - 7.02	*FGA, FGB, FGG, PLRG1*
chr4:21060000_21080000	20	*fst*	high	93 - 93	1.33 - 1.33	193 - 193	0.26 - 0.26	5.55 - 5.55	*ENSGALG00000049888 (LncRNA)*
chr4:21910000_21930000	20	*xp-ehh*	low	82 - 82	3.11 - 3.11	166 - 166	0.04 - 0.04	0.26 - 0.26	*C4H4ORF46, ETFDH, RXFP1*
chr4:40090000_40110000	20	*xp-ehh*	high	132 - 132	3.04 - 3.04	200 - 200	0.14 - 0.14	2.65 - 2.65	*TENM3*
chr4:40120000_40140000	20	*xp-ehh*	high	146 - 146	3.01 - 3.01	225 - 225	0.19 - 0.19	3.81 - 3.81	*TENM3*
chr4:40250000_40280000	30	*fst*	high	167 - 171	2.29 - 2.35	279 - 309	0.25 - 0.26	5.38 - 5.76	*TENM3*
chr4:40470000_40510000	40	*fst*	high	69 - 114	2.63 - 2.65	202 - 241	0.24 - 0.27	5.13 - 6.01	*TENM3*
chr4:40760000_40860000	100	*fst*	high	96 - 204	2.34 - 2.8	146 - 355	0.24 - 0.33	5.09 - 7.42	*ENSGALG00000050037 (LncRNA)*
chr4:40880000_40910000	30	*fst*	high	67 - 76	2.52 - 2.56	125 - 143	0.24 - 0.3	5.05 - 6.53	*ENSGALG00000050037 (LncRNA)*
chr4:40930000_40980000	50	*fst*	high	109 - 180	1.21 - 1.5	234 - 278	0.3 - 0.36	6.58 - 8.21	*ENSGALG00000050037 (LncRNA)*
chr4:42050000_42070000	20	*fst*	high	20 - 20	1.86 - 1.86	38 - 38	0.26 - 0.26	5.68 - 5.68	*ENSGALG00000047221 (LncRNA)*
chr4:42130000_42150000	20	*fst*	high	108 - 108	1.87 - 1.87	170 - 170	0.25 - 0.25	5.38 - 5.38	*ENSGALG00000054340 (LncRNA)*
chr4:42900000_42950000	50	*fst*	high	67 - 103	1.98 - 2.21	166 - 216	0.24 - 0.3	5.09 - 6.54	*GALNTL6*
chr4:43220000_43300000	80	*xp-ehh*	high	164 - 265	3.02 - 4.03	281 - 384	0.14 - 0.2	2.61 - 4.21	*ENSGALG00000046857 (LncRNA),* *GALNT7, HMGB2*
chr4:43390000_43420000	30	*fst*	high	103 - 128	2.38 - 2.39	137 - 198	0.24 - 0.27	5.15 - 5.99	*ENSGALG00000033311 (LncRNA),* *ENSGALG00000052572 (LncRNA)*
chr4:46290000_46310000	20	*fst*	high	207 - 207	1 - 1	386 - 386	0.24 - 0.24	5.22 - 5.22	*LIN54*
chr4:46450000_46470000	20	*fst*	high	200 - 200	2.24 - 2.24	395 - 395	0.26 - 0.26	5.68 - 5.68	*ABRAXAS1, HELQ, MRPS18C*
chr4:48290000_48310000	20	*fst*	high	173 - 173	0.53 - 0.53	248 - 248	0.32 - 0.32	7.26 - 7.26	*ADGRL3*
chr4:48300000_48330000	30	*fst*	low	183 - 185	0.5 - 0.53	285 - 296	0.28 - 0.35	6.19 - 8	*ADGRL3*
chr4:50860000_50970000	110	*fst*	high	117 - 201	0.6 - 0.77	238 - 347	0.24 - 0.34	5.05 - 7.73	*IL8, IL8L1, PPBP, RASSF6*
chr4:66820000_66840000	20	*xp-ehh*	low	281 - 281	3.19 - 3.19	479 - 479	0.08 - 0.08	1.29 - 1.29	*ENSGALG00000014200*
chr4:75670000_75710000	40	*xp-ehh*	high	143 - 176	3.06 - 3.17	227 - 282	0.13 - 0.19	2.45 - 3.79	*ENSGALG00000047307 (LncRNA),* *ENSGALG00000052764 (LncRNA)*
chr4:76500000_76560000	60	*xp-ehh*	high	238 - 305	3.05 - 3.57	366 - 491	0.12 - 0.14	2.05 - 2.54	*Prom1*

chr4:85560000_85580000	20	*xp-ehh*	high	364 - 364	3.01 - 3.01	565 - 565	0.06 - 0.06	0.59 - 0.59	*ENSGALG00000052751 (LncRNA), KDM3A*
chr5:1190000_1210000	20	*xp-ehh*	high	279 - 279	3.17 - 3.17	433 - 433	0.14 - 0.14	2.62 - 2.62	*gga-mir-1808, RMDN3*
chr5:1860000_1880000	20	*fst*	low	233 - 233	2.35 - 2.35	409 - 409	0.24 - 0.24	5.08 - 5.08	*BET1L, NADSYN1, RIC8A, VPS51*
chr5:13530000_13570000	40	*xp-ehh*	high	201 - 205	3.01 - 3.26	327 - 342	0.06 - 0.06	0.75 - 0.77	*KCNQ1*
chr5:17550000_17580000	30	*xp-ehh*	high	114 - 135	3.08 - 3.11	259 - 272	0.14 - 0.15	2.66 - 2.91	*ENSGALG00000051988 (LncRNA), ENSGALG00000027034 (ncRNA)*
chr5:48350000_48370000	20	*xp-ehh*	low	235 - 235	3.12 - 3.12	348 - 348	0.13 - 0.13	2.31 - 2.31	*CYP46A1, ENSGALG00000036286*
chr6:17050000_17090000	40	*xp-ehh*	low	322 - 334	3.01 - 3.47	478 - 526	0.04 - 0.09	0.07 - 1.38	*EIF3F, ENSGALG00000051819*
chr6:25320000_25350000	30	*fst*	high	254 - 283	1.8 - 1.96	409 - 435	0.24 - 0.26	5.08 - 5.59	*ENSGALG00000050593 (LncRNA)*
chr6:26870000_26890000	20	*fst*	high	181 - 181	0.63 - 0.63	308 - 308	0.24 - 0.24	5.22 - 5.22	*ENSGALG00000054720 (LncRNA), ENSGALG00000054969 (LncRNA)*
chr7:29270000_29290000	20	*fst*	high	97 - 97	1.23 - 1.23	245 - 245	0.27 - 0.27	5.8 - 5.8	*DPP10*
chr7:31520000_31560000	40	*xp-ehh*	high	137 - 212	3.03 - 3.39	298 - 322	0.19 - 0.22	3.98 - 4.67	*ENSGALG00000048656*
chr7:33280000_33310000	30	*xp-ehh*	low	156 - 174	3.39 - 3.47	312 - 324	0.08 - 0.11	1.06 - 1.91	*ENSGALG00000053032 (LncRNA), ZEB2*
chr7:33900000_33920000	20	*fst*	high	155 - 155	2.37 - 2.37	289 - 289	0.24 - 0.24	5.24 - 5.24	*ENSGALG00000041782 (LncRNA)*
chr7:35150000_35170000	20	*xp-ehh*	low	216 - 216	3.51 - 3.51	404 - 404	0.11 - 0.11	1.82 - 1.82	*RIF1*
chr7:35350000_35410000	60	*fst*	low	188 - 282	1.24 - 2.47	362 - 467	0.24 - 0.34	5.25 - 7.55	*CACNB4, STAM2*
chr7:36110000_36130000	20	*fst*	high	200 - 200	0.29 - 0.29	382 - 382	0.26 - 0.26	5.7 - 5.7	
chr8:8120000_8140000	20	*xp-ehh*	high	321 - 321	3.13 - 3.13	423 - 423	0.05 - 0.05	0.52 - 0.52	*COLGALT2*
chr8:8160000_8180000	20	*xp-ehh*	high	176 - 176	3.09 - 3.09	345 - 345	0.15 - 0.15	2.95 - 2.95	
chr8:25180000_25200000	20	*xp-ehh*	low	223 - 223	3.16 - 3.16	342 - 342	0.08 - 0.08	1.16 - 1.16	
chr9:10680000_10710000	30	*fst*	low	163 - 206	0.67 - 0.67	255 - 326	0.26 - 0.27	5.73 - 5.8	*PLS1, TRPC1*
chr9:14300000_14330000	30	*xp-ehh*	high	195 - 205	3.01 - 3.13	381 - 436	0.04 - 0.05	0.24 - 0.44	*ENSGALG00000047059 (LncRNA), ENSGALG00000052948 (LncRNA)*
chr9:18740000_18780000	40	*xp-ehh*	high	153 - 161	3.03 - 3.21	277 - 303	0.08 - 0.14	1.04 - 2.78	*NAALADL2*
chr9:18810000_18870000	60	*xp-ehh*	high	141 - 213	3.15 - 3.59	280 - 337	0.11 - 0.18	1.97 - 3.54	*ENSGALG00000048975 (LncRNA), NAALADL2*
chr9:18990000_19010000	20	*xp-ehh*	high	155 - 155	3.19 - 3.19	301 - 301	0.1 - 0.1	1.78 - 1.78	*NLGN1*
chr10:1520000_1580000	60	*xp-ehh*	high	125 - 179	3.06 - 3.28	299 - 337	0.11 - 0.16	1.94 - 3.13	*DUOX2, RBPMS2, ENSGALG00000036706 (ncRNA), SORD*
chr10:6490000_6520000	30	*fst*	high	57 - 105	2.2 - 2.28	170 - 250	0.25 - 0.29	5.41 - 6.5	*TJP1*

chr10:6820000_6870000	50	*xp-ehh* *fst*	high	153 - 183	3.29 - 3.37	254 - 299	0.24 - 0.31	5.25 - 6.82	*ENSGALG00000053176 (MiRNA), THSD4*
chr10:13660000_13690000	30	*xp-ehh*	high	213 - 261	3.21 - 3.32	414 - 503	0.19 - 0.22	3.97 - 4.69	*NTRK3*
chr10:20310000_20340000	30	*xp-ehh*	high	198 - 231	3.21 - 3.33	369 - 423	0.11 - 0.12	1.83 - 2.07	*FRMD5, WDR76*
chr11:13300000_13330000	30	*fst*	low	211 - 259	0.76 - 0.92	264 - 307	0.24 - 0.24	5.1 - 5.11	
chr11:15360000_15390000	30	*xp-ehh*	high	220 - 288	3.16 - 3.26	372 - 446	0.1 - 0.14	1.69 - 2.64	*ATMIN, CENPN, CMC2*
chr13:18270000_18290000	20	*xp-ehh*	high	116 - 116	3.05 - 3.05	191 - 191	0.04 - 0.04	0.07 - 0.07	*KCTD16*
chr14:4740000_4760000	20	*fst*	high	164 - 164	0.98 - 0.98	354 - 354	0.24 - 0.24	5.1 - 5.1	*ARPC1A*
chr14:7480000_7510000	30	*xp-ehh*	high	161 - 197	3.07 - 3.16	268 - 343	0.14 - 0.14	2.57 - 2.72	*HS3ST4*
chr14:13190000_13210000	20	*fst*	low	433 - 433	2.19 - 2.19	553 - 553	0.24 - 0.24	5.21 - 5.21	*CLUAP1*
chr14:15390000_15410000	20	*fst*	high	29 - 29	1.57 - 1.57	78 - 78	0.24 - 0.24	5.26 - 5.26	*ENSGALG00000041298, gga-mir-12225, UQCRC2*
chr15:7220000_7250000	30	*xp-ehh*	high	201 - 204	3.07 - 3.21	362 - 366	0.09 - 0.09	1.41 - 1.42	*KIAA1671, SGSM1*
chr15:9890000_9920000	30	*xp-ehh*	low	189 - 207	3.05 - 3.07	312 - 329	0.07 - 0.09	1 - 1.5	*CIT*
chr17:3160000_3180000	20	*xp-ehh*	low	219 - 219	3.12 - 3.12	360 - 360	0.01 - 0.01	-0.46 - -0.46	*ENSGALG00000050469 (LncRNA)*
chr17:5520000_5540000	20	*fst*	high	196 - 196	2.26 - 2.26	389 - 389	0.23 - 0.23	5.02 - 5.02	*SPTAN1*
chr18:850000_890000	40	*xp-ehh*	high	166 - 172	3.1 - 3.17	331 - 333	0.08 - 0.1	1.15 - 1.56	*MYOCD*
chr18:1360000_1390000	30	*xp-ehh*	high	151 - 197	3.11 - 3.11	310 - 392	0.14 - 0.15	2.7 - 2.91	
chr18:6530000_6610000	80	*xp-ehh*	high	172 - 291	3.05 - 3.16	235 - 422	0.1 - 0.13	1.57 - 2.36	*CRLF3, ENSGALG00000047593 (LncRNA), ENSGALG00000051036 (LncRNA), UTP6*
chr18:7220000_7240000	20	*xp-ehh*	high	164 - 164	3.02 - 3.02	284 - 284	0.1 - 0.1	1.55 - 1.55	*CACNG1, CACNG4*
chr18:9840000_9910000	70	*xp-ehh*	low	97 - 146	3.04 - 3.7	202 - 309	0.11 - 0.16	1.89 - 3.23	*AFMID, ARL16, ENSGALG00000019795 (LncRNA), ENSGALG00000027822, ENSGALG00000038671, GCGR, HGS, MRPL12, SLC25A10, SYNGR2, TK1, TMC6*
chr20:1570000_1590000	20	*fst*	high	147 - 147	1.06 - 1.06	198 - 198	0.25 - 0.25	5.4 - 5.4	*ASIP*
chr33:6470000_6500000	30	*fst*	low	4 - 5	0.76 - 1.38	12 - 22	0.4 - 0.5	9.14 - 11.7	*DRAP1, KAT5, UQCC3, RELA*

Table S4.2f. Candidate selection signature regions and overlapping genes detected from analyses of the LandUse. (**SSR_ID:** selection signature regions ID as chromosome:start_stop position. **DD:** directionality of selection).

SSR_ID	**size (kb)**	**detection method**	**DD**	**no. of SNPs per *XP-EHH* window (range)**	**absolute value of mean *XP-EHH*_std (range)**	**no. of SNPs per *Fst* window (range)**	**weighted *Fst* per window (range)**	***ZFst* per window (range)**	**Overlapping genes**
chr1:50910000_50930000	20	*xp-ehh*	low	143 - 143	3.22 - 3.22	266 - 266	0.06 - 0.06	1.23 - 1.23	*CSNK1E, ENSGALG00000053778, KCNJ4*
chr1:55450000_55470000	20	*fst*	high	107 - 107	1.5 - 1.5	339 - 339	0.23 - 0.23	6.26 - 6.26	*WASHC3*
chr1:60930000_60980000	50	*xp-ehh*	high	177 - 235	3.3 - 4.65	319 - 373	0.06 - 0.07	1.15 - 1.5	*ADIPOR2, CHUNK-1 ENSGALG00000036964*
chr1:64470000_64500000	30	*fst*	high	197 - 226	1.62 - 1.66	278 - 355	0.19 - 0.25	5.05 - 6.91	*PLEKHA5*
chr1:65410000_65430000	20	*fst*	low	171 - 171	2.93 - 2.93	285 - 285	0.19 - 0.19	5.18 - 5.18	*BCAT1, ENSGALG00000048674 (LncRNA)*
chr1:90440000_90540000	100	*xp-ehh*	low	163 - 280	3.24 - 3.94	253 - 418	0.08 - 0.15	1.71 - 3.81	*EPHA3*
chr1:92310000_92360000	50	*xp-ehh*	low	142 - 166	3.22 - 3.47	416 - 477	0.11 - 0.12	2.78 - 2.94	*ENSGALG00000049116 (LncRNA)*
chr1:97580000_97610000	30	*fst*	low	96 - 126	1.43 - 1.46	187 - 224	0.19 - 0.21	5.09 - 5.6	*ROBO1*
chr1:97640000_97660000	20	*fst*	low	71 - 71	1.98 - 1.98	156 - 156	0.21 - 0.21	5.79 - 5.79	*ROBO1*
chr1:104680000_104700000	20	*fst*	low	60 - 60	1.46 - 1.46	119 - 119	0.19 - 0.19	5.03 - 5.03	
chr1:110150000_110180000	30	*xp-ehh*	low	171 - 200	3.11 - 3.21	295 - 330	0.1 - 0.11	2.33 - 2.68	*BACE2, ENSGALG0000005055 (LncRNA)*
chr1:110220000_110270000	50	*xp-ehh*	low	133 - 166	3.18 - 3.53	259 - 308	0.1 - 0.13	2.25 - 3.23	*FAM3B, MX1, TMPRSS2*
chr1:110360000_110400000	40	*xp-ehh*	low	184 - 207	3.02 - 3.47	356 - 397	0.07 - 0.15	1.43 - 3.84	*ENSGALG00000051940 (LncRNA), PRDM15, RIPK4*
chr1:112400000_112430000	30	*fst*	low	10 - 10	2.84 - 2.84	361 - 403	0.2 - 0.21	5.31 - 5.61	*EFHC2*
chr1:112490000_112520000	30	*xp-ehh*	low	112 - 154	3.03 - 3.21	257 - 310	0.05 - 0.1	0.93 - 2.46	*NDP*
chr1:112830000_112850000	20	*fst*	low	63 - 63	2.31 - 2.31	191 - 191	0.21 - 0.21	5.73 - 5.73	
chr1:112870000_112890000	20	*fst*	low	119 - 119	2.51 - 2.51	202 - 202	0.19 - 0.19	5.01 - 5.01	
chr1:113570000_113610000	40	*xp-ehh*	low	192 - 295	3.06 - 3.24	352 - 424	0.07 - 0.1	1.49 - 2.28	*ATP6AP2, ENSGALG00000032773 (LncRNA), ENSGALG00000053145 (LncRNA)*
chr1:113630000_113700000	70	*xp-ehh*	low	47 - 190	3.01 - 3.36	139 - 354	0.06 - 0.13	1.1 - 3.15	*BCOR, ENSGALG00000046706 (LncRNA)*

chr1:115280000_115300000	20	*fst*	low	117 - 117	2.46 - 2.46	196 - 196	0.19 - 0.19	5.16 - 5.16	*TMEM47*
chr1:126420000_126450000	30	*xp-ehh*	low	218 - 274	3.08 - 3.11	356 - 441	0.02 - 0.03	-0.02 - 0.32	*SHROOM2*
chr1:126750000_126770000	20	*xp-ehh*	low	224 - 224	3.06 - 3.06	332 - 332	0.12 - 0.12	2.86 - 2.86	
chr1:127170000_127470000	300	*xp-ehh* and/or *fst*	low	65 - 219	3 - 4.05	105 - 349	0.03 - 0.22	0.17 - 5.91	*ENSGALG00000048400 (LncRNA), gga-mir-7448*
chr1:127530000_127560000	30	*xp-ehh*	low	47 - 58	3.12 - 3.15	149 - 190	0.13 - 0.15	3.19 - 3.97	
chr1:127750000_127980000	230	*xp-ehh* and/or *fst*	low	21 - 108	2.22 - 3.48	81 - 203	0.09 - 0.24	2.1 - 7.34	*ENSGALG00000052885 (LncRNA), PUDP, STS*
chr1:128010000_128300000	290	*xp-ehh* and/or *fst*	low	21 - 85	2.21 - 3.37	74 - 158	0.04 - 0.32	0.46 - 9.05	*NLGN4Y, gga-mir-1397*
chr1:128350000_128470000	120	*xp-ehh* and/or *fst*	low	14 - 59	3.13 - 3.53	52 - 101	0.03 - 0.2	0.28 - 5.27	*ENSGALG00000016628*
chr1:128510000_128560000	50	*fst*	low	10 - 23	2.46 - 2.55	44 - 55	0.2 - 0.22	5.33 - 5.97	
chr1:128960000_129350000	390	*xp-ehh* and/or *fst*	low	7 - 159	3.03 - 4.59	88 - 224	0.02 - 0.29	0.09 - 8.14	*ENSGALG00000049563, ENSGALG00000052489*
chr1:130650000_130670000	20	*xp-ehh*	low	141 - 141	3.06 - 3.06	237 - 237	0.1 - 0.1	2.23 - 2.23	
chr1:132470000_132490000	20	*fst*	low	189 - 189	2.05 - 2.05	375 - 375	0.2 - 0.2	5.21 - 5.21	*GABRG3*
chr1:137220000_137250000	30	*fst*	low	233 - 251	1.96 - 2.05	327 - 341	0.19 - 0.2	5.14 - 5.38	*ENSGALG00000049876 (LncRNA), ENSGALG00000052777 (LncRNA)*
chr1:137260000_137310000	50	*xp-ehh*	low	166 - 211	3.03 - 3.87	251 - 348	0.11 - 0.13	2.51 - 3.3	*ENSGALG00000049876 (LncRNA), ENSGALG00000052777 (LncRNA)*
chr1:144290000_144320000	30	*fst*	low	113 - 138	2.73 - 2.86	303 - 336	0.19 - 0.2	5.19 - 5.25	*TPP2*
chr1:156070000_156100000	30	*fst*	high	129 - 162	1.74 - 2.01	273 - 309	0.2 - 0.22	5.48 - 5.84	
chr1:163380000_163430000	50	*fst*	low	124 - 160	0.12 - 0.18	208 - 268	0.22 - 0.31	6 - 8.65	*ENSGALG00000048185*
chr1:163440000_163520000	80	*fst*	low	76 - 183	0.78 - 1.76	147 - 277	0.19 - 0.27	5.08 - 7.42	*ENSGALG00000050889 (LncRNA)*
chr1:167560000_167600000	40	*xp-ehh*	high	224 - 246	3.33 - 3.71	352 - 383	0.03 - 0.05	0.31 - 0.84	*DGKH*
chr1:173090000_173120000	30	*fst*	low	250 - 294	0.13 - 0.18	352 - 396	0.19 - 0.2	5.05 - 5.38	*ENSGALG00000049925 (LncRNA)*
chr1:187380000_187400000	20	*fst*	high	246 - 246	1.04 - 1.04	492 - 492	0.19 - 0.19	5.14 - 5.14	*SLC36A4*
chr1:194700000_194720000	20	*fst*	high	258 - 258	1.06 - 1.06	394 - 394	0.21 - 0.21	5.67 - 5.67	*CAPN5*
chr2:3530000_3550000	20	*xp-ehh*	low	182 - 182	3.35 - 3.35	442 - 442	0.13 - 0.13	3.2 - 3.2	*MYL3*
chr2:3970000_4000000	30	*xp-ehh*	high	236 - 247	3.11 - 3.23	376 - 386	0.08 - 0.08	1.72 - 1.89	*SETD2*
chr2:10240000_10270000	30	*xp-ehh*	low	205 - 224	3.27 - 3.44	408 - 422	0.16 - 0.16	4.07 - 4.1	*DIP2C*
chr2:26130000_26150000	20	*fst*	low	145 - 145	0.76 - 0.76	222 - 222	0.22 - 0.22	5.86 - 5.86	*PHF14*
chr2:26190000_26220000	30	*fst*	low	89 - 98	1.5 - 1.69	115 - 160	0.19 - 0.2	5.07 - 5.25	*PHF14*
chr2:27820000_27840000	20	*fst*	low	48 - 48	1.36 - 1.36	111 - 111	0.19 - 0.19	5.19 - 5.19	*AGMO, ENSGALG00000051512*

									(LncRNA)
chr2:27870000_27900000	30	*fst*	low	104 - 110	1.19 - 1.41	207 - 213	0.19 - 0.2	5 - 5.26	*AGMO*
chr2:27930000_28030000	100	*fst*	low	27 - 154	1.2 - 1.76	145 - 323	0.21 - 0.41	5.73 - 11.75	*AGMO*
chr2:30120000_30140000	20	*fst*	low	203 - 203	1.65 - 1.65	327 - 327	0.19 - 0.19	5.14 - 5.14	*ENSGALG00000052719 (LncRNA)*
chr2:32130000_32250000	120	*xp-ehh*	high	137 - 362	3.58 - 4.55	243 - 550	0.09 - 0.15	2.13 - 3.74	*CBX3, HNRNPA2B1, SNX10*
chr2:38790000_38850000	60	*xp-ehh*	low	170 - 216	3.33 - 3.48	316 - 510	0.04 - 0.18	0.52 - 4.84	*AZI2*
chr2:38990000_39020000	30	*fst*	low	94 - 98	0.65 - 0.78	204 - 242	0.19 - 0.2	5.05 - 5.41	*RBMS3*
chr2:39780000_39810000	30	*xp-ehh*	low	153 - 207	3.02 - 3.1	288 - 338	0.07 - 0.09	1.32 - 2.14	*TGFBR2*
chr2:40530000_40600000	70	*xp-ehh*	low	21 - 261	3.05 - 3.74	42 - 406	0.06 - 0.16	1.17 - 4.17	*TCAIM*
chr2:44890000_44960000	70	*xp-ehh*	high	203 - 218	3 - 3.28	293 - 336	0.07 - 0.1	1.56 - 2.48	
chr2:45670000_45700000	30	*xp-ehh*	high	183 - 190	3.11 - 3.21	251 - 260	0.07 - 0.07	1.54 - 1.55	
chr2:47480000_47500000	20	*xp-ehh*	low	191 - 191	3.12 - 3.12	276 - 276	0.09 - 0.09	1.92 - 1.92	*BMPER*
chr2:54590000_54620000	30	*fst*	low	203 - 217	1.2 - 1.23	286 - 289	0.2 - 0.22	5.4 - 5.9	
chr2:54640000_54720000	80	*fst*	low	53 - 132	0.21 - 0.91	192 - 271	0.21 - 0.38	5.71 - 10.75	*ADCY1*
chr2:73300000_73320000	20	*fst*	low	25 - 25	0.62 - 0.62	105 - 105	0.2 - 0.2	5.46 - 5.46	*ENSGALG00000012941*
chr2:73430000_73490000	60	*fst*	low	57 - 88	0.52 - 1.08	71 - 120	0.19 - 0.27	5.07 - 7.56	
chr2:82270000_82290000	20	*fst*	high	15 - 15	0.6 - 0.6	26 - 26	0.2 - 0.2	5.22 - 5.22	*VSTM2A*
chr2:82330000_82360000	30	*fst*	high	4 - 7	1.06 - 1.09	18 - 20	0.21 - 0.26	5.58 - 7.07	
chr2:82370000_82390000	20	*fst*	high	11 - 11	0.71 - 0.71	64 - 64	0.19 - 0.19	5.04 - 5.04	
chr2:82580000_82600000	20	*fst*	high	15 - 15	1.79 - 1.79	130 - 130	0.19 - 0.19	5.06 - 5.06	
chr2:82630000_82770000	140	*fst*	high	27 - 168	1.03 - 1.63	70 - 231	0.2 - 0.3	5.25 - 8.37	*ENSGALG00000049837, ENSGALG00000051256 (LncRNA)*
chr2:82790000_82840000	50	*fst*	high	61 - 71	1.42 - 1.7	74 - 102	0.19 - 0.22	5.01 - 5.97	*ENSGALG00000049962 (LncRNA)*
chr2:82850000_82940000	90	*fst*	high	47 - 83	1.19 - 1.69	76 - 119	0.2 - 0.35	5.47 - 9.92	*ENSGALG00000049962 (LncRNA), ENSGALG00000051615 (LncRNA)*
chr2:99490000_99530000	40	*fst*	low	174 - 221	0.55 - 0.83	311 - 417	0.19 - 0.21	5.19 - 5.52	*ENSGALG00000042652, LAMA1*
chr2:103540000_103570000	30	*xp-ehh*	low	58 - 90	3 - 3.03	217 - 264	0.1 - 0.12	2.25 - 2.91	
chr2:117230000_117290000	60	*xp-ehh*	high	170 - 243	3.34 - 3.5	246 - 352	0.09 - 0.13	2.13 - 3.23	
chr2:117840000_117860000	20	*fst*	high	193 - 193	2.61 - 2.61	268 - 268	0.19 - 0.19	5.09 - 5.09	*STAU2*
chr2:117960000_118010000	50	*fst*	high	142 - 183	1.1 - 1.62	213 - 308	0.21 - 0.27	5.73 - 7.47	*UBE2W*
chr2:128080000_128150000	70	*xp-ehh*	low	131 - 214	3.12 - 3.37	208 - 294	0.07 - 0.13	1.44 - 3.37	
chr3:10700000_10720000	20	*fst*	low	149 - 149	1.12 - 1.12	321 - 321	0.19 - 0.19	5.15 - 5.15	*SPRED2*

chr3:34260000_34320000	60	*xp-ehh* and/or *fst*	low	131 - 181	2.87 - 3	210 - 274	0.19 - 0.28	5.04 - 7.71	*KIF26B, SMYD3*
chr3:36150000_36170000	20	*fst*	high	63 - 63	1.7 - 1.7	132 - 132	0.2 - 0.2	5.48 - 5.48	*RGS7*
chr3:71120000_71150000	30	*fst*	low	100 - 102	1.46 - 1.59	232 - 236	0.19 - 0.19	5 - 5.06	
chr3:76980000_77000000	20	*fst*	low	188 - 188	1.15 - 1.15	267 - 267	0.19 - 0.19	5.06 - 5.06	*NT5E*
chr3:77630000_77650000	20	*xp-ehh*	low	216 - 216	3.01 - 3.01	355 - 355	0.03 - 0.03	0.36 - 0.36	*SNAP91*
chr3:91040000_91060000	20	*xp-ehh*	high	203 - 203	3.18 - 3.18	273 - 273	0.11 - 0.11	2.65 - 2.65	*DLGAP2*
chr3:94950000_94970000	20	*xp-ehh*	high	141 - 141	3.16 - 3.16	232 - 232	0.02 - 0.02	0.01 - 0.01	*ENSGALG00000026867 (ncRNA)*
chr3:97030000_97100000	70	*xp-ehh* / *fst*	high	166 - 200	2.76 - 3.23	285 - 341	0.09 - 0.19	2.17 - 5.16	
chr3:103450000_103490000	40	*xp-ehh*	high	103 - 165	3.05 - 3.68	174 - 220	0.05 - 0.07	0.94 - 1.44	
chr3:103660000_103680000	20	*xp-ehh*	high	203 - 203	3.21 - 3.21	280 - 280	0.06 - 0.06	1.01 - 1.01	
chr3:106970000_107000000	30	*xp-ehh*	low	294 - 299	3.27 - 3.34	443 - 482	0.07 - 0.13	1.36 - 3.11	
chr4:3190000_3220000	30	*xp-ehh*	low	143 - 166	3.05 - 3.09	357 - 362	0.13 - 0.15	3.39 - 3.94	*KLHL9*
chr4:4590000_4620000	30	*xp-ehh*	low	165 - 195	3.09 - 3.12	274 - 293	0.07 - 0.09	1.38 - 2.01	*ENSGALG00000047098 (LncRNA), ENSGALG00000049919 (LncRNA), ENSGALG00000052627 (LncRNA)*
chr4:12480000_12510000	30	*xp-ehh*	low	166 - 199	3.01 - 3.02	310 - 365	0.06 - 0.08	1.16 - 1.77	*NEXMIF*
chr4:24070000_24090000	20	*xp-ehh*	high	152 - 152	3.06 - 3.06	237 - 237	0.07 - 0.07	1.44 - 1.44	*TLL1*
chr4:24350000_24400000	50	*xp-ehh*	high	48 - 145	3.18 - 3.61	135 - 284	0.09 - 0.18	2.14 - 4.69	
chr4:24440000_24480000	40	*xp-ehh*	high	75 - 179	3 - 3.52	134 - 312	0.1 - 0.18	2.39 - 4.83	*ENSGALG00000051758 (LncRNA)*
chr4:25570000_25600000	30	*fst*	high	171 - 211	2.26 - 2.41	284 - 317	0.2 - 0.23	5.33 - 6.16	
chr4:25890000_26440000	550	*xp-ehh*	high	55 - 197	3.02 - 4.02	85 - 301	0.05 - 0.18	0.98 - 4.62	*ENSGALG00000046909 (LncRNA), ENSGALG00000047918 (LncRNA), ENSGALG00000049671 (LncRNA), ENSGALG00000050306 (LncRNA), ENSGALG00000050784 (LncRNA)*
chr4:26520000_26600000	80	*xp-ehh*	high	54 - 110	3.06 - 3.32	120 - 177	0.05 - 0.09	0.71 - 2.19	*ENSGALG00000053826 (LncRNA)*
chr4:27670000_27690000	20	*fst*	high	67 - 67	0.75 - 0.75	115 - 115	0.21 - 0.21	5.56 - 5.56	
chr4:29220000_29300000	80	*xp-ehh*	high	233 - 317	3.11 - 3.27	449 - 530	0.09 - 0.12	1.94 - 2.86	*ELF2, MGARP, NAA15*
chr4:36410000_36480000	70	*fst*	high	91 - 155	0.21 - 0.59	153 - 367	0.19 - 0.26	5.04 - 7.05	*ENSGALG00000047741*
chr4:54270000_54310000	40	*fst*	high	97 - 182	0.68 - 1.41	175 - 290	0.25 - 0.32	6.93 - 8.96	*PRDM5, ENSGALG00000026726 (ncRNA)*
chr4:54330000_54390000	60	*fst*	low	155 - 233	0.5 - 1.32	244 - 316	0.2 - 0.23	5.3 - 6.22	*ENSGALG00000053393 (LncRNA)*

chr4:56140000_56170000	30	*xp-ehh* and/or *fst*	low	112 - 125	2.89 - 3.01	182 - 223	0.19 - 0.21	5.12 - 5.54	*NDST4*
chr4:72710000_72730000	20	*fst*	low	75 - 75	2.27 - 2.27	125 - 125	0.2 - 0.2	5.32 - 5.32	
chr4:75600000_75620000	20	*fst*	high	5 - 5	2.07 - 2.07	79 - 79	0.19 - 0.19	5.13 - 5.13	
chr4:75690000_75810000	120	*xp-ehh* and/or *fst*	high	98 - 171	3.12 - 3.66	164 - 253	0.12 - 0.2	3.09 - 5.4	*ENSGALG00000052764 (LncRNA), NCAPG*
chr5:6840000_6860000	20	*xp-ehh*	high	229 - 229	3.07 - 3.07	383 - 383	0.03 - 0.03	0.12 - 0.12	
chr5:6880000_6910000	30	*xp-ehh*	high	245 - 279	3.11 - 3.19	377 - 416	0.06 - 0.08	1.29 - 1.87	
chr5:35600000_35660000	60	*xp-ehh*	high	88 - 142	3.06 - 3.14	126 - 191	0.11 - 0.13	2.59 - 3.39	*NPAS3*
chr5:40630000_40710000	80	*fst*	low	119 - 241	1.51 - 2.19	168 - 289	0.22 - 0.28	5.84 - 7.88	*ENSGALG00000026686 (MiRNA), NRXN3*
chr5:44410000_44430000	20	*xp-ehh*	high	276 - 276	3.04 - 3.04	324 - 324	0.05 - 0.05	0.99 - 0.99	*DGLUCY, GPR68*
chr5:47570000_47600000	30	*fst*	low	152 - 156	2.33 - 2.35	197 - 207	0.21 - 0.21	5.53 - 5.62	*ENSGALG00000050961 (LncRNA)*
chr5:48990000_49080000	90	*fst*	high	6 - 13	2.21 - 2.21	83 - 183	0.19 - 0.25	5.11 - 7	*BEGAIN, ENSGALG00000051728 (LncRNA), ENSGALG00000053427 (LncRNA)*
chr5:51290000_51350000	60	*fst*	low	147 - 205	0.15 - 0.74	233 - 308	0.19 - 0.26	5.01 - 7.25	*ENSGALG00000053209 (LncRNA), ENSGALG00000054591 (LncRNA)*
chr5:58190000_58230000	40	*xp-ehh*	low	214 - 217	3.19 - 3.32	347 - 424	0.03 - 0.04	0.21 - 0.56	*FRMD6, GNG2*
chr6:16050000_16080000	30	*fst*	low	88 - 99	2.71 - 2.82	264 - 272	0.22 - 0.23	6.05 - 6.24	*ADK*
chr6:20130000_20160000	30	*fst*	high	127 - 136	1.68 - 1.89	190 - 233	0.21 - 0.21	5.54 - 5.68	*ENSGALG00000048378 (LncRNA), ENSGALG00000053893 (LncRNA)*
chr6:20200000_20220000	20	*fst*	high	237 - 237	1.89 - 1.89	443 - 443	0.2 - 0.2	5.29 - 5.29	*SLC16A12*
chr6:21520000_21550000	30	*xp-ehh*	low	180 - 187	3.05 - 3.08	304 - 312	0.07 - 0.07	1.5 - 1.57	*CYP26A1, CYP26C1, EXOC6*
chr6:21690000_21760000	70	*fst*	low	285 - 398	0.9 - 2.93	400 - 568	0.19 - 0.23	5 - 6.33	*IDE, KIF11*
chr6:21780000_21820000	40	*fst*	low	234 - 271	1.51 - 1.98	366 - 439	0.19 - 0.19	5.01 - 5.08	*CPEB3, MARCH5*
chr6:31450000_31470000	20	*fst*	low	128 - 128	1.51 - 1.51	243 - 243	0.19 - 0.19	5.01 - 5.01	
chr7:26010000_26050000	40	*xp-ehh*	low	135 - 212	3.02 - 3.24	298 - 340	0.08 - 0.17	1.82 - 4.37	*TFCP2L1*
chr7:30200000_30220000	20	*xp-ehh*	high	164 - 164	3.01 - 3.01	342 - 342	0.07 - 0.07	1.35 - 1.35	*ENSGALG00000055018 (LncRNA)*
chr7:31520000_31550000	30	*fst*	high	167 - 178	0.93 - 0.97	305 - 306	0.19 - 0.19	5.11 - 5.15	*ENSGALG00000048656*
chr8:1520000_1570000	50	*fst*	low	176 - 247	1.26 - 2.29	258 - 312	0.19 - 0.22	5.09 - 5.91	*ENSGALG00000046646, VAV3*
chr8:3060000_3080000	20	*fst*	low	131 - 131	2.53 - 2.53	343 - 343	0.19 - 0.19	5.03 - 5.03	*ENSGALG00000002431*
chr8:3330000_3350000	20	*xp-ehh*	low	243 - 243	3.03 - 3.03	356 - 356	0.12 - 0.12	2.96 - 2.96	*ENSGALG00000049765 (LncRNA), ENSGALG00000050518 (LncRNA)*
chr8:5750000_5770000	20	*xp-ehh*	low	196 - 196	3.53 - 3.53	367 - 367	0.13 - 0.13	3.26 - 3.26	

chr8:20050000_20080000	30	*fst*	low	200 - 242	2.65 - 2.92	344 - 377	0.26 - 0.29	7.04 - 8.16	*MED8, SZT2*
chr8:23790000_23810000	20	*xp-ehh*	low	166 - 166	3.03 - 3.03	284 - 284	0.07 - 0.07	1.37 - 1.37	*BEND5*
chr8:24150000_24180000	30	*xp-ehh*	high	169 - 208	3.01 - 3.06	255 - 293	0.06 - 0.07	1.3 - 1.6	*FAF1*
chr8:29170000_29220000	50	*xp-ehh*	low	117 - 175	3.05 - 3.17	191 - 264	0.1 - 0.14	2.36 - 3.58	
chr9:8290000_8310000	20	*xp-ehh*	low	275 - 275	3.03 - 3.03	468 - 468	0.07 - 0.07	1.39 - 1.39	
chr9:8410000_8440000	30	*xp-ehh*	low	205 - 214	3.08 - 3.15	396 - 414	0.04 - 0.05	0.66 - 0.79	*AP1S3*
chr9:18190000_18420000	230	*xp-ehh*	low	135 - 219	3.02 - 3.91	260 - 388	0.01 - 0.12	-0.42 - 3.01	*ENSGALG00000047096 (LncRNA), ENSGALG00000049140 (LncRNA), ENSGALG00000050658 (LncRNA), ENSGALG00000050910 (LncRNA), ENSGALG00000051388 (LncRNA), ENSGALG00000054169 (LncRNA), ENSGALG00000054791 (LncRNA)*
chr10:2310000_2330000	20	*xp-ehh*	low	146 - 146	3.05 - 3.05	336 - 336	0.08 - 0.08	1.69 - 1.69	*NEO1*
chr10:10470000_10500000	30	*fst*	low	178 - 189	0.87 - 1.08	263 - 291	0.2 - 0.21	5.29 - 5.51	*ENSGALG00000051344 (LncRNA), SLC24A5*
chr11:5940000_5970000	30	*xp-ehh*	low	118 - 169	3.01 - 3.37	280 - 415	0.05 - 0.07	0.91 - 1.39	*SALL1*
chr11:13140000_13170000	30	*fst*	low	184 - 203	0.63 - 0.68	254 - 271	0.19 - 0.19	5.05 - 5.1	*ENSGALG00000050348 (LncRNA)*
chr11:13240000_13270000	30	*fst*	low	186 - 224	0.32 - 0.4	249 - 284	0.19 - 0.2	5.05 - 5.4	*ENSGALG00000050348 (LncRNA)*
chr12:2810000_2840000	30	*xp-ehh*	low	249 - 308	3.17 - 3.41	429 - 474	0.03 - 0.03	0.13 - 0.19	*IP6K1*
chr12:8330000_8370000	40	*xp-ehh*	low	129 - 148	3.05 - 3.11	195 - 224	0.12 - 0.13	2.92 - 3.36	
chr13:2150000_2170000	20	*fst*	low	9 - 9	0.26 - 0.26	36 - 36	0.19 - 0.19	5.03 - 5.03	*PFDN1*
chr13:7880000_7910000	30	*fst*	low	170 - 200	1.09 - 1.68	234 - 267	0.19 - 0.19	5.04 - 5.14	
chr13:11300000_11330000	30	*fst*	high	193 - 206	0.76 - 1.13	332 - 345	0.19 - 0.22	5.11 - 6.01	*SPINK6, UBLCP1*
chr13:17870000_17900000	30	*fst*	low	91 - 123	2 - 2.35	247 - 281	0.21 - 0.21	5.63 - 5.65	*GNPDA1, NDFIP1, RNF14*
chr14:13980000_14010000	30	*xp-ehh*	low	159 - 237	3.17 - 3.29	235 - 338	0.08 - 0.1	1.71 - 2.25	*ABCA3*
chr14:14660000_14690000	30	*xp-ehh*	low	294 - 338	3.08 - 3.14	419 - 459	0.03 - 0.04	0.33 - 0.62	*MAPK8IP3*
chr14:14740000_14760000	20	*xp-ehh*	low	172 - 172	3.06 - 3.06	296 - 296	0.05 - 0.05	0.96 - 0.96	*GLYR1*
chr15:4550000_4590000	40	*xp-ehh*	high	203 - 264	3.02 - 3.21	363 - 414	0.02 - 0.03	-0.02 - 0.21	*TMEM132B*
chr18:6570000_6590000	20	*xp-ehh*	high	193 - 193	3.15 - 3.15	296 - 296	0.1 - 0.1	2.35 - 2.35	*UTP6, UTP6*
chr22:850000_880000	30	*fst*	low	7 - 9	2.12 - 2.14	23 - 33	0.22 - 0.23	5.88 - 6.12	*EBF2*
chr23:5400000_5450000	50	*xp-ehh*	low	120 - 165	3.06 - 3.29	256 - 362	0.03 - 0.05	0.29 - 0.81	*FNDC5, HPCA, KIAA1522, RNF19B, YARS*
chr26:1810000_1840000	30	*xp-ehh*	low	157 - 176	3.4 - 3.56	357 - 380	0.09 - 0.11	1.92 - 2.64	*MDM4, PIK3C2B*

chr26:3940000_3970000	30	*fst*	low	244 - 260	1.65 - 2.22	378 - 384	0.21 - 0.21	5.59 - 5.73	*ENSGALG00000002170, ENSGALG00000002511, ENSGALG00000054224*
chr27:7660000_7710000	50	*xp-ehh*	low	48 - 103	3.18 - 3.61	112 - 184	0.05 - 0.09	0.97 - 2.17	*ENSGALG00000003282, ENSGALG00000003345, gga-mir-12256, gga-mir-12257, GHDC, HCRT, KCNH4, RAB5C*
chr28:4520000_4570000	50	*fst*	low	42 - 129	2.02 - 2.57	150 - 216	0.2 - 0.23	5.4 - 6.22	*CIB3, ENSGALG00000050334 (LncRNA), RAB8A, TPM4*

Table S4.3a. Candidate SNP list from signature analysis for minTemp population group. (Only SNPs with large differences in allele frequency (dAAF>0.5) between Low and High groups are included. At least one SNP with dAAF>0.5 for each candidate gene is included. |dAAF|: difference in allele frequency between Low and High group).

snpID	**rs_id**	**ref**	**alt**	**annotation**	**gene**	**gene ID**	**Low group pop1 AAF**	**Low group pop2 AAF**	**High group pop1 AAF**	**High group pop2 AAF**	**$\bar{X}$ AAF in Low group**	**$\bar{X}$ AAF in High group**	**|dAAF|**
chr21_3340742	rs315879878	G	T	intronic	*ENSGALG00000002506*	*SPSB1*	0.80	0.85	0.40	0.15	0.83	0.28	0.55
chr26_4160183	rs317851952	A	T	upstream	*ENSGALG00000002675*		0.95	0.85	0.40	0.25	0.90	0.33	0.58
chr26_4179927	rs736970763	C	G	intronic	*ENSGALG00000002692*		0.90	0.80	0.30	0.20	0.85	0.25	0.60
chr18_5095581	rs741306978	G	A	intronic	*ENSGALG00000002954*	*MBTD1*	0.85	0.80	0.25	0.15	0.83	0.20	0.63
chr18_5105184	rs732629913	T	C	downstream	*ENSGALG00000002985*	*UTP18*	0.85	0.80	0.00	0.20	0.83	0.10	0.73
chr18_5099822	rs738364253	A	G	nonsynonymous	*ENSGALG00000002985*	*UTP18*	0.80	0.80	0.00	0.20	0.80	0.10	0.70
chr18_5097158	rs314168808	C	T	nonsynonymous	*ENSGALG00000002985*	*UTP18*	0.85	0.80	0.25	0.15	0.83	0.20	0.63
chr18_5097978	rs313980559	C	A	nonsynonymous	*ENSGALG00000002985*	*UTP18*	0.80	0.80	0.25	0.20	0.80	0.23	0.58
chr18_5156846	rs317312418	T	C	intronic	*ENSGALG00000002993*	*CA10*	0.95	0.80	0.05	0.40	0.88	0.23	0.65
chr18_6579283	rs15826593	G	A	upstream	*ENSGALG00000003260*	*UTP6*	0.15	0.25	0.90	0.80	0.20	0.85	0.65
chr1_197380178	na	A	G	intronic	*ENSGALG00000003416*	*FAM160A2*	0.90	0.85	0.40	0.30	0.88	0.35	0.53
chr14_1629928	rs735581328	A	G	intronic	*ENSGALG00000003562*	*NPTX2*	0.80	0.75	0.05	0.40	0.78	0.23	0.55
chr21_5235097	na	G	T	downstream	*ENSGALG00000004097*	*LRRC38*	0.75	0.75	0.30	0.10	0.75	0.20	0.55
chr14_3947909	rs736665750	G	T	intronic	*ENSGALG00000004420*	*SDK1*	0.75	0.90	0.05	0.30	0.83	0.18	0.65
chr14_5131529	na	A	C	intronic	*ENSGALG00000004816*		0.70	0.75	0.15	0.15	0.73	0.15	0.58
chr14_5142884	rs738767944	C	T	UTR3	*ENSGALG00000004834*	*COPS3*	0.75	0.75	0.15	0.15	0.75	0.15	0.60
chr14_5155289	na	G	C	intronic	*ENSGALG00000004842*	*NT5M*	0.70	0.80	0.10	0.20	0.75	0.15	0.60
chr5_17263013	na	G	A	downstream	*ENSGALG00000004845*	*TIMM10*	0.85	0.75	0.10	0.15	0.80	0.13	0.68
chr19_7460814	rs14122580	A	G	intronic	*ENSGALG00000005160*	*VMP1*	0.75	0.85	0.30	0.25	0.80	0.28	0.53
chr2_4618697	rs317672568	A	G	intronic	*ENSGALG00000005710*	*CTDSPL*	0.25	0.25	0.95	0.70	0.25	0.83	0.58
chr7_6760479	rs10723271	G	T	nonsynonymous	*ENSGALG00000005974*	*COL6A1*	0.80	0.85	0.15	0.30	0.83	0.23	0.60
chr2_5051731	rs734773228	A	G	intronic	*ENSGALG00000006025*	*XIRP1*	0.65	0.80	0.30	0.10	0.73	0.20	0.53
chr5_16925158	rs316806623	G	A	UTR3	*ENSGALG00000007127*		0.75	0.90	0.05	0.10	0.83	0.08	0.75
chr5_16957578	rs315213418	A	T	UTR3	*ENSGALG00000007178*	*FADS2*	0.10	0.15	0.90	0.80	0.13	0.85	0.73
chr5_16944732	rs15673184	A	G	nonsynonymous	*ENSGALG00000007178*	*FADS2*	0.55	0.75	0.00	0.15	0.65	0.08	0.58
chr5_16967014	rs1059511401	A	C	intronic	*ENSGALG00000007203*	*RAB3IL1*	0.85	0.90	0.10	0.25	0.88	0.18	0.70

chr5_16973935	rs734597481	C	T	nonsynonymous	*ENSGALG00000007203*	*RAB3IL1*	0.90	0.85	0.30	0.35	0.88	0.33	0.55
chr10_18147314	na	C	T	intronic	*ENSGALG00000007252*	*ANKDD1A*	0.85	0.75	0.25	0.15	0.80	0.20	0.60
chr10_18164060	rs315348044	A	G	intronic	*ENSGALG00000007262*	*SPG21*	0.90	0.75	0.25	0.20	0.83	0.23	0.60
chr5_17182206	rs1059217695	G	A	UTR5	*ENSGALG00000007287*		0.85	0.95	0.30	0.15	0.90	0.23	0.68
chr10_18171192	rs14952795	T	C	intronic	*ENSGALG00000007289*	*CLPX*	0.10	0.35	0.75	0.75	0.23	0.75	0.53
chr5_17197342	rs15673614	C	T	intronic	*ENSGALG00000007330*	*CTNND1*	0.20	0.10	0.85	0.85	0.15	0.85	0.70
chr5_17231117	rs317045995	T	C	nonsynonymous	*ENSGALG00000007341*	*MED19*	0.55	0.75	0.15	0.10	0.65	0.13	0.53
chr5_17242235	rs740651106	T	C	upstream	*ENSGALG00000007362*	*ZDHHC5*	0.85	0.70	0.10	0.05	0.78	0.08	0.70
chr5_17248017	rs739291012	G	C	downstream	*ENSGALG00000007370*	*CLP1*	0.95	0.80	0.15	0.10	0.88	0.13	0.75
chr5_17252599	rs1057834097	C	T	intronic	*ENSGALG00000007374*	*YPEL4*	0.75	0.75	0.10	0.20	0.75	0.15	0.60
chr5_17261040	rs1058920117	T	C	UTR5	*ENSGALG00000007381*		0.95	0.95	0.20	0.35	0.95	0.28	0.68
chr5_17256023	rs739581191	C	G	nonsynonymous	*ENSGALG00000007381*		0.05	0.15	0.70	0.60	0.10	0.65	0.55
chr5_17256193	rs739661866	C	T	nonsynonymous	*ENSGALG00000007381*		0.20	0.25	0.75	0.75	0.23	0.75	0.53
chr1_196930434	na	T	C	upstream	*ENSGALG00000007441*	*MAP6*	0.90	0.80	0.35	0.25	0.85	0.30	0.55
chr5_17283707	na	C	T	intronic	*ENSGALG00000007520*	*SSRP1*	0.80	0.75	0.05	0.30	0.78	0.18	0.60
chr18_10410366	na	C	T	UTR3	*ENSGALG00000007713*		0.70	0.70	0.30	0.00	0.70	0.15	0.55
chr4_14107250	rs737145126	T	C	intronic	*ENSGALG00000008282*	*ATG4A*	0.85	0.75	0.10	0.25	0.80	0.18	0.63
chr1_27533032	rs741250331	G	A	intronic	*ENSGALG00000009485*	*IMMP2L*	0.10	0.20	0.70	0.65	0.15	0.68	0.53
chr27_5719783	rs315239810	G	T	intronic	*ENSGALG00000009943*	*FAM117A*	0.70	0.90	0.05	0.35	0.80	0.20	0.60
chr27_5715201	rs317359923	C	T	intronic	*ENSGALG00000009944*	*KAT7*	0.25	0.05	0.90	0.80	0.15	0.85	0.70
chr5_34721145	rs314147964	A	G	intronic	*ENSGALG00000009983*	*NUBPL*	0.85	0.85	0.00	0.45	0.85	0.23	0.63
chr8_24344580	rs313880413	C	T	intronic	*ENSGALG00000010539*	*RNF11*	0.05	0.10	0.75	0.70	0.08	0.73	0.65
chr3_36713864	rs80704816	C	T	intronic	*ENSGALG00000010778*	*CHRM1*	0.25	0.15	0.90	0.75	0.20	0.83	0.63
chr1_49902052	rs738966909	A	G	nonsynonymous	*ENSGALG00000012003*	*XPNPEP3*	0.90	0.75	0.40	0.15	0.83	0.28	0.55
chr1_49891637	rs730966112	C	G	nonsynonymous	*ENSGALG00000012003*	*XPNPEP3*	0.90	0.70	0.30	0.25	0.80	0.28	0.53
chr1_49932449	na	G	A	intronic	*ENSGALG00000012013*	*SLC25A17*	0.90	0.75	0.25	0.20	0.83	0.23	0.60
chr1_50054793	rs740590617	G	A	nonsynonymous	*ENSGALG00000012024*	*MRTFA*	0.95	0.80	0.30	0.20	0.88	0.25	0.63
chr1_50054827	rs734326086	C	T	nonsynonymous	*ENSGALG00000012024*	*MRTFA*	0.95	0.80	0.30	0.20	0.88	0.25	0.63
chr1_56032224	rs13870546	G	A	intronic	*ENSGALG00000012792*	*HIPK2*	0.15	0.10	0.70	0.85	0.13	0.78	0.65
chr1_56228517	rs312403160	G	T	intronic	*ENSGALG00000012796*		0.65	0.95	0.20	0.25	0.80	0.23	0.58
chr1_56239018	rs315472386	A	G	nonsynonymous	*ENSGALG00000012796*		0.55	0.10	0.85	0.85	0.33	0.85	0.53

chr1_56240399	rs15277089	T	A	nonsynonymous	*ENSGALG00000012796*		0.55	0.10	0.85	0.85	0.33	0.85	0.53
chr2_68088249	rs316937055	G	A	intronic	*ENSGALG00000012888*	*PHLPP1*	0.10	0.10	0.95	0.60	0.10	0.78	0.68
chr4_63136574	rs733978669	A	G	UTR3	*ENSGALG00000013646*	*VPS37A*	0.70	0.90	0.10	0.25	0.80	0.18	0.63
chr4_63912289	na	T	C	intronic	*ENSGALG00000013684*	*SGCZ*	0.80	0.75	0.25	0.10	0.78	0.18	0.60
chr4_65679957	rs318134095	G	A	intronic	*ENSGALG00000013940*	*LNX1*	0.80	0.95	0.15	0.55	0.88	0.35	0.53
chr5_16420305	rs315888203	T	C	upstream	*ENSGALG00000014152*		0.10	0.15	0.90	0.60	0.13	0.75	0.63
chr1_77832373	rs313794401	T	C	intronic	*ENSGALG00000014468*	*MLF2*	0.20	0.40	0.85	0.80	0.30	0.83	0.53
chr1_109134117	rs732503598	A	G	intronic	*ENSGALG00000016059*	*ETS2*	0.85	0.90	0.30	0.40	0.88	0.35	0.53
chr1_109639334	rs739509160	G	A	intronic	*ENSGALG00000016138*	*DSCAM*	0.15	0.05	0.65	0.60	0.10	0.63	0.53
chr1_111026718	rs317526225	C	T	intronic	*ENSGALG00000016196*		0.70	0.95	0.30	0.05	0.83	0.18	0.65
chr3_108234307	rs740753540	T	C	intronic	*ENSGALG00000016680*	*TFAP2B*	0.90	0.90	0.20	0.40	0.90	0.30	0.60
chr1_145479945	rs316657473	T	C	intronic	*ENSGALG00000016872*	*PCCA*	0.85	0.90	0.35	0.20	0.88	0.28	0.60
chr1_197374203	na	G	C	downstream	*ENSGALG00000017299*	*RPS11*	0.90	0.80	0.30	0.15	0.85	0.23	0.63
chr1_195837260	na	C	G	intronic	*ENSGALG00000017304*	*PGM2L1*	0.95	0.85	0.15	0.25	0.90	0.20	0.70
chr1_195827151	na	C	G	nonsynonymous	*ENSGALG00000017304*	*PGM2L1*	0.90	0.80	0.15	0.35	0.85	0.25	0.60
chr1_197289816	na	G	A	upstream	*ENSGALG00000017347*	*HBBR*	0.95	0.80	0.40	0.20	0.88	0.30	0.58
chr18_10416434	na	T	A	nonsynonymous	*ENSGALG00000019755*		0.70	0.70	0.35	0.00	0.70	0.18	0.53
chr3_108310019	rs312701210	C	A	intronic	*ENSGALG00000019842*	*TFAP2D*	0.95	0.85	0.30	0.20	0.90	0.25	0.65
chr5_17269803	rs740320098	A	G	intronic	*ENSGALG00000020479*	*SLC43A3*	0.90	0.75	0.10	0.10	0.83	0.10	0.73
chr5_17278237	rs731898985	A	G	nonsynonymous	*ENSGALG00000020479*	*SLC43A3*	0.80	0.70	0.15	0.15	0.75	0.15	0.60
chr5_16904462	rs737081707	G	C	intronic	*ENSGALG00000020485*		0.65	0.80	0.05	0.15	0.73	0.10	0.63
chr21_2007660	rs315268906	G	C	intronic	*ENSGALG00000021676*		0.85	0.80	0.00	0.30	0.83	0.15	0.68
chr1_197390872	na	G	A	nonsynonymous	*ENSGALG00000022704*		0.90	0.80	0.15	0.40	0.85	0.28	0.58
chr18_10429868	rs738416634	G	A	intronic	*ENSGALG00000027789*	*RAB37*	0.70	0.70	0.30	0.05	0.70	0.18	0.53
chr1_197283369	na	G	A	UTR3	*ENSGALG00000028273*	*HBE1*	0.95	0.80	0.45	0.20	0.88	0.33	0.55
chr4_89590013	rs735474386	G	A	intronic	*ENSGALG00000030602*	*ADAM33*	0.60	0.60	0.00	0.15	0.60	0.08	0.53
chr1_195808059	na	A	G	downstream	*ENSGALG00000030652*		0.90	0.80	0.05	0.20	0.85	0.13	0.73
chr1_197267802	na	G	A	nonsynonymous	*ENSGALG00000030885*	*OR51M1*	0.95	0.75	0.30	0.10	0.85	0.20	0.65
chr23_2992673	rs734052742	C	T	intronic	*ENSGALG00000031440*	*PTPRU*	0.90	0.80	0.30	0.20	0.85	0.25	0.60
chr21_1998185	rs739470072	A	T	UTR3	*ENSGALG00000032413*	*NADK*	0.80	0.80	0.00	0.35	0.80	0.18	0.63
chr5_17264521	na	G	A	nonsynonymous	*ENSGALG00000032903*	*RTN4RL2*	0.85	0.75	0.10	0.20	0.80	0.15	0.65
chr1_468193	rs735888526	A	G	ncRNA_intronic	*ENSGALG00000033919*		0.05	0.10	0.70	0.70	0.08	0.70	0.63
chr1_197417380	na	C	G	intronic	*ENSGALG00000034004*	*TRIM3*	0.95	0.85	0.15	0.55	0.90	0.35	0.55
chr1_197418047	na	A	G	nonsynonymous	*ENSGALG00000034004*	*TRIM3*	0.95	0.90	0.25	0.55	0.93	0.40	0.53

chr1_196804406	na	T	G	intronic	*ENSGALG00000034967*	*GDPD5*	0.90	0.80	0.25	0.35	0.85	0.30	0.55
chr1_197277566	na	A	G	upstream	*ENSGALG00000035309*	*HBE*	0.95	0.80	0.45	0.20	0.88	0.33	0.55
chr12_3121160	na	C	T	intronic	*ENSGALG00000035578*	*RBM6*	0.75	0.75	0.00	0.20	0.75	0.10	0.65
chr1_197401060	na	C	T	nonsynonymous	*ENSGALG00000036106*	*SMPD1*	0.90	0.80	0.15	0.35	0.85	0.25	0.60
chr1_197401171	na	G	A	nonsynonymous	*ENSGALG00000036106*	*SMPD1*	0.90	0.80	0.15	0.35	0.85	0.25	0.60
chr2_43430775	rs735453043	A	G	upstream	*ENSGALG00000036110*		0.75	0.80	0.25	0.20	0.78	0.23	0.55
chr8_23765614	rs741698490	C	T	intronic	*ENSGALG00000036909*	*BEND5*	0.60	0.65	0.00	0.10	0.63	0.05	0.58
chr1_197438282	na	A	G	intronic	*ENSGALG00000037839*	*ARFIP2*	0.05	0.20	0.75	0.70	0.13	0.73	0.60
chr2_43920885	na	G	A	upstream	*ENSGALG00000038646*	*EIF1B*	0.80	0.75	0.05	0.40	0.78	0.23	0.55
chr1_197395093	na	C	T	downstream	*ENSGALG00000039837*	*CCKBR*	0.90	0.80	0.15	0.35	0.85	0.25	0.60
chr5_17280906	na	G	A	ncRNA_intronic	*ENSGALG00000040479*	*P2RX3*	0.80	0.75	0.00	0.30	0.78	0.15	0.63
chr5_17280624	na	C	T	ncRNA_exonic	*ENSGALG00000040479*	*P2RX3*	0.80	0.75	0.05	0.30	0.78	0.18	0.60
chr5_17280173	rs314385185	G	A	ncRNA_exonic	*ENSGALG00000040479*	*P2RX3*	0.20	0.25	0.95	0.70	0.23	0.83	0.60
chr5_17280159	rs318189997	T	C	ncRNA_exonic	*ENSGALG00000040479*	*P2RX3*	0.15	0.25	0.90	0.65	0.20	0.78	0.58
chr14_4909738	rs736382989	C	A	intronic	*ENSGALG00000041285*	*MAP2K3*	0.65	0.80	0.10	0.20	0.73	0.15	0.58
chr2_43959021	rs317408968	T	C	intronic	*ENSGALG00000041372*		0.05	0.15	0.85	0.45	0.10	0.65	0.55
chr1_197438993	na	G	C	intronic	*ENSGALG00000042572*		0.00	0.20	0.75	0.70	0.10	0.73	0.63
chr2_42164063	na	G	T	intronic	*ENSGALG00000043415*	*ACPP*	0.60	0.80	0.10	0.25	0.70	0.18	0.53
chr5_17275619	rs794312629	G	A	UTR3	*ENSGALG00000043860*	*PGR2/3*	0.75	0.65	0.20	0.15	0.70	0.18	0.53
chr1_197225428	na	G	A	intronic	*ENSGALG00000044591*	*OR51E2*	1.00	0.80	0.25	0.25	0.90	0.25	0.65
chr4_89603283	rs16453444	G	A	intronic	*ENSGALG00000044620*		0.80	0.90	0.00	0.45	0.85	0.23	0.63
chr5_17258173	rs1059780782	A	G	nonsynonymous	*ENSGALG00000046098*	*UBE2L6*	0.90	0.85	0.15	0.25	0.88	0.20	0.68
chr1_197404221	na	T	C	intronic	*ENSGALG00000046164*	*APBB1*	0.95	0.80	0.15	0.35	0.88	0.25	0.63
chr4_43496808	rs738114890	G	C	ncRNA_intronic	*ENSGALG00000046699*		0.75	0.80	0.05	0.25	0.78	0.15	0.63
chr1_49880947	rs736613178	T	A	ncRNA_intronic	*ENSGALG00000046744*		0.90	0.70	0.30	0.25	0.80	0.28	0.53
chr20_4516141	rs314237836	G	C	ncRNA_intronic	*ENSGALG00000046818*		0.85	0.70	0.20	0.15	0.78	0.18	0.60
chr3_108130465	rs312702646	G	C	ncRNA_intronic	*ENSGALG00000047869*		0.90	0.95	0.45	0.25	0.93	0.35	0.58
chr1_109515507	rs732840204	G	T	intronic	*ENSGALG00000047890*	*PCP4*	0.75	0.90	0.15	0.35	0.83	0.25	0.58
chr4_43505891	rs741120306	T	C	ncRNA_intronic	*ENSGALG00000048030*		0.05	0.15	0.95	0.55	0.10	0.75	0.65
chr4_43502734	rs313378340	G	A	ncRNA_exonic	*ENSGALG00000048030*		0.75	0.80	0.05	0.25	0.78	0.15	0.63
chr1_197263129	na	T	A	downstream	*ENSGALG00000048553*	*OR52R1*	0.95	0.90	0.40	0.35	0.93	0.38	0.55
chr2_43538627	na	T	C	downstream	*ENSGALG00000048965*		0.75	0.80	0.25	0.25	0.78	0.25	0.53
chr6_27327782	na	T	A	ncRNA_exonic	*ENSGALG00000049592*		0.60	0.85	0.10	0.10	0.73	0.10	0.63

chr2_43926924	na	A	G	downstream	*ENSGALG00000049942*	*gga-mir-7457*	0.85	0.70	0.15	0.35	0.78	0.25	0.53
chr4_43442934	rs314348884	C	G	ncRNA_splicing	*ENSGALG00000050160*		0.80	0.75	0.05	0.15	0.78	0.10	0.68
chr4_43434441	rs315414306	C	T	ncRNA_exonic	*ENSGALG00000050160*		0.80	0.70	0.05	0.25	0.75	0.15	0.60
chr6_25333139	na	G	A	ncRNA_intronic	*ENSGALG00000050593*		0.70	0.70	0.00	0.20	0.70	0.10	0.60
chr6_25331885	na	G	T	ncRNA_exonic	*ENSGALG00000050593*		0.65	0.70	0.00	0.20	0.68	0.10	0.58
chr3_107835173	rs317318272	A	G	nonsynonymous	*ENSGALG00000050753*		0.75	0.80	0.15	0.35	0.78	0.25	0.53
chr3_43733903	rs312669753	G	A	ncRNA_intronic	*ENSGALG00000051296*		0.80	0.90	0.25	0.20	0.85	0.23	0.63
chr6_27270359	rs312608923	G	A	ncRNA_exonic	*ENSGALG00000051688*		0.75	0.90	0.25	0.30	0.83	0.28	0.55
chr1_97184538	rs736739044	A	G	ncRNA_intronic	*ENSGALG00000051825*		0.75	0.55	0.10	0.15	0.65	0.13	0.53
chr5_17564981	rs741666407	T	G	ncRNA_intronic	*ENSGALG00000051988*		0.75	0.80	0.10	0.15	0.78	0.13	0.65
chr6_27321826	rs739968199	T	A	ncRNA_exonic	*ENSGALG00000052216*		0.60	0.85	0.10	0.10	0.73	0.10	0.63
chr4_43422571	rs734941354	T	C	ncRNA_splicing	*ENSGALG00000052572*		0.80	0.75	0.05	0.35	0.78	0.20	0.58
chr4_43422589	rs313243968	C	A	ncRNA_exonic	*ENSGALG00000052572*		0.80	0.75	0.05	0.35	0.78	0.20	0.58
chr14_5124954	rs732505121	G	A	downstream	*ENSGALG00000052641*	*PLD6*	0.70	0.80	0.10	0.15	0.75	0.13	0.63
chr4_52721011	rs312579587	C	T	ncRNA_intronic	*ENSGALG00000052662*		0.15	0.25	0.80	0.90	0.20	0.85	0.65
chr1_197280636	na	G	A	nonsynonymous	*ENSGALG00000052767*		0.95	0.80	0.45	0.20	0.88	0.33	0.55
chr1_197280967	na	T	C	nonsynonymous	*ENSGALG00000052767*		0.95	0.80	0.45	0.20	0.88	0.33	0.55
chr1_197280970	na	G	A	nonsynonymous	*ENSGALG00000052767*		0.95	0.80	0.45	0.20	0.88	0.33	0.55
chr1_37164801	rs315096302	T	A	ncRNA_exonic	*ENSGALG00000052790*		0.70	0.95	0.15	0.25	0.83	0.20	0.63
chr20_4587434	rs738582207	G	C	ncRNA_intronic	*ENSGALG00000052811*		0.95	0.90	0.30	0.30	0.93	0.30	0.63
chr2_42108399	rs14170918	C	T	ncRNA_intronic	*ENSGALG00000052910*		0.65	0.90	0.00	0.30	0.78	0.15	0.63
chr1_197231010	na	C	T	intronic	*ENSGALG00000053527*		0.95	0.80	0.25	0.15	0.88	0.20	0.68
chr5_17193485	rs315097836	A	C	UTR3	*ENSGALG00000053601*	*LPXN*	0.80	0.90	0.15	0.25	0.85	0.20	0.65
chr5_17191121	na	G	T	nonsynonymous	*ENSGALG00000053601*	*LPXN*	0.75	0.75	0.15	0.10	0.75	0.13	0.63
chr5_17190919	na	G	T	nonsynonymous	*ENSGALG00000053601*	*LPXN*	0.70	0.75	0.15	0.10	0.73	0.13	0.60
chr5_17191232	rs731490047	A	G	nonsynonymous	*ENSGALG00000053601*	*LPXN*	0.20	0.20	0.85	0.65	0.20	0.75	0.55
chr1_197252179	na	G	A	nonsynonymous	*ENSGALG00000053623*		0.85	0.75	0.20	0.10	0.80	0.15	0.65
chr3_43756481	rs315338151	T	C	ncRNA_exonic	*ENSGALG00000053691*		0.80	0.90	0.30	0.30	0.85	0.30	0.55
chr3_107854024	rs312965298	G	C	ncRNA_exonic	*ENSGALG00000053865*		0.75	0.85	0.15	0.40	0.80	0.28	0.53
chr2_43463850	rs316715125	T	G	intronic	*ENSGALG00000053882*		0.80	0.85	0.30	0.15	0.83	0.23	0.60
chr2_43924111	na	G	T	nonsynonymous	*ENSGALG00000054003*		0.80	0.70	0.05	0.40	0.75	0.23	0.53

chr1_197243456	na	C	T	downstream	*ENSGALG00000054014*		0.95	0.80	0.25	0.10	0.88	0.18	0.70
chr23_2186018	na	C	T	intronic	*ENSGALG00000055006*	*AHDC1*	0.85	0.85	0.25	0.20	0.85	0.23	0.63

Table S4.3b. Candidate SNP list from signature analysis for preSeasonality population group. (Only SNPs with large differences in allele frequency (dAAF>0.5) between Low and High groups are included. At least one SNP with dAAF>0.5 for each candidate gene is included. |dAAF|: difference in allele frequency between Low and High group).

snpID	**rs_id**	**ref**	**alt**	**annotation**	**gene**	**gene ID**	**Low group pop1 AAF**	**Low group pop2 AAF**	**High group pop1 AAF**	**High group pop2 AAF**	$\bar{\bar{X}}$ **AAF in Low group**	$\bar{\bar{X}}$ **AAF in High group**	**|dAAF|**
chr19_1762984	rs431878943	C	T	intronic	*ENSGALG00000001153*	*AUTS2*	0.94	0.70	0.50	0.11	0.82	0.31	0.52
chr23_4374321	rs318042596	A	G	intronic	*ENSGALG00000002346*	*CLSPN*	0.67	0.65	0.15	0.06	0.66	0.10	0.56
chr23_4372637	rs733769530	A	G	nonsynonymous	*ENSGALG00000002346*	*CLSPN*	0.67	0.70	0.30	0.00	0.68	0.15	0.53
chr11_1672079	rs14958343	A	G	intronic	*ENSGALG00000002458*	*VAC14*	0.22	0.30	0.80	1.00	0.26	0.90	0.64
chr22_2620465	na	C	T	intronic	*ENSGALG00000003299*	*NSD3*	0.50	0.85	0.10	0.00	0.68	0.05	0.63
chr6_12475073	rs740736721	A	G	intronic	*ENSGALG00000004569*	*UNC5B*	0.83	0.75	0.30	0.22	0.79	0.26	0.53
chr20_6577639	rs317369497	T	C	intronic	*ENSGALG00000004758*	*KCNB1*	0.78	0.70	0.15	0.11	0.74	0.13	0.61
chr4_2746828	rs317632014	A	T	intronic	*ENSGALG00000005853*	*HTR2C*	0.83	0.50	0.00	0.06	0.67	0.03	0.64
chr1_13966666	rs316266545	C	T	intronic	*ENSGALG00000008154*	*SRPK2*	0.06	0.35	0.75	0.78	0.20	0.76	0.56
chr1_13953265	rs317668098	A	G	intronic	*ENSGALG00000008167*	*KMT2E*	0.00	0.15	0.60	0.67	0.08	0.63	0.56
chr1_13950303	rs316302097	G	A	nonsynonymous	*ENSGALG00000008167*	*KMT2E*	0.00	0.15	0.60	0.56	0.08	0.58	0.50
chr1_12937368	rs313777231	A	T	intronic	*ENSGALG00000008301*	*FAM185A*	0.72	0.70	0.25	0.17	0.71	0.21	0.50
chr1_12920502	rs13831816	T	C	UTR3	*ENSGALG00000008306*	*FGL2*	0.00	0.15	0.65	0.72	0.08	0.69	0.61
chr1_12914788	rs739994118	G	A	intronic	*ENSGALG00000008309*		0.89	0.70	0.15	0.06	0.79	0.10	0.69
chr6_25058737	rs15800502	G	A	intronic	*ENSGALG00000008318*	*SLK*	0.44	0.35	0.90	0.94	0.40	0.92	0.53
chr3_5561217	rs315780117	G	A	intronic	*ENSGALG00000008725*	*KIF16B*	0.94	0.70	0.15	0.39	0.82	0.27	0.55
chr5_28923353	rs316888789	G	C	intronic	*ENSGALG00000009491*	*RAD51B*	0.78	0.75	0.40	0.11	0.76	0.26	0.51
chr3_20239718	rs741653393	C	T	intronic	*ENSGALG00000009645*	*ESRRG*	0.83	0.80	0.40	0.22	0.82	0.31	0.51
chr3_21969178	rs740864445	A	G	intronic	*ENSGALG00000009822*	*TMEM206*	0.56	0.60	0.00	0.00	0.58	0.00	0.58
chr4_30278252	rs316385750	C	T	intronic	*ENSGALG00000009880*	*INPP4B*	0.72	0.75	0.10	0.11	0.74	0.11	0.63
chr5_48776717	rs315241281	C	G	UTR5	*ENSGALG00000011228*	*WARS*	0.61	0.50	0.00	0.06	0.56	0.03	0.53

chr4_60959651	rs316313507	A	C	intronic	*ENSGALG00000012319*	*SLC9B2*	0.67	0.65	0.25	0.06	0.66	0.15	0.51
chr4_60979707	rs737134967	A	T	intronic	*ENSGALG00000012338*	*BDH2*	0.89	0.80	0.35	0.28	0.84	0.31	0.53
chr7_33207317	rs731995024	C	T	intronic	*ENSGALG00000012424*	*GTDC1*	0.78	0.80	0.25	0.17	0.79	0.21	0.58
chr7_35151474	rs317912233	G	A	intronic	*ENSGALG00000012484*	*RIF1*	0.17	0.10	0.80	0.72	0.13	0.76	0.63
chr7_35164267	rs16614439	T	G	nonsynonymous	*ENSGALG00000012484*	*RIF1*	0.06	0.05	0.60	0.67	0.05	0.63	0.58
chr7_36105598	rs313025143	A	G	intronic	*ENSGALG00000012543*	*GPD2*	0.22	0.30	0.75	0.83	0.26	0.79	0.53
chr2_69272916	rs314642495	G	C	intronic	*ENSGALG00000012917*	*CDH6*	0.78	0.80	0.30	0.11	0.79	0.21	0.58
chr3_70967885	rs736804304	C	T	intronic	*ENSGALG00000015466*	*SIM1*	0.17	0.20	0.65	0.72	0.18	0.69	0.50
chr3_80453177	rs314160172	C	T	intronic	*ENSGALG00000015898*	*MYO6*	0.44	0.35	0.90	0.94	0.40	0.92	0.53
chr4_89968685	rs737780359	A	T	intronic	*ENSGALG00000016098*	*EXOC6B*	0.06	0.45	0.80	0.78	0.25	0.79	0.54
chr3_86092927	rs314877271	T	C	intronic	*ENSGALG00000016276*	*KHDRBS2*	0.94	0.80	0.35	0.06	0.87	0.20	0.67
chr3_86692518	rs734012915	C	T	intronic	*ENSGALG00000016278*	*PRIM2*	0.72	0.80	0.15	0.06	0.76	0.10	0.66
chr3_86745844	rs730944091	A	T	intronic	*ENSGALG00000016279*	*RAB23*	0.11	0.25	0.60	0.94	0.18	0.77	0.59
chr1_178331570	rs314191305	A	T	intronic	*ENSGALG00000017106*	*ATP8A2*	0.94	0.70	0.35	0.28	0.82	0.31	0.51
chr2_94079426	rs741513456	G	T	intronic	*ENSGALG00000019509*		0.17	0.05	0.70	0.56	0.11	0.63	0.52
chr3_86764582	rs15417056	A	G	intronic	*ENSGALG00000026605*	*BAG2*	1.00	0.75	0.45	0.22	0.88	0.34	0.54
chr5_48793409	rs733995268	G	A	intronic	*ENSGALG00000028897*	*WDR25*	0.39	0.40	0.95	0.89	0.39	0.92	0.53
chr1_464307	rs733041205	G	A	ncRNA_intronic	*ENSGALG00000033919*		0.78	0.90	0.30	0.33	0.84	0.32	0.52
chr3_32612124	rs14336777	A	G	ncRNA_intronic	*ENSGALG00000035137*		0.83	0.70	0.15	0.06	0.77	0.10	0.66
chr7_35389628	rs317824661	T	A	intronic	*ENSGALG00000035466*	*CACNB4*	0.83	0.70	0.10	0.06	0.77	0.08	0.69
chr1_52409419	rs15267228	A	G	ncRNA_intronic	*ENSGALG00000036131*		0.78	0.75	0.15	0.17	0.76	0.16	0.61
chr6_12545812	rs736886375	G	A	intronic	*ENSGALG00000036824*	*CDH23*	0.67	0.65	0.15	0.00	0.66	0.08	0.58
chr9_12559730	rs13607342	T	C	downstream	*ENSGALG00000038512*	*AGTR1*	0.33	0.35	0.85	0.89	0.34	0.87	0.53
chr23_4365109	rs314218782	C	T	intronic	*ENSGALG00000039713*	*AGO4*	0.72	0.70	0.30	0.00	0.71	0.15	0.56
chr20_1104430	rs731765162	C	T	intronic	*ENSGALG00000040856*	*CPNE1*	0.17	0.20	0.80	0.72	0.18	0.76	0.58
chr7_33832456	rs315244744	A	G	ncRNA_intronic	*ENSGALG00000041782*		0.94	0.85	0.00	0.28	0.90	0.14	0.76
chr1_134822995	na	C	A	intronic	*ENSGALG00000045388*	*RFX8*	1.00	0.70	0.20	0.28	0.85	0.24	0.61
chr4_74231552	rs80710144	G	T	ncRNA_intronic	*ENSGALG00000046053*		0.22	0.35	0.90	0.89	0.29	0.89	0.61
chr2_148164284	rs314286815	A	G	ncRNA_intronic	*ENSGALG00000047330*		0.72	0.70	0.10	0.06	0.71	0.08	0.63
chr7_33829465	rs15886171	C	T	ncRNA_exonic	*ENSGALG00000047423*		0.11	0.10	0.70	0.56	0.11	0.63	0.52
chr3_50852764	rs16275293	A	C	ncRNA_intronic	*ENSGALG00000047502*		0.94	0.80	0.45	0.28	0.87	0.36	0.51
chr3_71652581	rs16298732	A	G	ncRNA_intronic	*ENSGALG00000047943*		0.78	0.80	0.25	0.17	0.79	0.21	0.58
chr3_71653988	rs739129986	C	T	ncRNA_exonic	*ENSGALG00000047943*		0.78	0.70	0.20	0.17	0.74	0.18	0.56
chr4_74184142	rs733020532	G	A	ncRNA_intronic	*ENSGALG00000048521*		0.11	0.10	0.65	0.78	0.11	0.71	0.61

chr4_74190773	rs14490489	G	A	ncRNA_exonic	*ENSGALG00000048521*		0.89	0.85	0.35	0.22	0.87	0.29	0.58
chr2_148244192	rs312881030	A	C	ncRNA_intronic	*ENSGALG00000048947*		0.17	0.05	0.65	0.94	0.11	0.80	0.69
chr2_148199210	rs314647706	G	T	ncRNA_exonic	*ENSGALG00000048947*		0.72	0.85	0.40	0.17	0.79	0.28	0.50
chr3_21964852	rs14325060	G	T	intronic	*ENSGALG00000049422*	*ATF3*	0.39	0.30	1.00	0.94	0.34	0.97	0.63
chr2_103509237	rs317545067	G	A	ncRNA_intronic	*ENSGALG00000049694*		0.89	0.85	0.35	0.28	0.87	0.31	0.56
chr4_74218577	na	T	C	ncRNA_exonic	*ENSGALG00000050078*		0.17	0.35	0.90	0.89	0.26	0.89	0.64
chr1_164032291	rs741501392	T	C	ncRNA_intronic	*ENSGALG00000050284*		0.39	0.30	0.90	0.94	0.34	0.92	0.58
chr2_95916659	rs318197480	T	A	ncRNA_intronic	*ENSGALG00000050654*		0.33	0.30	0.90	0.94	0.32	0.92	0.61
chr5_28955788	rs313814136	A	G	ncRNA_intronic	*ENSGALG00000051208*		0.89	0.90	0.50	0.22	0.89	0.36	0.53
chr4_74290857	rs741369370	A	G	ncRNA_exonic	*ENSGALG00000051573*		0.22	0.25	0.65	0.83	0.24	0.74	0.51
chr3_18289928	rs15283236	T	G	downstream	*ENSGALG00000052742*		0.67	0.90	0.15	0.17	0.78	0.16	0.63
chr9_11423609	rs14661003	G	C	ncRNA_intronic	*ENSGALG00000052830*		0.94	0.90	0.60	0.22	0.92	0.41	0.51
chr3_71643736	rs16298716	A	G	downstream	*ENSGALG00000052916*		0.44	0.40	0.95	0.94	0.42	0.95	0.53
chr2_114061204	rs80744771	C	T	ncRNA_intronic	*ENSGALG00000053888*		0.22	0.20	0.75	0.72	0.21	0.74	0.53
chr2_114062296	rs316490734	G	A	ncRNA_exonic	*ENSGALG00000053888*		0.72	0.75	0.15	0.28	0.74	0.21	0.52
chr13_7328547	rs317710966	G	C	ncRNA_intronic	*ENSGALG00000054885*		0.44	0.60	0.00	0.00	0.52	0.00	0.52

Table S4.3c. Candidate SNP list from signature analysis for precWQ population group. (Only SNPs with large differences in allele frequency (dAAF>0.5) between Low and High groups are included. At least one SNP with dAAF>0.5 for each candidate gene is included. |dAAF|: difference in allele frequency between Low and High group).

snpID	**rs_id**	**ref**	**alt**	**annotation**	**gene**	**gene ID**	**Low group pop1 AAF**	**Low group pop2 AAF**	**High group pop1 AAF**	**High group pop2 AAF**	**$\bar{\bar{X}}$ AAF in Low group**	**$\bar{\bar{X}}$ AAF in High group**	**|dAAF|**
chr27_4836254	rs733495695	T	C	intronic	*ENSGALG00000000461*	*MRC2*	0.15	0.32	0.55	1.00	0.24	0.78	0.54
chr17_7815584	rs732528570	G	A	intronic	*ENSGALG00000002546*	*COL5A1*	0.75	0.86	0.00	0.06	0.80	0.03	0.78
chr15_4938506	rs316673516	G	A	intronic	*ENSGALG00000003120*	*NCOR2*	0.30	0.25	0.95	0.94	0.28	0.95	0.67
chr23_5596146	rs314063519	G	T	intronic	*ENSGALG00000003693*	*MACF1*	0.90	0.86	0.00	0.06	0.88	0.03	0.85
chr23_5558718	rs316933203	A	G	nonsynonymous	*ENSGALG00000003693*	*MACF1*	0.90	0.82	0.25	0.17	0.86	0.21	0.65
chr23_5551243	rs314855848	G	A	nonsynonymous	*ENSGALG00000003693*	*MACF1*	0.70	0.82	0.15	0.17	0.76	0.16	0.60
chr14_4418236	na	G	A	ncRNA_exonic	*ENSGALG00000004504*		0.35	0.07	0.70	0.83	0.21	0.77	0.56

chr14_4435770	rs733978568	C	T	downstream	*ENSGALG00000004530*	*MMD2*	0.25	0.39	0.80	0.89	0.32	0.84	0.52
chr8_10619626	na	G	A	intronic	*ENSGALG00000005065*	*PLA2G4A*	0.15	0.21	0.80	0.83	0.18	0.82	0.63
chr11_16304785	rs316253144	T	C	intronic	*ENSGALG00000005483*	*CDH13*	0.15	0.21	0.85	0.89	0.18	0.87	0.69
chr19_8780087	rs737099702	A	G	intronic	*ENSGALG00000005489*	*DDX52*	0.90	0.96	0.35	0.17	0.93	0.26	0.67
chr19_8785804	rs1058285452	T	A	nonsynonymous	*ENSGALG00000005489*	*DDX52*	0.90	0.93	0.40	0.17	0.91	0.28	0.63
chr19_8784622	rs431897287	T	C	nonsynonymous	*ENSGALG00000005489*	*DDX52*	0.90	1.00	0.50	0.33	0.95	0.42	0.53
chr19_8795246	rs733764249	C	T	intronic	*ENSGALG00000005504*	*HNF1B*	0.95	0.93	0.35	0.17	0.94	0.26	0.68
chr15_7961739	rs732175619	G	A	intronic	*ENSGALG00000005676*	*TTC28*	0.10	0.29	0.90	0.78	0.19	0.84	0.65
chr15_7970153	rs736876872	T	A	intronic	*ENSGALG00000005695*	*CHEK2*	0.10	0.21	0.80	0.78	0.16	0.79	0.63
chr2_8169684	rs15060462	G	A	intronic	*ENSGALG00000006379*	*SHH*	0.00	0.11	0.70	0.67	0.05	0.68	0.63
chr24_2954062	rs312753491	G	A	intronic	*ENSGALG00000006494*	*GRAMD1B*	0.80	0.64	0.20	0.00	0.72	0.10	0.62
chr15_10657436	rs317784425	G	A	intronic	*ENSGALG00000007732*	*RTN4R*	0.15	0.18	0.60	1.00	0.16	0.80	0.64
chr4_15366691	rs313896728	A	T	intronic	*ENSGALG00000008442*	*TENM1*	0.95	0.75	0.05	0.11	0.85	0.08	0.77
chr2_21986521	rs15072919	A	G	intronic	*ENSGALG00000009062*	*CDK14*	0.95	0.93	0.15	0.11	0.94	0.13	0.81
chr5_29744386	rs740952797	C	T	intronic	*ENSGALG00000009619*	*GPR176*	0.10	0.11	0.80	0.89	0.10	0.84	0.74
chr4_25200009	rs16377041	C	T	downstream	*ENSGALG00000009650*		0.05	0.07	0.60	0.56	0.06	0.58	0.52
chr4_25219821	rs314617422	C	T	upstream	*ENSGALG00000009655*		0.85	0.79	0.25	0.22	0.82	0.24	0.58
chr4_25500060	rs735396688	G	A	intronic	*ENSGALG00000009692*	*SH3RF1*	0.85	0.71	0.20	0.11	0.78	0.16	0.63
chr4_25520137	na	A	G	nonsynonymous	*ENSGALG00000009692*	*SH3RF1*	0.85	0.64	0.20	0.11	0.75	0.16	0.59
chr4_25522813	rs736520830	G	A	nonsynonymous	*ENSGALG00000009692*	*SH3RF1*	0.85	0.68	0.25	0.17	0.76	0.21	0.56
chr4_25548208	rs733070316	C	T	intronic	*ENSGALG00000009714*	*CBR4*	0.85	0.75	0.10	0.11	0.80	0.11	0.69
chr4_29853731	rs739130538	T	C	intronic	*ENSGALG00000009865*	*RNF150*	0.55	0.86	0.05	0.28	0.70	0.16	0.54
chr4_31964457	rs740688037	A	G	UTR3	*ENSGALG00000010019*	*TMEM184C*	0.80	0.75	0.20	0.22	0.78	0.21	0.56
chr4_31970860	rs741235690	A	C	intronic	*ENSGALG00000010022*	*PRMT9*	1.00	0.82	0.25	0.28	0.91	0.26	0.65
chr4_31967765	rs741489346	G	A	nonsynonymous	*ENSGALG00000010022*	*PRMT9*	0.80	0.75	0.25	0.28	0.78	0.26	0.51
chr1_38032780	rs317828600	T	C	intronic	*ENSGALG00000010238*	*NAP1L1*	0.15	0.32	0.90	0.78	0.24	0.84	0.60
chr5_39346834	rs15706986	C	T	UTR3	*ENSGALG00000010406*	*TMEM63C*	0.20	0.00	0.80	0.61	0.10	0.71	0.61
chr5_39348173	rs313519656	G	A	UTR3	*ENSGALG00000010413*	*NGB*	0.15	0.00	0.60	0.56	0.08	0.58	0.50
chr8_22207096	rs737316662	A	T	upstream	*ENSGALG00000010444*	*ATPAF1*	0.85	0.64	0.25	0.22	0.75	0.24	0.51
chr5_43322206	na	G	C	upstream	*ENSGALG00000010616*	*PTPN21*	0.05	0.04	0.55	0.83	0.04	0.69	0.65
chr5_43344947	rs740498423	A	G	intronic	*ENSGALG00000010622*	*ZC3H14*	0.20	0.18	0.65	0.78	0.19	0.71	0.52
chr4_39436369	rs316883178	A	G	intronic	*ENSGALG00000010628*	*ACSL1*	0.60	0.82	0.05	0.06	0.71	0.05	0.66
chr5_43358208	rs313510958	T	A	intronic	*ENSGALG00000010639*	*EML5*	0.80	0.64	0.20	0.22	0.72	0.21	0.51
chr8_27970038	rs313589909	T	C	downstream	*ENSGALG00000010985*		0.90	0.75	0.30	0.17	0.83	0.23	0.59

chr8_28337013	rs732083903	G	A	intronic	*ENSGALG00000011019*	*ROR1*	0.70	0.64	0.00	0.17	0.67	0.08	0.59
chr2_37133230	rs731492166	A	G	intronic	*ENSGALG00000011289*	*NKIRAS1*	0.15	0.21	0.60	0.78	0.18	0.69	0.51
chr2_37138591	rs14165315	G	A	UTR3	*ENSGALG00000011290*	*RPL15*	0.15	0.43	0.80	0.94	0.29	0.87	0.58
chr1_49385771	rs316254574	C	T	intronic	*ENSGALG00000011878*	*TCF20*	0.10	0.11	0.85	0.78	0.10	0.81	0.71
chr7_30451014	rs316457934	C	T	intronic	*ENSGALG00000012197*	*ACMSD*	0.25	0.00	0.60	0.72	0.13	0.66	0.54
chr4_59790747	rs316072128	G	A	intronic	*ENSGALG00000012227*	*TSPAN5*	0.70	0.00	0.85	0.89	0.35	0.87	0.52
chr2_61180106	rs735904880	C	T	intronic	*ENSGALG00000012702*	*JARID2*	0.65	0.96	0.20	0.11	0.81	0.16	0.65
chr1_55789175	rs737764409	A	G	intronic	*ENSGALG00000012791*	*TBXAS1*	0.50	0.18	0.80	0.94	0.34	0.87	0.53
chr2_65646408	rs316485261	G	A	intronic	*ENSGALG00000012808*		0.95	0.96	0.45	0.44	0.96	0.45	0.51
chr2_68093262	rs314522770	C	T	intronic	*ENSGALG00000012888*	*PHLPP1*	0.95	0.75	0.10	0.00	0.85	0.05	0.80
chr2_83534670	rs731782657	C	T	intronic	*ENSGALG00000013129*	*RPRD1A*	0.95	0.82	0.15	0.17	0.89	0.16	0.73
chr4_65495024	rs733193951	C	T	intronic	*ENSGALG00000013931*	*CHIC2*	0.65	0.61	0.20	0.06	0.63	0.13	0.50
chr4_89983847	rs315686852	A	G	intronic	*ENSGALG00000016098*	*EXOC6B*	0.80	0.54	0.00	0.00	0.67	0.00	0.67
chr3_87865731	rs315425423	G	C	downstream	*ENSGALG00000016300*	*TINAG*	0.95	0.64	0.35	0.17	0.80	0.26	0.54
chr14_4402914	na	T	G	downstream	*ENSGALG00000025659*	*gga-mir-2129*	0.25	0.00	0.70	0.89	0.13	0.79	0.67
chr13_2584104	rs317675790	G	A	intronic	*ENSGALG00000028565*	*PCDHGA2*	1.00	0.96	0.20	0.28	0.98	0.24	0.74
chr2_41885078	rs739064554	G	A	intronic	*ENSGALG00000029235*	*CPNE4*	0.20	0.11	0.55	0.89	0.15	0.72	0.57
chr2_136099676	rs317956773	C	A	intronic	*ENSGALG00000029664*	*SAMD12*	0.15	0.25	0.80	0.89	0.20	0.84	0.64
chr4_89589419	rs14505596	T	A	intronic	*ENSGALG00000030602*	*ADAM33*	1.00	0.89	0.15	0.06	0.95	0.10	0.84
chr10_20309795	rs740025971	G	C	intronic	*ENSGALG00000030858*	*FRMD5*	0.55	0.43	1.00	1.00	0.49	1.00	0.51
chr2_135788002	rs316552794	C	A	intronic	*ENSGALG00000030862*	*MED30*	0.05	0.11	0.60	0.72	0.08	0.66	0.58
chr9_8048225	rs736036381	G	A	intronic	*ENSGALG00000030944*	*Pax3*	0.85	0.75	0.00	0.06	0.80	0.03	0.77
chr2_42583923	rs15966693	T	C	downstream	*ENSGALG00000032316*	*CDV3*	0.55	0.11	0.80	0.89	0.33	0.84	0.52
chr9_20261447	rs317252309	A	G	intronic	*ENSGALG00000032577*	*MECOM*	0.50	0.89	0.15	0.11	0.70	0.13	0.57
chr22_5298163	na	C	G	intronic	*ENSGALG00000032656*	*TCF7L1*	0.30	0.18	0.75	0.83	0.24	0.79	0.55
chr2_42592533	rs312420170	C	T	intronic	*ENSGALG00000034461*	*TOPBP1*	0.40	0.21	0.75	0.89	0.31	0.82	0.51
chr1_38022988	rs312921343	C	T	intronic	*ENSGALG00000034788*	*PHLDA1*	0.10	0.14	0.85	0.72	0.12	0.79	0.66
chr8_23371814	rs313293937	G	C	intronic	*ENSGALG00000036909*	*BEND5*	0.75	0.75	0.05	0.33	0.75	0.19	0.56
chr17_2822458	rs1060013588	G	C	nonsynonymous	*ENSGALG00000037031*		0.30	0.00	0.60	0.78	0.15	0.69	0.54
chr1_20448549	rs732568414	T	C	upstream	*ENSGALG00000038467*	*SELENOO*	0.00	0.00	0.45	0.61	0.00	0.53	0.53
chr23_5607357	rs737930251	T	A	intronic	*ENSGALG00000038984*		0.75	0.68	0.05	0.11	0.71	0.08	0.63
chr23_5608924	rs732739090	C	T	nonsynonymous	*ENSGALG00000038984*		0.75	0.64	0.05	0.11	0.70	0.08	0.62
chr2_135873518	rs733107179	A	C	downstream	*ENSGALG00000039013*	*EXT1*	0.30	0.39	0.80	0.94	0.35	0.87	0.53

chr1_133250883	rs316643536	A	G	intronic	*ENSGALG00000039826*	*CNGA3*	0.90	0.54	0.15	0.17	0.72	0.16	0.56
chr17_2820193	rs1058702762	A	G	downstream	*ENSGALG00000040003*	*STXBP1*	1.00	0.96	0.50	0.39	0.98	0.44	0.54
chr10_13683086	rs733336439	G	A	intronic	*ENSGALG00000040241*	*NTRK3*	0.15	0.14	0.80	0.83	0.15	0.82	0.67
chr2_135422594	rs15157999	A	T	intronic	*ENSGALG00000040522*	*EIF3H*	0.05	0.07	0.70	0.61	0.06	0.66	0.59
chr14_4404742	rs15006315	T	C	upstream	*ENSGALG00000040528*	*AP5Z1*	0.75	1.00	0.30	0.11	0.88	0.21	0.67
chr14_4410232	rs313746392	A	T	nonsynonymous	*ENSGALG00000040528*	*AP5Z1*	0.35	0.04	0.70	0.89	0.19	0.79	0.60
chr4_1113291	rs740662094	C	T	intronic	*ENSGALG00000041437*	*FAM155B*	0.90	0.64	0.30	0.11	0.77	0.21	0.57
chr13_2950064	rs315966597	G	A	ncRNA_exonic	*ENSGALG00000042617*		0.65	0.82	0.00	0.00	0.74	0.00	0.74
chr8_10354829	rs15909685	G	A	ncRNA_exonic	*ENSGALG00000043258*		0.15	0.29	0.75	0.78	0.22	0.76	0.55
chr8_10356392	rs312992425	G	A	ncRNA_exonic	*ENSGALG00000043258*		0.15	0.29	0.75	0.78	0.22	0.76	0.55
chr4_89598953	rs314488388	T	C	intronic	*ENSGALG00000044620*		0.00	0.14	0.50	0.83	0.07	0.67	0.60
chr8_10598175	rs737176813	C	T	nonsynonymous	*ENSGALG00000047029*		0.15	0.29	0.80	0.83	0.22	0.82	0.60
chr8_10598177	na	C	T	nonsynonymous	*ENSGALG00000047029*		0.15	0.29	0.80	0.83	0.22	0.82	0.60
chr8_27984732	rs316843375	A	C	ncRNA_exonic	*ENSGALG00000047521*		0.95	0.86	0.30	0.50	0.90	0.40	0.50
chr14_4400041	rs15006307	G	T	upstream	*ENSGALG00000047899*	*gga-mir-12224*	0.75	0.79	0.25	0.06	0.77	0.15	0.62
chr3_21652067	rs14324669	T	C	ncRNA_intronic	*ENSGALG00000047938*		0.10	0.21	0.75	0.94	0.16	0.85	0.69
chr2_58713136	rs734679973	G	A	ncRNA_intronic	*ENSGALG00000048667*		0.75	0.82	0.20	0.28	0.79	0.24	0.55
chr5_39324615	rs738864593	A	T	ncRNA_intronic	*ENSGALG00000048945*		0.75	0.86	0.30	0.22	0.80	0.26	0.54
chr5_39265317	rs317766185	T	G	ncRNA_exonic	*ENSGALG00000049515*		0.15	0.04	0.75	0.61	0.09	0.68	0.59
chr4_4572234	rs739047199	A	G	ncRNA_intronic	*ENSGALG00000049919*		0.25	0.00	0.80	0.83	0.13	0.82	0.69
chr4_4589945	na	G	T	ncRNA_exonic	*ENSGALG00000049919*		0.15	0.04	0.60	0.83	0.09	0.72	0.62
chr3_87830267	rs15419986	T	G	upstream	*ENSGALG00000049936*		0.90	0.96	0.30	0.00	0.93	0.15	0.78
chr3_87824322	rs740528692	G	A	ncRNA_intronic	*ENSGALG00000049936*		0.90	0.86	0.30	0.06	0.88	0.18	0.70
chr4_39419074	rs733655018	C	T	ncRNA_exonic	*ENSGALG00000050111*		0.30	0.04	0.95	0.89	0.17	0.92	0.75
chr4_39419075	rs738419841	G	A	ncRNA_exonic	*ENSGALG00000050111*		0.30	0.04	0.95	0.89	0.17	0.92	0.75
chr2_65520806	rs16024288	A	G	intronic	*ENSGALG00000051665*	*LYRM4*	0.20	0.18	0.90	0.72	0.19	0.81	0.62
chr13_2951147	rs317492478	A	G	ncRNA_intronic	*ENSGALG00000051725*		0.35	0.14	1.00	1.00	0.25	1.00	0.75
chr1_150082655	rs315667921	T	C	upstream	*ENSGALG00000052317*		0.30	0.07	0.70	0.78	0.19	0.74	0.55
chr2_83535517	rs738834479	A	G	ncRNA_exonic	*ENSGALG00000052392*		0.95	0.82	0.10	0.17	0.89	0.13	0.75
chr2_83535712	rs315559938	G	C	ncRNA_exonic	*ENSGALG00000052392*		0.95	0.82	0.10	0.17	0.89	0.13	0.75
chr1_99146493	rs318136054	G	T	ncRNA_intronic	*ENSGALG00000052669*		0.85	1.00	0.30	0.17	0.93	0.23	0.69
chr1_99139549	rs730904570	A	G	ncRNA_exonic	*ENSGALG00000052669*		0.85	1.00	0.40	0.17	0.93	0.28	0.64
chr3_21692302	rs736698483	T	A	ncRNA_intronic	*ENSGALG00000052926*		0.05	0.29	0.85	0.78	0.17	0.81	0.65

chr13_1290501	na	C	G	upstream	*ENSGALG00000052992*		1.00	0.93	0.50	0.22	0.96	0.36	0.60
chr10_17834159	rs735955190	T	G	ncRNA_intronic	*ENSGALG00000053260*		0.05	0.29	0.65	0.89	0.17	0.77	0.60
chr8_27976851	rs732426878	G	T	ncRNA_intronic	*ENSGALG00000053679*		0.95	0.89	0.35	0.28	0.92	0.31	0.61
chr13_2979931	rs14057329	G	A	ncRNA_exonic	*ENSGALG00000054361*		0.65	0.79	0.20	0.11	0.72	0.16	0.56
chr1_83105496	rs738446375	C	T	ncRNA_intronic	*ENSGALG00000054748*		0.15	0.07	0.90	0.89	0.11	0.89	0.78

Table S4.3d. Candidate SNP list from signature analysis for precDQ population group. (Only SNPs with large differences in allele frequency (dAAF>0.5) between Low and High groups are included. At least one SNP with dAAF>0.5 for each candidate gene is included. |dAAF|: difference in allele frequency between Low and High group).

snpID	**rs_id**	**ref**	**alt**	**annotation**	**gene**	**gene ID**	**Low group pop1 AAF**	**Low group pop2 AAF**	**High group pop1 AAF**	**High group pop2 AAF**	**$\bar{\bar{X}}$ AAF in Low group**	**$\bar{\bar{X}}$ AAF in High group**	**|dAAF|**
chr21_544526	rs740716939	G	A	intronic	*ENSGALG00000000627*	*DNAJC11*	0.78	0.72	0.15	0.28	0.75	0.21	0.54
chr19_1847689	rs316908858	A	G	intronic	*ENSGALG00000001153*	*AUTS2*	0.83	0.83	0.30	0.17	0.83	0.23	0.60
chr17_9453975	rs738022500	A	G	intronic	*ENSGALG00000001157*	*DENND1A*	0.78	0.83	0.30	0.17	0.81	0.23	0.57
chr26_3951479	na	A	C	UTR3	*ENSGALG00000002170*		0.83	0.78	0.15	0.44	0.81	0.30	0.51
chr11_1675064	rs735405883	G	A	intronic	*ENSGALG00000002458*	*VAC14*	0.94	0.83	0.35	0.22	0.89	0.29	0.60
chr22_2620465	na	C	T	intronic	*ENSGALG00000003299*	*NSD3*	0.00	0.33	0.85	0.50	0.17	0.68	0.51
chr11_18122306	rs317389472	G	C	intronic	*ENSGALG00000005872*	*BANP*	0.72	0.89	0.15	0.22	0.81	0.19	0.62
chr2_5471305	rs318104934	T	C	intronic	*ENSGALG00000006072*		0.72	0.94	0.15	0.17	0.83	0.16	0.67
chr6_21873025	na	C	T	intronic	*ENSGALG00000006912*	*CPEB3*	0.56	0.89	0.05	0.06	0.72	0.05	0.67
chr18_9863048	rs314018412	C	T	UTR3	*ENSGALG00000007191*	*TK1*	0.67	0.67	0.20	0.11	0.67	0.16	0.51
chr6_25055476	rs16555443	G	C	intronic	*ENSGALG00000008318*	*SLK*	0.06	0.06	0.65	0.61	0.06	0.63	0.58
chr3_8446794	rs314939603	A	G	intronic	*ENSGALG00000009078*	*GALNT14*	0.67	0.83	0.15	0.06	0.75	0.10	0.65
chr4_23364554	rs314992123	C	T	intronic	*ENSGALG00000009499*	*NPY5R*	0.50	0.78	0.10	0.06	0.64	0.08	0.56
chr4_23415547	rs314607081	G	A	intronic	*ENSGALG00000009501*	*TMA16*	0.72	0.83	0.25	0.17	0.78	0.21	0.57
chr8_20453398	rs732469649	G	A	intronic	*ENSGALG00000010053*	*PTPRF*	0.78	0.72	0.20	0.17	0.75	0.18	0.57
chr4_35188945	rs14451188	G	A	intronic	*ENSGALG00000010371*	*FAM13A*	0.89	0.94	0.35	0.39	0.92	0.37	0.55
chr4_43282141	rs312368233	G	T	intronic	*ENSGALG00000010744*	*GALNT7*	0.39	0.00	0.70	0.72	0.19	0.71	0.52
chr2_28015850	rs317484297	G	A	intronic	*ENSGALG00000010792*	*AGMO*	0.22	0.17	0.65	0.83	0.19	0.74	0.55

chr2_28021322	rs315829959	T	C	nonsynonymous	*ENSGALG00000010792*	*AGMO*	0.11	0.11	0.60	0.67	0.11	0.63	0.52
chr2_28014205	rs15077551	C	T	nonsynonymous	*ENSGALG00000010792*	*AGMO*	0.00	0.00	0.35	0.67	0.00	0.51	0.51
chr2_28067993	rs738289988	T	A	intronic	*ENSGALG00000010794*	*MEOX2*	0.11	0.00	0.50	0.78	0.06	0.64	0.58
chr8_26059205	rs732929160	C	T	intronic	*ENSGALG00000010818*	*PLPP3*	0.72	0.83	0.15	0.39	0.78	0.27	0.51
chr7_31045601	rs314250782	C	G	intronic	*ENSGALG00000012362*	*THSD7B*	0.22	0.22	0.70	0.78	0.22	0.74	0.52
chr7_35151474	rs317912233	G	A	intronic	*ENSGALG00000012484*	*RIF1*	0.72	0.83	0.10	0.17	0.78	0.13	0.64
chr7_35164267	rs16614439	T	G	nonsynonymous	*ENSGALG00000012484*	*RIF1*	0.67	0.67	0.05	0.06	0.67	0.05	0.61
chr7_36101421	rs736081272	G	A	intronic	*ENSGALG00000012543*	*GPD2*	0.56	0.56	0.00	0.00	0.56	0.00	0.56
chr1_52807448	rs15268225	A	T	intronic	*ENSGALG00000012559*	*LARGE1*	0.44	0.17	0.80	0.89	0.31	0.84	0.54
chr3_49703837	rs314052498	C	T	upstream	*ENSGALG00000013607*	*FBXO5*	0.72	0.78	0.30	0.17	0.75	0.23	0.52
chr3_49707564	rs315686448	A	G	intronic	*ENSGALG00000013609*	*MTRF1L*	0.72	0.78	0.25	0.17	0.75	0.21	0.54
chr1_78066038	rs313409191	T	C	upstream	*ENSGALG00000014659*	*C1R*	0.22	0.11	0.65	0.78	0.17	0.71	0.55
chr1_78077393	rs317830961	C	T	intronic	*ENSGALG00000014674*	*CLSTN3*	0.17	0.28	0.85	0.89	0.22	0.87	0.65
chr1_81243475	rs736091251	T	G	intronic	*ENSGALG00000015018*	*CASQ2*	0.00	0.00	0.45	0.67	0.00	0.56	0.56
chr2_103792728	rs16089132	G	A	intronic	*ENSGALG00000015112*	*ZNF521*	0.33	0.17	0.80	0.83	0.25	0.82	0.57
chr1_88329905	rs316074735	A	G	intronic	*ENSGALG00000015353*	*BBX*	0.11	0.06	0.65	0.56	0.08	0.60	0.52
chr3_70961225	rs739469705	C	T	intronic	*ENSGALG00000015466*	*SIM1*	0.61	0.83	0.25	0.11	0.72	0.18	0.54
chr3_72505566	rs315786802	T	C	intronic	*ENSGALG00000015513*	*MMS22L*	0.11	0.06	0.70	0.56	0.08	0.63	0.54
chr3_72517241	rs314869658	C	A	nonsynonymous	*ENSGALG00000015513*	*MMS22L*	0.72	0.83	0.30	0.22	0.78	0.26	0.52
chr3_73187113	rs15389516	A	T	upstream	*ENSGALG00000015590*	*MANEA*	0.83	0.78	0.10	0.17	0.81	0.13	0.67
chr3_73938509	rs740881829	A	C	intronic	*ENSGALG00000015593*	*EPHA7*	0.89	0.89	0.25	0.33	0.89	0.29	0.60
chr4_89961208	rs314008162	G	A	intronic	*ENSGALG00000016098*	*EXOC6B*	0.72	0.78	0.30	0.00	0.75	0.15	0.60
chr1_109721134	rs14871885	C	T	intronic	*ENSGALG00000016138*	*DSCAM*	0.06	0.06	0.65	0.50	0.06	0.58	0.52
chr1_113714276	rs733028159	G	A	intronic	*ENSGALG00000016245*	*BCOR*	0.94	1.00	0.35	0.50	0.97	0.43	0.55
chr3_86703122	rs315677415	G	T	intronic	*ENSGALG00000016278*	*PRIM2*	0.72	0.61	0.00	0.06	0.67	0.03	0.64
chr3_86710637	rs313188681	T	C	nonsynonymous	*ENSGALG00000016278*	*PRIM2*	0.28	0.22	0.80	0.72	0.25	0.76	0.51
chr3_86754389	rs315833014	C	A	intronic	*ENSGALG00000016279*	*RAB23*	0.17	0.39	0.75	0.94	0.28	0.85	0.57
chr3_86751585	rs317473946	T	C	nonsynonymous	*ENSGALG00000016279*	*RAB23*	0.11	0.22	0.60	0.78	0.17	0.69	0.52
chr3_86786031	rs733735638	G	C	intronic	*ENSGALG00000016282*	*ZNF451*	0.72	0.72	0.15	0.06	0.72	0.10	0.62
chr3_87511307	rs315455718	C	T	intronic	*ENSGALG00000016294*	*HMGCLL1*	0.17	0.06	0.80	0.78	0.11	0.79	0.68
chr8_15400113	rs735472125	C	T	intronic	*ENSGALG00000020805*	*LRRC8D*	0.61	0.94	0.20	0.22	0.78	0.21	0.57
chr8_15392581	rs10724918	T	A	nonsynonymous	*ENSGALG00000020805*	*LRRC8D*	0.56	0.83	0.15	0.22	0.69	0.19	0.51
chr3_73206656	rs312336006	A	G	upstream	*ENSGALG00000025686*	*RF00026*	0.83	0.78	0.25	0.17	0.81	0.21	0.60
chr17_9508178	rs315550755	T	C	downstream	*ENSGALG00000026417*	*RF02271*	0.78	0.83	0.30	0.22	0.81	0.26	0.54

chr3_86764582	rs15417056	A	G	intronic	*ENSGALG00000026605*	*BAG2*	0.22	0.33	0.75	1.00	0.28	0.88	0.60
chr3_106484531	rs739397828	G	T	intronic	*ENSGALG00000028032*	*MSRA*	0.78	0.83	0.10	0.11	0.81	0.11	0.70
chr21_560325	rs10725890	A	G	UTR3	*ENSGALG00000028148*	*PHF13*	0.22	0.28	0.85	0.67	0.25	0.76	0.51
chr6_21704316	na	C	T	upstream	*ENSGALG00000030597*	*KIF11*	0.39	0.83	0.10	0.11	0.61	0.11	0.51
chr19_949859	rs741125990	C	T	intronic	*ENSGALG00000032220*		0.61	0.78	0.10	0.11	0.69	0.11	0.59
chr6_21804826	na	C	T	intronic	*ENSGALG00000032234*	*MARCH5*	0.50	0.94	0.05	0.11	0.72	0.08	0.64
chr21_538509	rs13602412	A	T	intronic	*ENSGALG00000033569*		0.78	0.72	0.15	0.22	0.75	0.19	0.56
chr3_72056954	na	C	T	downstream	*ENSGALG00000034564*		0.56	0.83	0.20	0.17	0.69	0.18	0.51
chr21_566581	rs315528106	C	T	intronic	*ENSGALG00000035211*	*KLHL21*	0.22	0.28	0.85	0.67	0.25	0.76	0.51
chr7_35393566	rs317060282	G	A	intronic	*ENSGALG00000035466*	*CACNB4*	0.89	0.72	0.15	0.11	0.81	0.13	0.68
chr2_46670986	rs731926865	A	G	downstream	*ENSGALG00000035505*	*AOAH*	0.28	0.22	0.80	0.78	0.25	0.79	0.54
chr3_72397353	rs316455543	G	A	ncRNA_intronic	*ENSGALG00000036204*		0.89	0.83	0.25	0.17	0.86	0.21	0.65
chr9_12548812	rs738244929	G	A	intronic	*ENSGALG00000038512*	*AGTR1*	0.00	0.06	0.50	0.56	0.03	0.53	0.50
chr23_4367774	rs731167823	C	T	intronic	*ENSGALG00000039713*	*AGO4*	0.00	0.39	0.70	0.72	0.19	0.71	0.52
chr1_110363154	rs794211905	G	A	upstream	*ENSGALG00000040189*	*RIPK4*	0.67	0.83	0.10	0.17	0.75	0.13	0.62
chr13_17881220	na	C	T	intronic	*ENSGALG00000040594*	*GNPDA1*	0.50	0.94	0.15	0.22	0.72	0.19	0.54
chr20_1104292	rs731212396	G	A	intronic	*ENSGALG00000040856*	*CPNE1*	0.78	0.83	0.20	0.28	0.81	0.24	0.57
chr3_73013447	rs313632779	C	T	intronic	*ENSGALG00000042116*	*FUT9*	0.78	0.89	0.10	0.22	0.83	0.16	0.67
chr13_17872778	na	C	T	intronic	*ENSGALG00000043085*	*RNF14*	0.50	0.94	0.20	0.22	0.72	0.21	0.51
chr2_46680571	rs741033568	C	T	intronic	*ENSGALG00000043642*	*ANLN*	0.50	0.78	0.05	0.11	0.64	0.08	0.56
chr13_17889787	na	G	A	intronic	*ENSGALG00000043744*	*NDFIP1*	0.50	0.94	0.20	0.22	0.72	0.21	0.51
chr1_134823969	rs314518705	T	G	intronic	*ENSGALG00000045388*	*RFX8*	0.67	0.50	0.15	0.00	0.58	0.08	0.51
chr4_74251752	rs737719175	A	G	ncRNA_intronic	*ENSGALG00000046053*		0.72	0.72	0.10	0.11	0.72	0.11	0.62
chr4_23352860	rs16374605	A	G	ncRNA_intronic	*ENSGALG00000046688*		0.50	0.22	0.85	0.94	0.36	0.90	0.54
chr12_18106384	rs316127096	G	A	ncRNA_intronic	*ENSGALG00000046751*		0.94	0.94	0.60	0.22	0.94	0.41	0.53
chr3_73342074	rs315272289	A	T	ncRNA_intronic	*ENSGALG00000046901*		0.72	0.83	0.20	0.17	0.78	0.18	0.59
chr3_74469948	rs15390883	C	T	ncRNA_exonic	*ENSGALG00000047305*		0.56	0.83	0.20	0.11	0.69	0.16	0.54
chr3_71654227	rs734579716	T	C	ncRNA_exonic	*ENSGALG00000047943*		0.83	0.78	0.20	0.22	0.81	0.21	0.59
chr4_74184142	rs733020532	G	A	ncRNA_intronic	*ENSGALG00000048521*		0.78	0.78	0.10	0.11	0.78	0.11	0.67
chr9_6290012	rs14667193	G	T	downstream	*ENSGALG00000049259*		0.61	0.61	0.10	0.06	0.61	0.08	0.53
chr6_12512820	rs15780500	G	A	UTR3	*ENSGALG00000049446*	*SLC29A3*	0.61	0.67	0.10	0.17	0.64	0.13	0.51
chr4_74213430	rs315917373	G	A	upstream	*ENSGALG00000050078*		0.11	0.22	0.65	0.83	0.17	0.74	0.58
chr4_74218577	na	T	C	ncRNA_exonic	*ENSGALG00000050078*		0.89	0.78	0.35	0.17	0.83	0.26	0.58
chr3_73386736	na	G	C	upstream	*ENSGALG00000050194*		0.78	0.83	0.35	0.22	0.81	0.29	0.52

chr2_26960211	rs317304288	G	A	ncRNA_intronic	*ENSGALG00000050874*		0.33	0.17	0.70	0.89	0.25	0.79	0.54
chr1_155052318	rs312499696	C	T	ncRNA_intronic	*ENSGALG00000051545*		0.11	0.00	0.65	0.78	0.06	0.71	0.66
chr4_74290857	rs741369370	A	G	ncRNA_exonic	*ENSGALG00000051573*		0.83	0.67	0.25	0.22	0.75	0.24	0.51
chr1_110369650	rs15380390	T	C	ncRNA_exonic	*ENSGALG00000051940*		0.39	0.00	0.80	0.94	0.19	0.87	0.68
chr4_9970045	rs313910852	T	A	ncRNA_intronic	*ENSGALG00000052590*		0.78	0.78	0.05	0.22	0.78	0.14	0.64
chr3_71643736	rs16298716	A	G	downstream	*ENSGALG00000052916*		0.94	0.94	0.40	0.44	0.94	0.42	0.52
chr8_15658008	rs734760197	G	C	ncRNA_intronic	*ENSGALG00000052940*		0.56	0.94	0.25	0.17	0.75	0.21	0.54
chr3_71085059	rs735672719	A	T	upstream	*ENSGALG00000053011*		0.67	0.72	0.15	0.11	0.69	0.13	0.56
chr7_34388391	rs313766753	A	C	ncRNA_intronic	*ENSGALG00000053295*		0.72	0.83	0.25	0.06	0.78	0.15	0.63
chr26_3952469	rs736838272	A	G	intronic	*ENSGALG00000054224*		0.61	0.78	0.10	0.28	0.69	0.19	0.51
chr1_97326982	rs741111731	C	T	ncRNA_intronic	*ENSGALG00000054585*		0.78	0.72	0.20	0.22	0.75	0.21	0.54
chr13_7328547	rs317710966	G	C	ncRNA_intronic	*ENSGALG00000054885*		0.00	0.00	0.60	0.44	0.00	0.52	0.52
chr2_28134759	rs13543650	C	A	ncRNA_intronic	*ENSGALG00000054958*		1.00	1.00	0.60	0.33	1.00	0.47	0.53

Table S4.3e. Candidate SNP list from signature analysis for SoilOrgC population group. (Only SNPs with large differences in allele frequency (dAAF>0.5) between Low and High groups are included. At least one SNP with dAAF>0.5 for each candidate gene is included. |dAAF|: difference in allele frequency between Low and High group).

snpID	rs_id	ref	alt	annotation	gene	gene ID	Low group pop1 AAF	Low group pop2 AAF	High group pop1 AAF	High group pop2 AAF	$\bar{\bar{X}}$ AAF in Low group	$\bar{\bar{X}}$ AAF in High group	|dAAF|
chr1_197282265	na	G	A	upstream	*ENSGALG00000047152*	*HBBA*	0.44	0.30	0.95	0.85	0.37	0.90	0.53
chr33_6484810	na	G	A	downstream	*ENSGALG00000047273*	*DRAP1*	0.00	0.05	0.75	0.50	0.03	0.63	0.60
chr1_197282265	na	G	A	downstream	*ENSGALG00000052767*		0.44	0.30	0.95	0.85	0.37	0.90	0.53
chr18_852214	rs317419481	G	A	intronic	*ENSGALG00000001021*	*MYOCD*	0.56	0.70	0.10	0.15	0.63	0.13	0.50
chr9_10691546	rs312400259	C	G	intronic	*ENSGALG00000002639*	*TRPC1*	1.00	0.85	0.40	0.35	0.93	0.38	0.55
chr9_10683587	rs740485975	G	A	intronic	*ENSGALG00000002647*	*PLS1*	0.00	0.15	0.60	0.60	0.08	0.60	0.53
chr10_6837508	rs15570701	A	G	intronic	*ENSGALG00000004087*	*THSD4*	0.78	0.85	0.15	0.20	0.81	0.18	0.64
chr5_1878183	rs15650675	T	C	intronic	*ENSGALG00000004153*	*NADSYN1*	0.72	0.70	0.10	0.25	0.71	0.18	0.54
chr5_1873527	rs738367843	C	T	intronic	*ENSGALG00000004166*	*VPS51*	0.78	0.75	0.10	0.20	0.76	0.15	0.61
chr5_1866493	rs15650715	T	C	UTR3	*ENSGALG00000004172*	*BET1L*	0.83	0.65	0.15	0.10	0.74	0.13	0.62

chr5_1862115	rs316635524	C	T	intronic	*ENSGALG00000004191*	*RIC8A*	0.78	0.60	0.10	0.05	0.69	0.08	0.61
chr14_4757031	rs315337156	T	G	intronic	*ENSGALG00000004693*	*ARPC1A*	0.22	0.40	0.85	0.90	0.31	0.88	0.56
chr17_5539734	rs317472633	G	A	intronic	*ENSGALG00000004719*	*SPTAN1*	0.50	0.75	0.05	0.10	0.63	0.08	0.55
chr2_5882084	rs316359926	T	A	intronic	*ENSGALG00000006158*	*ACVR2B*	0.17	0.00	0.75	0.75	0.08	0.75	0.67
chr15_9915198	rs738971750	A	G	intronic	*ENSGALG00000007354*	*CIT*	0.67	0.65	0.15	0.15	0.66	0.15	0.51
chr2_17229852	rs317151158	A	G	intronic	*ENSGALG00000007766*	*KIAA1217*	0.83	0.85	0.15	0.20	0.84	0.18	0.67
chr4_13126545	rs735864582	G	A	intronic	*ENSGALG00000007972*	*TRPC5*	0.22	0.10	0.65	0.75	0.16	0.70	0.54
chr3_28356219	rs317888833	C	A	intronic	*ENSGALG00000010050*	*LRFN2*	0.83	0.75	0.15	0.30	0.79	0.23	0.57
chr4_43282033	rs317960491	T	G	intronic	*ENSGALG00000010744*	*GALNT7*	0.72	0.95	0.25	0.20	0.84	0.23	0.61
chr4_43291522	na	A	T	intronic	*ENSGALG00000010745*	*HMGB2*	0.72	0.75	0.15	0.05	0.74	0.10	0.64
chr4_46301618	rs738519856	C	T	intronic	*ENSGALG00000011133*	*LIN54*	0.22	0.10	0.75	0.75	0.16	0.75	0.59
chr4_46462249	rs317703256	T	C	intronic	*ENSGALG00000011208*	*HELQ*	0.28	0.05	0.80	0.70	0.16	0.75	0.59
chr4_46469240	rs14462895	G	A	intronic	*ENSGALG00000011210*	*ABRAXAS1*	0.33	0.10	0.90	0.90	0.22	0.90	0.68
chr4_46467397	rs315716648	T	C	nonsynonymous	*ENSGALG00000011210*	*ABRAXAS1*	0.17	0.05	0.60	0.80	0.11	0.70	0.59
chr4_46467877	rs317124316	T	C	nonsynonymous	*ENSGALG00000011210*	*ABRAXAS1*	0.17	0.10	0.60	0.80	0.13	0.70	0.57
chr4_46467281	rs314846798	T	C	nonsynonymous	*ENSGALG00000011210*	*ABRAXAS1*	0.22	0.10	0.60	0.80	0.16	0.70	0.54
chr1_44957795	rs315703916	C	T	intronic	*ENSGALG00000011297*	*CRADD*	0.33	0.20	0.90	0.80	0.27	0.85	0.58
chr4_48305392	rs314276252	T	C	intronic	*ENSGALG00000011319*	*ADGRL3*	0.06	0.20	0.90	0.70	0.13	0.80	0.67
chr4_50884060	rs314082155	T	C	intronic	*ENSGALG00000011664*	*RASSF6*	0.06	0.10	0.70	0.80	0.08	0.75	0.67
chr4_50924373	rs316517697	C	A	upstream	*ENSGALG00000011668*	*IL8L1*	0.11	0.10	0.75	0.75	0.11	0.75	0.64
chr7_35151474	rs317912233	G	A	intronic	*ENSGALG00000012484*	*RIF1*	0.17	0.10	0.60	0.75	0.13	0.68	0.54
chr7_35402192	rs14630853	G	A	intronic	*ENSGALG00000012513*	*STAM2*	0.72	0.80	0.15	0.15	0.76	0.15	0.61
chr1_55982800	rs317669778	A	C	intronic	*ENSGALG00000012791*	*TBXAS1*	0.17	0.35	0.80	0.75	0.26	0.78	0.52
chr1_60820017	rs316758233	A	G	intronic	*ENSGALG00000012994*	*ERC1*	0.17	0.30	0.85	0.80	0.23	0.83	0.59
chr1_60860135	rs14833762	C	T	intronic	*ENSGALG00000012998*	*WNT5B*	0.83	0.60	0.10	0.30	0.72	0.20	0.52
chr1_64902090	rs740935649	G	T	intronic	*ENSGALG00000013143*	*PDE3A*	0.78	0.65	0.20	0.10	0.71	0.15	0.56
chr4_76540574	rs315713688	A	G	intronic	*ENSGALG00000014496*	*Prom1*	0.83	0.70	0.05	0.45	0.77	0.25	0.52
chr1_78048067	rs740296375	T	A	nonsynonymous	*ENSGALG00000014603*	*C1S*	0.00	0.10	0.65	0.55	0.05	0.60	0.55
chr1_78049446	rs14845732	A	C	nonsynonymous	*ENSGALG00000014603*	*C1S*	1.00	0.90	0.30	0.50	0.95	0.40	0.55
chr1_78048418	rs13887326	G	A	nonsynonymous	*ENSGALG00000014603*	*C1S*	0.06	0.15	0.65	0.60	0.10	0.63	0.52
chr1_78065070	rs313713208	T	G	intronic	*ENSGALG00000014659*	*C1R*	0.00	0.15	0.70	0.60	0.08	0.65	0.58
chr1_78070795	rs314288648	C	T	intronic	*ENSGALG00000014662*	*RBP5*	0.00	0.10	0.60	0.60	0.05	0.60	0.55
chr1_78077393	rs317830961	C	T	intronic	*ENSGALG00000014674*	*CLSTN3*	0.89	0.85	0.30	0.25	0.87	0.28	0.59
chr3_66629773	rs15377681	A	C	intronic	*ENSGALG00000015061*	*CDC40*	0.11	0.15	0.75	0.60	0.13	0.68	0.54

chr3_67213317	rs737877959	T	G	upstream	*ENSGALG00000015289*	*ARMC2*	0.00	0.15	0.90	0.45	0.08	0.68	0.60
chr3_67294992	rs739018988	G	T	intronic	*ENSGALG00000015297*	*FOXO3*	0.06	0.00	0.75	0.55	0.03	0.65	0.62
chr1_105401546	rs740014385	G	A	intronic	*ENSGALG00000015822*		0.33	0.15	0.85	0.75	0.24	0.80	0.56
chr1_108010859	rs316401785	G	A	intronic	*ENSGALG00000016042*	*CHAF1B*	0.67	0.45	0.00	0.00	0.56	0.00	0.56
chr1_117673165	rs738448021	T	C	intronic	*ENSGALG00000016288*	*IL1RAPL1*	0.17	0.20	0.80	0.75	0.18	0.78	0.59
chr3_87499222	rs317097894	G	C	intronic	*ENSGALG00000016294*	*HMGCLL1*	0.00	0.20	0.75	0.80	0.10	0.78	0.68
chr3_108224362	rs740527901	A	G	intronic	*ENSGALG00000016680*	*TFAP2B*	0.22	0.10	0.80	0.65	0.16	0.73	0.56
chr1_146498503	rs737557685	T	C	intronic	*ENSGALG00000016886*	*FARP1*	0.72	0.45	0.00	0.15	0.59	0.08	0.51
chr1_167340986	rs318131789	G	A	intronic	*ENSGALG00000016955*	*VWA8*	0.28	0.20	0.90	0.65	0.24	0.78	0.54
chr1_168941361	rs317293763	C	T	intronic	*ENSGALG00000016972*	*NUFIP1*	0.94	0.75	0.25	0.10	0.85	0.18	0.67
chr1_171744646	rs737161831	A	C	intronic	*ENSGALG00000017026*	*VPS36*	0.33	0.25	0.85	0.80	0.29	0.83	0.53
chr1_197287932	na	G	A	downstream	*ENSGALG00000017347*	*HBBR*	0.39	0.25	0.95	0.75	0.32	0.85	0.53
chr4_46465110	rs15564231	A	G	intronic	*ENSGALG00000023191*	*MRPS18C*	0.28	0.05	0.90	0.70	0.16	0.80	0.64
chr9_18772140	rs740431523	G	A	intronic	*ENSGALG00000023670*	*NAALADL2*	0.06	0.15	0.70	0.60	0.10	0.65	0.55
chr4_50940750	rs16409260	C	T	UTR3	*ENSGALG00000026098*	*IL8*	0.83	0.85	0.20	0.15	0.84	0.18	0.67
chr3_106483692	rs735464908	G	A	intronic	*ENSGALG00000028032*	*MSRA*	0.06	0.10	0.90	0.65	0.08	0.78	0.70
chr1_197283909	na	C	T	intronic	*ENSGALG00000028273*	*HBE1*	0.28	0.30	0.95	0.75	0.29	0.85	0.56
chr4_40261626	rs15551203	G	A	intronic	*ENSGALG00000030065*	*TENM3*	0.17	0.40	0.95	0.80	0.28	0.88	0.59
chr4_40127982	na	G	A	nonsynonymous	*ENSGALG00000030065*	*TENM3*	0.17	0.30	0.80	0.70	0.23	0.75	0.52
chr2_137929184	rs317580060	A	G	intronic	*ENSGALG00000030460*	*TBC1D31*	0.06	0.05	0.55	0.65	0.05	0.60	0.55
chr4_42916104	rs14458637	A	G	intronic	*ENSGALG00000030588*	*GALNTL6*	0.11	0.40	0.90	0.80	0.26	0.85	0.59
chr14_13193066	rs731928779	C	G	intronic	*ENSGALG00000030877*	*CLUAP1*	0.22	0.15	0.85	0.75	0.19	0.80	0.61
chr10_1538694	rs733384767	C	T	intronic	*ENSGALG00000031408*	*SORD*	0.33	0.25	0.85	0.75	0.29	0.80	0.51
chr1_168960849	rs738879763	C	A	intronic	*ENSGALG00000031567*	*GPALPP1*	0.94	0.70	0.20	0.20	0.82	0.20	0.62
chr18_9897787	na	C	T	UTR3	*ENSGALG00000032336*	*SLC25A10*	0.83	0.80	0.35	0.15	0.82	0.25	0.57
chr3_35172014	rs734091960	A	T	intronic	*ENSGALG00000034081*	*AKT3*	0.06	0.20	0.90	0.60	0.13	0.75	0.62
chr5_1196708	rs315998563	G	T	intronic	*ENSGALG00000035242*	*RMDN3*	0.17	0.40	0.85	0.80	0.28	0.83	0.54
chr7_35393566	rs317060282	G	A	intronic	*ENSGALG00000035466*	*CACNB4*	0.11	0.15	0.90	0.95	0.13	0.93	0.79
chr18_9872059	rs16347414	T	C	downstream	*ENSGALG00000038875*	*TMC6*	0.78	0.75	0.40	0.10	0.76	0.25	0.51
chr20_1587973	rs317149019	G	A	intronic	*ENSGALG00000039101*	*ASIP*	0.11	0.20	0.70	0.75	0.16	0.73	0.57
chr33_6498218	na	G	T	UTR5	*ENSGALG00000039409*	*RELA*	0.00	0.05	0.70	0.45	0.03	0.58	0.55
chr10_13674790	rs314848346	G	A	intronic	*ENSGALG00000040241*	*NTRK3*	0.89	0.65	0.10	0.15	0.77	0.13	0.64
chr9_19001952	rs735432390	C	A	intronic	*ENSGALG00000040938*	*NLGN1*	0.11	0.40	0.85	0.75	0.26	0.80	0.54
chr3_66606219	rs316456097	A	G	intronic	*ENSGALG00000041120*	*METTL24*	0.11	0.15	0.80	0.65	0.13	0.73	0.59

chr3_106838116	rs431829052	G	A	intronic	*ENSGALG00000041139*	*PINX1*	0.22	0.30	0.85	0.75	0.26	0.80	0.54
chr1_174193558	rs736954116	C	T	intronic	*ENSGALG00000041263*	*DCLK1*	0.11	0.15	0.85	0.85	0.13	0.85	0.72
chr14_15399971	rs741294578	T	C	intronic	*ENSGALG00000041298*		0.28	0.25	0.85	0.85	0.26	0.85	0.59
chr7_33913011	rs794420089	C	T	ncRNA_intronic	*ENSGALG00000041782*		0.11	0.15	1.00	0.60	0.13	0.80	0.67
chr7_33917087	rs733609911	T	C	ncRNA_exonic	*ENSGALG00000041782*		0.17	0.15	1.00	0.60	0.16	0.80	0.64
chr4_50964457	rs14468633	T	C	downstream	*ENSGALG00000046160*	*PPBP*	0.11	0.10	0.75	0.75	0.11	0.75	0.64
chr4_50963368	rs312785207	G	A	UTR3	*ENSGALG00000046160*	*PPBP*	0.11	0.05	0.70	0.70	0.08	0.70	0.62
chr4_50962848	rs16409303	G	A	nonsynonymous	*ENSGALG00000046160*	*PPBP*	0.22	0.20	0.70	0.80	0.21	0.75	0.54
chr4_43294449	rs317275943	G	A	ncRNA_intronic	*ENSGALG00000046857*		0.22	0.20	0.75	0.85	0.21	0.80	0.59
chr4_42054421	rs315189131	A	G	ncRNA_intronic	*ENSGALG00000047221*		0.17	0.35	0.90	0.70	0.26	0.80	0.54
chr3_108218185	rs16342493	T	C	intronic	*ENSGALG00000047589*		0.72	0.80	0.20	0.25	0.76	0.23	0.54
chr4_6492223	rs740869514	A	G	ncRNA_intronic	*ENSGALG00000047994*		0.72	0.60	0.15	0.05	0.66	0.10	0.56
chr2_9798934	rs734424330	C	T	ncRNA_exonic	*ENSGALG00000048545*		0.11	0.35	0.80	0.70	0.23	0.75	0.52
chr7_31522897	rs313169111	C	T	intronic	*ENSGALG00000048656*		0.22	0.25	0.95	0.75	0.24	0.85	0.61
chr9_18864400	rs737848314	G	A	downstream	*ENSGALG00000048975*		0.22	0.35	0.85	0.80	0.29	0.83	0.54
chr2_56240597	rs741421980	C	A	intronic	*ENSGALG00000049812*	*PQLC1*	0.17	0.05	0.60	0.70	0.11	0.65	0.54
chr4_40973754	rs15552615	T	C	ncRNA_intronic	*ENSGALG00000050037*		0.83	0.85	0.20	0.20	0.84	0.20	0.64
chr3_107774437	rs316811446	G	A	downstream	*ENSGALG00000050553*		0.78	0.85	0.05	0.30	0.81	0.18	0.64
chr6_25343087	rs741144048	G	C	ncRNA_intronic	*ENSGALG00000050593*		0.00	0.10	0.75	0.70	0.05	0.73	0.68
chr6_25331885	na	G	T	ncRNA_exonic	*ENSGALG00000050593*		0.00	0.05	0.65	0.60	0.03	0.63	0.60
chr3_107808168	rs317702451	C	T	upstream	*ENSGALG00000050753*		0.72	0.90	0.05	0.05	0.81	0.05	0.76
chr3_91851269	rs318216900	A	G	upstream	*ENSGALG00000050914*		0.67	0.55	0.10	0.10	0.61	0.10	0.51
chr33_6484465	na	G	C	ncRNA_exonic	*ENSGALG00000051169*	*UQCC3*	0.00	0.10	0.75	0.60	0.05	0.68	0.63
chr1_173810330	rs731851082	A	G	downstream	*ENSGALG00000051912*		0.33	0.35	0.85	0.90	0.34	0.88	0.53
chr1_173813052	rs312496906	T	C	ncRNA_exonic	*ENSGALG00000051912*		0.22	0.30	0.75	0.80	0.26	0.78	0.51
chr5_17570223	rs15674380	G	A	ncRNA_intronic	*ENSGALG00000051988*		0.67	0.60	0.05	0.05	0.63	0.05	0.58
chr1_110526373	rs316127210	C	T	intronic	*ENSGALG00000052523*		0.00	0.35	0.80	0.65	0.18	0.73	0.55
chr1_110525301	rs740893938	T	C	nonsynonymous	*ENSGALG00000052523*		0.00	0.40	0.80	0.65	0.20	0.73	0.53
chr1_110525253	rs317142169	A	G	nonsynonymous	*ENSGALG00000052523*		0.33	0.45	0.90	0.90	0.39	0.90	0.51
chr1_93736639	rs317435027	C	G	ncRNA_intronic	*ENSGALG00000052608*		0.06	0.40	0.70	0.90	0.23	0.80	0.57
chr2_6128884	rs315845943	G	C	downstream	*ENSGALG00000053252*		0.33	0.50	0.90	0.95	0.42	0.93	0.51
chr1_197241952	na	A	G	intronic	*ENSGALG00000054014*		0.39	0.30	0.95	0.75	0.34	0.85	0.51
chr4_42135670	rs317255318	A	G	ncRNA_intronic	*ENSGALG00000054340*		0.83	0.65	0.10	0.30	0.74	0.20	0.54
chr6_26881924	rs317363623	A	G	ncRNA_intronic	*ENSGALG00000054720*		0.28	0.20	0.75	0.80	0.24	0.78	0.54

Table S4.3f. Candidate SNP list from signature analysis for LandUse population group. (Only SNPs with large differences in allele frequency (dAAF>0.5) between Low and High groups are included. At least one SNP with dAAF>0.5 for each candidate gene is included. |dAAF|: difference in allele frequency between Low and High group).

snpID	**rs_id**	**ref**	**alt**	**annotation**	**gene**	**gene ID**	**Low group pop1 AAF**	**Low group pop2 AAF**	**High group pop1 AAF**	**High group pop2 AAF**	**$\bar{\bar{X}}$ AAF in Low group**	**$\bar{\bar{X}}$ AAF in High group**	**|dAAF|**
chr3_34295140	rs317838524	T	C	intronic	*ENSGALG00000010664*	*KIF26B*	0.89	1.00	0.25	0.55	0.94	0.40	0.54
chr22_851111	rs314845387	A	G	intronic	*ENSGALG00000000242*	*EBF2*	0.72	0.89	0.25	0.30	0.81	0.28	0.53
chr13_11317880	rs313834857	G	A	downstream	*ENSGALG00000003672*	*UBLCP1*	0.11	0.06	0.65	0.75	0.08	0.70	0.62
chr6_16069210	rs736844754	C	T	intronic	*ENSGALG00000005057*	*ADK*	1.00	0.89	0.40	0.45	0.94	0.43	0.52
chr2_3539810	rs14129599	A	G	intronic	*ENSGALG00000005448*	*MYL3*	0.94	0.83	0.20	0.45	0.89	0.33	0.56
chr6_20203680	rs315457244	A	G	upstream	*ENSGALG00000006421*	*SLC16A12*	0.28	0.28	0.65	0.95	0.28	0.80	0.52
chr2_10257999	rs735572037	A	C	intronic	*ENSGALG00000006652*	*DIP2C*	0.94	0.94	0.30	0.45	0.94	0.38	0.57
chr3_10701300	rs317403184	A	T	intronic	*ENSGALG00000008778*	*SPRED2*	0.72	0.89	0.20	0.30	0.81	0.25	0.56
chr8_20061558	rs312780602	G	C	intronic	*ENSGALG00000009993*	*MED8*	0.72	0.89	0.25	0.05	0.81	0.15	0.66
chr3_34275191	rs312954554	C	T	intronic	*ENSGALG00000010653*	*SMYD3*	0.11	0.06	0.70	0.50	0.08	0.60	0.52
chr2_27960216	rs317883849	A	G	intronic	*ENSGALG00000010792*	*AGMO*	0.72	0.94	0.10	0.25	0.83	0.18	0.66
chr2_32216921	rs735947960	G	A	upstream	*ENSGALG00000011046*	*SNX10*	0.17	0.22	0.60	0.90	0.19	0.75	0.56
chr2_38801757	rs317204693	G	T	upstream	*ENSGALG00000011428*	*AZI2*	0.94	0.94	0.40	0.45	0.94	0.43	0.52
chr2_40559061	rs732153441	T	A	UTR3	*ENSGALG00000011498*	*TCAIM*	0.72	0.94	0.25	0.25	0.83	0.25	0.58
chr4_54303073	rs737880929	T	C	intronic	*ENSGALG00000011963*	*PRDM5*	0.78	0.78	0.15	0.20	0.78	0.18	0.60
chr4_56156061	rs313952468	T	C	intronic	*ENSGALG00000012015*	*NDST4*	0.56	0.89	0.05	0.35	0.72	0.20	0.52
chr1_64476422	rs13879021	C	T	intronic	*ENSGALG00000013134*	*PLEKHA5*	0.89	0.72	0.30	0.15	0.81	0.23	0.58
chr4_75792680	rs739647947	A	G	intronic	*ENSGALG00000014425*	*NCAPG*	0.67	0.67	0.30	0.00	0.67	0.15	0.52
chr2_99526132	rs735982242	C	T	intronic	*ENSGALG00000014615*	*LAMA1*	0.89	0.72	0.20	0.30	0.81	0.25	0.56
chr1_187388925	rs316030124	T	A	intronic	*ENSGALG00000017227*	*SLC36A4*	0.39	0.50	0.90	1.00	0.44	0.95	0.51
chr5_40641043	rs733812232	G	A	intronic	*ENSGALG00000027255*	*NRXN3*	0.00	0.22	0.80	0.55	0.11	0.68	0.56
chr8_20069462	rs15923605	A	G	intronic	*ENSGALG00000028561*	*SZT2*	0.39	0.11	0.80	0.95	0.25	0.88	0.63
chr8_20069280	rs313088300	T	C	nonsynonymous	*ENSGALG00000028561*	*SZT2*	0.39	0.17	0.80	0.95	0.28	0.88	0.60
chr6_21698226	na	T	C	intronic	*ENSGALG00000030597*	*KIF11*	0.67	0.78	0.15	0.15	0.72	0.15	0.57
chr2_54711769	rs314836526	A	T	intronic	*ENSGALG00000031170*	*ADCY1*	0.28	0.33	0.85	0.85	0.31	0.85	0.54
chr2_117843872	rs315792028	A	T	intronic	*ENSGALG00000031390*	*STAU2*	0.33	0.44	0.90	0.95	0.39	0.93	0.54
chr6_21795782	rs736082992	T	A	intronic	*ENSGALG00000032234*	*MARCH5*	0.22	0.00	0.45	0.80	0.11	0.63	0.51

chr8_1520789	rs313288599	T	C	intronic	*ENSGALG00000039327*	*VAV3*	0.94	0.61	0.30	0.25	0.78	0.28	0.50
chr3_36167656	rs317303812	T	A	intronic	*ENSGALG00000039772*	*RGS7*	0.67	0.56	0.10	0.05	0.61	0.08	0.54
chr2_117976975	rs312653243	C	T	intronic	*ENSGALG00000040097*	*UBE2W*	0.72	0.56	0.05	0.05	0.64	0.05	0.59
chr1_194710902	rs15051046	C	T	intronic	*ENSGALG00000040478*	*CAPN5*	0.89	0.94	0.40	0.20	0.92	0.30	0.62
chr1_132474638	rs741583645	G	A	intronic	*ENSGALG00000041042*	*GABRG3*	0.72	0.67	0.10	0.20	0.69	0.15	0.54
chr13_17886924	na	G	C	intronic	*ENSGALG00000043744*	*NDFIP1*	0.89	0.94	0.25	0.55	0.92	0.40	0.52
chr4_25957558	na	A	G	ncRNA_intronic	*ENSGALG00000046909*		0.22	0.17	0.55	0.85	0.19	0.70	0.51
chr1_65421972	rs315335738	T	C	upstream	*ENSGALG00000048674*		0.17	0.17	0.75	0.65	0.17	0.70	0.53
chr1_92331553	rs316354370	T	C	ncRNA_intronic	*ENSGALG00000049116*		0.00	0.06	0.60	0.50	0.03	0.55	0.52
chr4_25916124	na	T	C	ncRNA_intronic	*ENSGALG00000049671*		0.22	0.17	0.55	0.85	0.19	0.70	0.51
chr1_173116056	rs317447038	C	A	ncRNA_exonic	*ENSGALG00000049925*		0.33	0.00	0.85	0.65	0.17	0.75	0.58
chr11_13260695	rs317241311	A	G	ncRNA_intronic	*ENSGALG00000050348*		0.06	0.11	0.65	0.55	0.08	0.60	0.52
chr1_163441754	rs731900954	C	T	ncRNA_intronic	*ENSGALG00000050889*		0.89	0.83	0.25	0.35	0.86	0.30	0.56
chr2_82658367	rs314148401	T	A	ncRNA_intronic	*ENSGALG00000051256*		0.56	0.72	0.15	0.10	0.64	0.13	0.51
chr10_10479871	rs316438713	G	A	ncRNA_exonic	*ENSGALG00000051344*		0.78	0.61	0.05	0.30	0.69	0.18	0.52
chr2_82925607	rs741742877	A	G	ncRNA_intronic	*ENSGALG00000051615*		0.56	0.72	0.15	0.05	0.64	0.10	0.54
chr1_110369132	rs13920393	G	A	downstream	*ENSGALG00000051940*		0.22	0.06	0.80	0.55	0.14	0.68	0.54
chr2_30130035	rs741615728	T	C	ncRNA_intronic	*ENSGALG00000052719*		1.00	0.83	0.35	0.40	0.92	0.38	0.54
chr5_51336712	rs314805918	A	G	downstream	*ENSGALG00000053209*		1.00	0.50	0.20	0.25	0.75	0.23	0.53

Table S4.4. Candidate genes from selection signature regions overlapping with QTLs (ChickenQTLdb). (Only QTLs with size > 1 Mb were considered.

Environmental parameter	**genes**	**gene ID**	**selection direction**	**Method of detection**	**Overlap with significant QTLs from Chicken QTLdb**
minTemp	*ENSGALG00000002692*		low	*XP-EHH*	Breast parameters, Digestive system property
minTemp	*ENSGALG00000003243*	*RAD54L2*	low	*XP-EHH*	Egg production related, Eggshell property
minTemp	*ENSGALG00000003562*	*NPTX2*	low	*XP-EHH*	Growth
minTemp	*ENSGALG00000004420*	*SDK1*	low	*Fst*	Body temperature, Breast parameters, Growth
minTemp	*ENSGALG00000005139*	*CLTC*	low	*XP-EHH*	FCR
minTemp	*ENSGALG00000007038*	*ASTN2*	low	*XP-EHH*	Immune response related
minTemp	*ENSGALG00000008282*	*ATG4A*	low	*XP-EHH*	Blood parameters
minTemp	*ENSGALG00000009943*	*FAM117A*	low	*XP-EHH*	Skin related
minTemp	*ENSGALG00000009944*	*KAT7*	low	*XP-EHH*	Feed intake, Skin related
minTemp	*ENSGALG00000009983*	*NUBPL*	low	*XP-EHH*	Immune response related
minTemp	*ENSGALG00000012007*	*ST13*	low	*XP-EHH*	FCR
minTemp	*ENSGALG00000012888*	*PHLPP1*	low	*XP-EHH*	Abdominal fat properties
minTemp	*ENSGALG00000013646*	*VPS37A*	low	*XP-EHH*	Immune response related
minTemp	*ENSGALG00000013649*	*CNOT7*	low	*XP-EHH*	Immune response related
minTemp	*ENSGALG00000014467*	*COPS7A*	high	*Fst*	FCR
minTemp	*ENSGALG00000016689*	*ASMTL*	high	*XP-EHH*	Immune response related
minTemp	*ENSGALG00000025230*		low	*XP-EHH*	Immune response-related
minTemp	*ENSGALG00000031182*	*FANCM*	high	*XP-EHH*	Disease susceptibility
minTemp	*ENSGALG00000031440*	*PTPRU*	low	*XP-EHH*	Growth
minTemp	*ENSGALG00000031487*	*FSTL4*	low	*XP-EHH*	Disease susceptibility
minTemp	*ENSGALG00000034967*	*GDPD5*	low	*Fst*	FCR
minTemp	*ENSGALG00000035578*	*RBM6*	low	*XP-EHH*	Egg production related, Eggshell property, FCR
minTemp	*ENSGALG00000036021*	*MTMR7*	low	*Fst*	Immune response related
minTemp	*ENSGALG00000036909*	*BEND5*	high	*XP-EHH*	Feather properties
precSeasonality	*ENSGALG00000000824*		low	*XP-EHH*	Body temperature
precSeasonality	*ENSGALG00000001153*	*AUTS2*	high	*XP-EHH*	Feed intake, Growth, Albumen property, Egg property
precSeasonality	*ENSGALG00000001768*	*TENM2*	low	*XP-EHH*	Age at sexual maturity
precSeasonality	*ENSGALG00000003958*	*PRMT3*	high	*XP-EHH*	Growth
precSeasonality	*ENSGALG00000005853*	*HTR2C*	high	*XP-EHH*, *Fst*	Immune response-related, Feather pecking, Feather properties

precSeasonality	*ENSGALG00000006743*	*YWHAH*	low	*XP-EHH*	Body temperature, Age at sexual maturity
precSeasonality	*ENSGALG00000007172*	*FHIT*	high	*XP-EHH*	Digestive system property, Growth, Eggshell property
precSeasonality	*ENSGALG00000008309*		low	*XP-EHH*	Thymus weight
precSeasonality	*ENSGALG00000009491*	*RAD51B*	low	*XP-EHH*	Breast parameters, Feather pecking
precSeasonality	*ENSGALG00000009493*	*TMEM229B*	low	*XP-EHH*	Breast parameters
precSeasonality	*ENSGALG00000009880*	*INPP4B*	high	*XP-EHH*	Growth, Immune response related
precSeasonality	*ENSGALG00000010053*	*PTPRF*	high	*XP-EHH*	Breast parameters
precSeasonality	*ENSGALG00000011630*	*GLI2*	low	*XP-EHH*	FCR, Growth, Immune response related
precSeasonality	*ENSGALG00000012338*	*BDH2*	low	*XP-EHH*	Spleen property
precSeasonality	*ENSGALG00000012424*	*GTDC1*	low	*XP-EHH*	FCR
precSeasonality	*ENSGALG00000014425*	*NCAPG*	high	*XP-EHH*	Egg property, Gizzard property, Heart condition, Liver related, Ovary property, Proventriculus weight, Growth
precSeasonality	*ENSGALG00000015897*	*IMPG1*	high	*XP-EHH*	Eggshell property
precSeasonality	*ENSGALG00000015898*	*MYO6*	high	*XP-EHH*	Eggshell property
precSeasonality	*ENSGALG00000016955*	*VWA8*	high	*XP-EHH*	Thymus weight
precSeasonality	*ENSGALG00000029022*	*NPAS3*	high	*Fst*	Feather properties
precSeasonality	*ENSGALG00000030054*	*PTPRN2*	high	*Fst*	Feather properties, Age at sexual maturity
precSeasonality	*ENSGALG00000030684*		low	*Fst*	Body temperature, Age at sexual maturity
precSeasonality	*ENSGALG00000032658*	*FMNL1*	low	*XP-EHH*	Body temperature
precSeasonality	*ENSGALG00000033470*	*ADGRA3*	high	*XP-EHH*	Eggshell property
precSeasonality	*ENSGALG00000035137*		low	*XP-EHH*	Growth
precSeasonality	*ENSGALG00000035554*		low	*XP-EHH*	Body temperature
precSeasonality	*ENSGALG00000041121*	*SLIT2*	high	*Fst*	Eggshell property, Growth, Abdominal fat properties, Egg property
precSeasonality	*ENSGALG00000041782*		high	*XP-EHH*	Immune response-related
precSeasonality	*ENSGALG00000042814*	*MCRIP1*	low	*XP-EHH*	Blood parameters
precSeasonality	*ENSGALG00000045753*	*MRPL12*	low	*Fst*	Bursa of Fabricius weight
precWQ	*ENSGALG00000000461*	*MRC2*	high	*XP-EHH*	Bursa of Fabricius weight
precWQ	*ENSGALG00000002546*	*COL5A1*	low	*XP-EHH*	FCR
precWQ	*ENSGALG00000004504*		high	*XP-EHH*	Growth
precWQ	*ENSGALG00000004530*	*MMD2*	high	*XP-EHH*	Growth
precWQ	*ENSGALG00000005483*	*CDH13*	high	*Fst*	Yolk property
precWQ	*ENSGALG00000006379*	*SHH*	high	*XP-EHH*	Heart condition
precWQ	*ENSGALG00000008256*	*BLM*	high	*XP-EHH*	FCR
precWQ	*ENSGALG00000009865*	*RNF150*	low	*Fst*	Growth

precWQ	*ENSGALG00000010616*	*PTPN21*	low	*XP-EHH*	Egg production related
precWQ	*ENSGALG00000010622*	*ZC3H14*	low	*XP-EHH*	Egg production related
precWQ	*ENSGALG00000012888*	*PHLPP1*	high	*Fst*	Abdominal fat properties
precWQ	*ENSGALG00000025659*	*gga-mir-2129*	high	*XP-EHH*	Body temperature, Growth
precWQ	*ENSGALG00000031960*	*LEF1*	low	*XP-EHH*	Body condition, Breast parameters, Leg muscle properties
precWQ	*ENSGALG00000032577*	*MECOM*	low	*XP-EHH*	Bursa of Fabricius weight, Growth
precWQ	*ENSGALG00000036909*	*BEND5*	low	*Fst*	Feather properties
precWQ	*ENSGALG00000037031*		low	*XP-EHH*	Comb property
precWQ	*ENSGALG00000038223*	*WHRN*	low	*XP-EHH*	Comb property
precWQ	*ENSGALG00000039826*	*CNGA3*	low	*XP-EHH*	FCR
precWQ	*ENSGALG00000040003*	*STXBP1*	low	*XP-EHH*	Comb property
precWQ	*ENSGALG00000040528*	*AP5Z1*	high	*XP-EHH*	Growth
precWQ	*ENSGALG00000041437*	*FAM155B*	low	*XP-EHH*	Comb property
precWQ	*ENSGALG00000042453*	*gga-mir-7442*	low	*XP-EHH*	Comb property
precDQ	*ENSGALG00000000627*	*DNAJC11*	high	*XP-EHH*	Growth, Eggshell property, Dry matter digestibility
precDQ	*ENSGALG00000001153*	*AUTS2*	low	*XP-EHH*	Feed intake, Growth, Albumen property, Egg property
precDQ	*ENSGALG00000002170*		low	*XP-EHH*	Growth, Breast parameters, Digestive system property
precDQ	*ENSGALG00000004128*	*CYSLTR1*	low	*XP-EHH*	Bursa of Fabricius weight, Comb property, Skin related, Albumen property
precDQ	*ENSGALG00000005139*	*CLTC*	high	*XP-EHH*	FCR
precDQ	*ENSGALG00000005872*	*BANP*	low	*XP-EHH*, *Fst*	Immune response related
precDQ	*ENSGALG00000006912*	*CPEB3*	low	*XP-EHH*	Breast parameters
precDQ	*ENSGALG00000009078*	*GALNT14*	low	*XP-EHH*	Breast parameters
precDQ	*ENSGALG00000009813*	*SCOC*	high	*Fst*	Growth
precDQ	*ENSGALG00000009826*	*CLGN*	high	*Fst*	Growth
precDQ	*ENSGALG00000010053*	*PTPRF*	low	*XP-EHH*	Breast parameters
precDQ	*ENSGALG00000010744*	*GALNT7*	low	*XP-EHH*	Feather properties
precDQ	*ENSGALG00000012424*	*GTDC1*	high	*Fst*	FCR
precDQ	*ENSGALG00000012559*	*LARGE1*	high	*XP-EHH*	beta 2-globulin level
precDQ	*ENSGALG00000012975*	*IQSEC3*	low	*XP-EHH*	Eggshell property
precDQ	*ENSGALG00000013177*	*BCAT1*	low	*XP-EHH*	Eggshell property
precDQ	*ENSGALG00000013606*	*FGL1*	high	*Fst*	Immune response related
precDQ	*ENSGALG00000015018*	*CASQ2*	low	*Fst*	FCR
precDQ	*ENSGALG00000015112*	*ZNF521*	low	*Fst*	Eggshell property

precDQ	*ENSGALG00000016282*	*ZNF451*	high	*XP-EHH*	Thymus weight
precDQ	*ENSGALG00000016282*	*ZNF451*	low	*XP-EHH*	Thymus weight
precDQ	*ENSGALG00000028032*	*MSRA*	low	*XP-EHH, Fst*	Disease susceptibility
precDQ	*ENSGALG00000028148*	*PHF13*	high	*XP-EHH*	Growth, Eggshell property, Dry matter digestibility
precDQ	*ENSGALG00000030597*	*KIF11*	low	*XP-EHH*	Eggshell property
precDQ	*ENSGALG00000032234*	*MARCH5*	low	*XP-EHH*	Eggshell property
precDQ	*ENSGALG00000032672*	*KRT5*	high	*Fst*	Feather properties
precDQ	*ENSGALG00000033157*	*ADAMTS10*	high	*XP-EHH*	Eggshell property, Small yellow follicle number, Ovary property, Small yellow follicle number
precDQ	*ENSGALG00000033470*	*ADGRA3*	low	*Fst*	Eggshell property
precDQ	*ENSGALG00000033569*		high	*Fst*	Growth, Eggshell property, Dry matter digestibility
precDQ	*ENSGALG00000035211*	*KLHL21*	high	*XP-EHH*	Growth, Eggshell property, Dry matter digestibility
precDQ	*ENSGALG00000035492*	*ELMO1*	high	*XP-EHH*	Growth, Eggshell property
precDQ	*ENSGALG00000035505*	*AOAH*	high	*XP-EHH*	Eggshell property
precDQ	*ENSGALG00000037856*	*CHL1*	low	*XP-EHH*	Feed intake
precDQ	*ENSGALG00000038924*	*IDE*	low	*XP-EHH*	Eggshell property
precDQ	*ENSGALG00000039472*	*SART3*	high	*Fst*	Skin related, Albumen property
precDQ	*ENSGALG00000043642*	*ANLN*	high	*Fst*	Eggshell property
precDQ	*ENSGALG00000043689*		high	*XP-EHH*	Feather properties
precDQ	*ENSGALG00000045753*	*MRPL12*	high	*XP-EHH*	Bursa of Fabricius weight
Landuse	*ENSGALG00000000636*	*MDM4*	low	*Fst*	Abdominal fat properties
Landuse	*ENSGALG00000002170*		low	*XP-EHH*	Growth, Breast parameters, Digestive system property
Landuse	*ENSGALG00000002510*	*IP6K1*	low	*Fst*	Egg production related, Eggshell property, Abdominal fat properties
Landuse	*ENSGALG00000002511*		low	*Fst*	Growth, Breast parameters, Digestive system property
Landuse	*ENSGALG00000003563*	*YARS*	low	*XP-EHH*	FCR
Landuse	*ENSGALG00000006763*	*EXOC6*	low	*XP-EHH*	Eggshell property
Landuse	*ENSGALG00000006912*	*CPEB3*	low	*XP-EHH*	Breast parameters
Landuse	*ENSGALG00000010536*	*FAF1*	high	*Fst*	Feather properties, Egg property
Landuse	*ENSGALG00000011428*	*AZI2*	low	*XP-EHH*	Comb property
Landuse	*ENSGALG00000011439*	*RBMS3*	low	*Fst*	Feather properties
Landuse	*ENSGALG00000011635*	*TFCP2L1*	low	*XP-EHH*	FCR, Growth
Landuse	*ENSGALG00000012015*	*NDST4*	low	*XP-EHH, Fst*	Abdominal fat properties
Landuse	*ENSGALG00000013008*	*CHUNK-1*	high	*XP-EHH*	Eggshell property
Landuse	*ENSGALG00000013134*	*PLEKHA5*	high	*XP-EHH*	Eggshell property

Landuse	*ENSGALG00000013177*	*BCAT1*	low	*XP-EHH*	Eggshell property
Landuse	*ENSGALG00000014425*	*NCAPG*	high	*XP-EHH, Fst*	Albumen property, Egg production related, Eggshell property, Growth, Yolk property, Gizzard property, Liver property
Landuse	*ENSGALG00000014615*	*LAMA1*	low	*XP-EHH*	Cholesterol level, FCR
Landuse	*ENSGALG00000015846*	*SNAP91*	low	*Fst*	Eggshell property
Landuse	*ENSGALG00000016959*	*DGKH*	high	*Fst*	Growth
Landuse	*ENSGALG00000028851*	*RAB8A*	low	*Fst*	FCR
Landuse	*ENSGALG00000029022*	*NPAS3*	high	*XP-EHH*	Feather properties
Landuse	*ENSGALG00000030597*	*KIF11*	low	*Fst*	Eggshell property
Landuse	*ENSGALG00000031390*	*STAU2*	high	*XP-EHH*	FCR
Landuse	*ENSGALG00000032234*	*MARCH5*	low	*XP-EHH*	Eggshell property
Landuse	*ENSGALG00000033695*	*CYP26A1*	low	*Fst*	Eggshell property
Landuse	*ENSGALG00000035533*	*ADIPOR2*	high	*XP-EHH*	Eggshell property, Growth
Landuse	*ENSGALG00000036909*	*BEND5*	low	*Fst*	Feather properties
Landuse	*ENSGALG00000036964*		high	*Fst*	Eggshell property
Landuse	*ENSGALG00000037043*	*PIK3C2B*	low	*XP-EHH*	Abdominal fat properties
Landuse	*ENSGALG00000038924*	*IDE*	low	*XP-EHH*	Eggshell property
Landuse	*ENSGALG00000039327*	*VAV3*	low	*XP-EHH*	Spleen property
Landuse	*ENSGALG00000039772*	*RGS7*	high	*Fst*	Growth
Landuse	*ENSGALG00000041042*	*GABRG3*	low	*Fst*	Growth
Landuse	*ENSGALG00000043213*	*CYP26C1*	low	*Fst*	Eggshell property
soilOrgC	*ENSGALG00000004087*	*THSD4*	high	*XP-EHH, Fst*	Bursa of Fabricius weight
soilOrgC	*ENSGALG00000004719*	*SPTAN1*	high	*XP-EHH*	Growth
soilOrgC	*ENSGALG00000005379*	*SGSM1*	high	*Fst*	Skin related
soilOrgC	*ENSGALG00000007354*	*CIT*	low	*XP-EHH*	Age at sexual maturity
soilOrgC	*ENSGALG00000009424*	*FOXP2*	high	*Fst*	Earlobe property
soilOrgC	*ENSGALG00000010744*	*GALNT7*	high	*Fst*	Feather properties
soilOrgC	*ENSGALG00000012989*	*NINJ2*	high	*XP-EHH*	Eggshell property
soilOrgC	*ENSGALG00000012994*	*ERC1*	high	*XP-EHH*	Eggshell property
soilOrgC	*ENSGALG00000012998*	*WNT5B*	high	*Fst*	Wing porperty
soilOrgC	*ENSGALG00000013143*	*PDE3A*	low	*Fst*	Eggshell property
soilOrgC	*ENSGALG00000014496*	*Prom1*	high	*XP-EHH*	Proventriculus weight, Carcass property
soilOrgC	*ENSGALG00000015061*	*CDC40*	high	*Fst*	Breast parameters, FCR
soilOrgC	*ENSGALG00000016886*	*FARP1*	high	*XP-EHH*	FCR

soilOrgC	*ENSGALG00000016955*	*VWA8*	high	*Fst*	Thymus weight
soilOrgC	*ENSGALG00000016974*	*GTF2F2*	high	*XP-EHH*	Gizzard property, Heart condition, Liver related, Proventriculus weight
soilOrgC	*ENSGALG00000017025*	*CKAP2*	high	*XP-EHH*	Comb property, Feed intake, Ovary property
soilOrgC	*ENSGALG00000017026*	*VPS36*	high	*Fst*	FCR, Comb property, Feed intake, Ovary property
soilOrgC	*ENSGALG00000023670*	*NAALADL2*	high	*XP-EHH*	Breast parameters
soilOrgC	*ENSGALG00000028032*	*MSRA*	high	*Fst*	Disease susceptibility
soilOrgC	*ENSGALG00000033311*		high	*XP-EHH*	Immune response-related
soilOrgC	*ENSGALG00000040938*	*NLGN1*	high	*Fst*	Breast parameters
soilOrgC	*ENSGALG00000041120*	*METTL24*	high	*Fst*	Breast parameters
soilOrgC	*ENSGALG00000041263*	*DCLK1*	high	*XP-EHH*	Feed intake, Pre-ovulatory follicle number
soilOrgC	*ENSGALG00000041510*	*DUOX2*	high	*XP-EHH*	Breast parameters
soilOrgC	*ENSGALG00000041782*		high	*XP-EHH*	Immune response-related
soilOrgC	*ENSGALG00000045753*	*MRPL12*	low	*XP-EHH*	Bursa of Fabricius weight

Chapter 5

Dissecting genome-environment associations in Ethiopian chicken ecotypes

Table S5.1. The number of candidates *ZHp* selective sweep regions per ecotype and chromosome.

	ecotypes											
	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
chr1	1	5	6	3	6	3		1	10	14	1	3
chr2	6	6	7	2	4	3		2	5	6	3	7
chr3	4		4	4	5	4		1	1	2	2	4
chr4	1	5	8	1	2	4		2	6	13	4	4
chr5	3	3	9	2	3	6	2	2	17	11	3	2
chr6	1	3	3	1		1			1	2	1	2
chr7	1	5	2	2	1	2			2		4	4
chr8		1	7	3		3			2	5		1
chr9	3	1		1	1	2		1	1	3		1
chr10		2	1			1			1	2		1
chr11			2	1	1				3	1		
chr12	2	3	1			2		1	2	1		3
chr13												
chr14	1		1									1
chr15						1			2			
chr16												
chr17												
chr18	1		2	1	1				1	2		
chr19	1									1		
chr20						1			1	1		
chr21												
chr22	1		1			1				2		
chr23					1	3			1			
chr24		1								1		1
chr25												
chr26			1			4						
chr27	2		1	1	3	2						
chr28												
chr29												
chr30					1				1			
chr32									1			
chr33	1					2						
Total	29	35	56	22	29	45	2	10	58	67	18	34

Table S5.2. The number of candidates *iHS* selective sweep regions per ecotype and chromosome.

	ecotypes											
	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
chr1	3	17	7	4	12	2	10	10	25	7	2	2
chr2	6	12	6	2	16	2	6	5	10	3	27	9
chr3	7	3		2	1	3			1	4		9
chr4	5			4	9	7	5		5	3	28	5
chr5	3	2	4	2	3	3	4	8	15		18	2
chr6		2	3		7	2		4	2			1
chr7	3	1		2	2	1	2					6
chr8	1	1	1		2		2	6		1	1	3
chr9			1	4	1		2		6	2		1
chr10	1	1	1		2		2				64	
chr11			1		1			1		1		3
chr12	1		1		1		1	6	3			
chr13		1		1				3				
chr14	1		2							1	11	
chr15	4			2	3							1
chr16									1			
chr17		7					2	1			20	
chr18					1	1	2					
chr19	2	1			1			3	1			1
chr20				1	2		1	1			3	
chr21					1	1	1	5				2
chr22												1
chr23							1		1			
chr24					1							
chr25	1	1							2			
chr26				4	2	1	3		1			
chr27	1		1						3		3	
chr28	1										3	
chr30	1	1	2	1								
chr31		1										
chr32									2			
chr33		2	2		6			3	2			
Total	41	53	32	29	74	23	44	56	80	22	180	46

Table S5.3 Common *Hp* sweep regions between ecotypes (* indicates lowest *ZHp* value for that particular ecotype).

Sweep regions	**Overlapping gene**	**Ecotypes shared (*ZHp* value)**	**# ecotypes sharing candidate sweep**
1_120930000_120960000	*ENSGALG00000049135*	E2 (-4.55), E9 (-4.536)	2
1_162440000_162460000	*ENSGALG00000047522*	E5 (-4.09), E10 (-4.038)	2
1_193490000_193510000	*TENM4*	E5 (-4.344), E9 (-4.165)	2
1_32700000_32750000	*ENSGALG00000050070,* *ENSGALG00000052108*	E9 (-4.877), E10 (-5.143)	2
1_32730000_32750000	*ENSGALG00000052108*	E1 (-4.081), E3 (-4.587)	2
1_32940000_32960000	*ENSGALG00000049735*	E3 (-4.29), E10 (-4.29)	2
2_141290000_141310000	*EFR3A*	E5 (-4.276), E8 (-4.045)	2
2_147370000_147390000	*TSNARE1*	E1 (-4.2), E2 (-4.42)	2
2_48280000_48300000	*PDE1C*	E1 (-4.319), E2 (-4.132)	2
3_50280000_50300000	*ENSGALG00000027159*	E3 (-4.158), E12 (-4.097)	2
3_51550000_51570000	*ARID1B*	E1 (-4.634), E3 (-4.408)	2
3_51610000_51630000	*TMEM242*	E4 (-4.006), E5 (-4.172)	2
3_85440000_85460000	*ENSGALG00000047999*	E4 (-4.546), E6 (-4.013)	2
3_69170000_69210000	*ENSGALG00000041688*	E1 (-4.43), E6 (-4.226)	2
5_41180000_41200000	*ENSGALG00000053642*	E5 (-4.08), E12 (-4.099)	2
5_59680000_59700000	*LRFN5*	E1 (-4.122), E6 (-4.031)	2
7_19820000_19850000	*SCN2A*	E3 (-4.31), E12 (-4.118)	2
9_23530000_23570000	*MBNL1*	E1 (-4.086), E2 (-4.358)	2
8_10000000_10030000	*ENSGALG00000049647*	E6 (-4.046), E10 (-4.983)	2
8_19380000_19410000	*ENSGALG00000008988*	E6 (-4.289), E9 (-5.226)	2
8_24640000_24670000	*ZFYVE9*	E2 (-4.272), E9 (-4.19)	2
8_9110000_9130000	*ENSGALG00000054728*	E3 (-4.349), E6 (-4.252)	2
11_1550000_1570000	*HYDIN*	E4 (-4.822), E9 (-4.359)	2
18_40000_60000	*ZNF302*	E5 (-4.114), E10 (-4.543)	2
22_3670000_3690000	*TMEM127*	E1 (-4.52), E6 (-4.669)	2
27_7800000_7830000	*TUBG1, CCR10, CNTNAP1*	E1 (-4.176), E3 (-4.637)	2
1_112400000_112430000	*EFHC2*	E4 (-4.301), E8 (-4.659)*, E10 (-4.242)	3
2_147140000_147160000	*TSNARE1*	E4 (-4.809), E6 (-4.178), E10 (-5.134)*	3
3_20100000_20120000	*ESRRG*	E1 (-4.023), E6 (-4.03), E12 (-4.35)	3
4_39200000_39220000	*TACR3*	E1 (-4.614), E5 (-4.421), E8 (-4.513)	3
4_77860000_77880000	*ENSGALG00000050913*	E6 (-4.21), E9 (-4.226), E10 (-4.41)	3
4_81460000_81480000	*ENSGALG00000053356*	E4 (-4.085), E9 (-4.181), E10 (-4.41)	3
9_12520000_12560000	*AGTR1*	E1 (-4.865), E8 (-4.177), E9 (-4.041)	3
10_6320000_6340000	*APBA2*	E2 (-4.366), E6 (-4.477), E10 (-4.154)	3
12_4170000_4190000	*ENSGALG00000030908*	E1 (-4.251), E2 (-4.523), E6 (-4.16)	3

18_10850000_10870000	*CASKIN2*	E4 (-4.245), E9 (-4.244), E10 (-4.214)	3
23_5660000_5680000	*TRIT1*	E5 (-4.871), E6 (-4.807), E9 (-4.081)	3
24_6150000_6180000	*BCO2, TEX12, IL18, SDHD*	E2 (-4.157), E10 (-5.027), E12 (-4.343)	3
4_75820000_75840000	*NCAPG, LCORL*	E3 (-4.451), E9 (-4.897), E10 (-4.544), E11 (-4.339)	4
7_19830000_19850000	*SCN2A*	E4 (-4.348), E6 (-4.772), E9 (-4.281), E11 (-4.196)	4
7_6950000_6970000	DIP2A	E2 (-4.183), E5 (-4.139), E6 (-4.793), E9 (-4.115)	4
18_10880000_10900000	*TSEN54, LLGL2*	E3 (-4.022), E5 (-4.3859), E11 (-4.232), E12 (-4.395)	4
5_40990000_41040000	*TSHR, GTF2A1*	E4 (-4.996), E5 (-5.339)*, E6 (-5.101)*, E10 (-5.192), E12 (-4.676)	5
6_19150000_19170000	*ENSGALG00000006137*	E1 (-4.0408), E2 (-4.925), E3 (-5.207), E6 (-4.906), E10 (-4.18)	5
5_41000000_41030000	*TSHR*	E1 (-5.388)*, E2 (-5.452)*, E3 (-6.04)*, E7 (-4.268)*, E9 (-5.574), E11 (-4.826)	6
12_8120000_8150000	*CACNA2D3*	E1 (-4.6), E3 (-4.051), E6 (-5.517), E9 (-4.626), E10 (-4.202), E12 (-5.078)*	6
5_22750000_22800000	*ENSGALG00000052351*	E1 (-4.714), E2 (-4.749), E3 (-4.641), E4 (-4.123), E5 (-4.694), E6 (-4.183), E7 (-4.035), E8 (-4.285), E9 (-4.314), E10 (-4.886), E11 (-4.534)	11

Table S5.4 Common *iHS* sweep regions between ecotypes.

Sweep regions	**Overlapping gene**	**Ecotypes shared (|*iHS*| value)**	**# ecotypes sharing candidate sweep**
2_102650000_102670000	*GATA6*	E11 (3.099), E12 (2.852)	2
2_109290000_109310000	*ENSGALG00000053102*	E1 (2.396), E3 (2.153)	2
2_109360000_109400000	*ENSGALG00000049655*	E1 (2.903), E3 (2.489)	2
2_147150000_147170000	*TSNARE1*	E2 (2.164), E5 (2.206)	2
2_147290000_147320000	*TSNARE1*	E1 (2.562), E2 (2.181)	2
2_41470000_41490000	*ENSGALG00000054423*	E7 (2.405), E9 (2.854)	2
2_44530000_44550000	*FBXL2*	E2 (2.318), E9 (2.566)	2
2_82850000_82940000	*ENSGALG00000051615*	E1 (2.393), E9 (2.429)	2
4_81600000_81620000	*ENSGALG00000015609*	E9 (2.344), E12 (2.268)	2
4_90360000_90380000	*ENSGALG00000053684*	E5 (2.388), E10 (2.026)	2
5_27510000_27550000	*MAP3K9*	E2 (2.345), E8 (2.349)	2
5_38240000_38270000	*FCF1*, *YLPM1*	E4 (2.243), E8 (2.407)	2
12_2390000_2410000	*DOCK3*	E5 (2.248), E9 (2.269)	2
12_8520000_8540000	*ERC2*	E1 (2.673), E8 (2.163)	2
21_3340000_3370000	*SPSB1*	E7 (2.381), E8 (2.304)	2
26_60000_110000	*TEAD3*, *TULP1*, *FKBP5*	E4 (2.14), E7 (2.823)	2
33_4120000_4140000	*MARS*	E2 (2.804), E8 (2.626)	2
6_9980000_10000000	*CHUK*	E3 (2.666), E5 (2.216), E8 (2.449)	3
33_6110000_6130000	*ENSGALG00000052298*	E3 (2.268), E5 (2.841), E9 (2.281)	3

Table S5.5.1. Candidate selective sweep regions detected in Ecotype 1 using the *Hp* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u*ZHp* ± s.d**.: mean and standard deviation for *ZHp* for selected windows within a region. GENE: annotated gene(s) overlapping selected region). *Red highlighted genes were found uniquely in this ecotype.

Chr	**Start**	**Stop**	**# win**	**Fragment (Kb)**	**u*ZHp* ± s.d.**	**GENE***
1	32730000	32750000	1	20	-4.08	
2	119410000	119460000	4	50	-4.39 ± 0.25	*ZFHX4*
2	119460000	119490000	2	30	-4.19 ± 0.00	*PEX2*
2	147130000	147160000	2	30	4.24 ± 0.31	*TSNARE1*
2	147370000	147390000	1	20	-4.20	*TSNARE1*
2	48280000	48300000	1	20	-4.31	*PDE1C*
2	50230000	50250000	1	20	-4.73	*SUGCT*
3	20100000	20120000	1	20	-4.02	*ESRRG*
3	21010000	21030000	1	20	-4.17	*KCNK2*
3	51550000	51570000	1	20	-4.63	*ARID1B*
3	69170000	69210000	3	40	-4.43 ± 0.27	
4	39200000	39220000	1	20	-4.61	*TACR3*
5	22750000	22800000	4	50	-4.71 ± 0.43	
5	41000000	41030000	2	30	-5.38 ± 0.74	*TSHR*
5	59680000	59700000	1	20	-4.12	*LRFN5*
6	19150000	19170000	1	20	-4.04	
7	19850000	19870000	1	20	-4.19	*SCN2A*
9	10440000	10460000	1	20	-4.16	*GRK7, ATP1B3*
9	12520000	12560000	3	40	-4.86 ± 0.13	*AGTR1*
9	23530000	23570000	2	40	-4.08 ± 0.05	*MBNL1*
12	4170000	4190000	1	20	-4.25	
12	8120000	8150000	2	30	-4.60 ± 0.30	*CACNA2D3*
14	5670000	5690000	1	20	-4.03	*CACNA1H*
18	620000	650000	2	30	-4.38 ± 0.43	*MYH1C*
19	9290000	9310000	1	20	-4.00	
22	3670000	3690000	1	20	-4.52	*TMEM127*
27	4660000	4680000	1	20	-4.15	*CRHR1*
27	7800000	7830000	3	30	-4.17 ± 0.18	*TUBG1, CCR10, CNTNAP1*
33	6470000	6570000	7	100000	-4.21 ± 0.16	*DRAP1, RELA, TMEM223, STX5, WDR74, SNORD30, SNORD22, TAF6L, MTA2, EML3*

Table S5.5.2. Candidate selective sweep regions detected in Ecotype 2 using the *Hp* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u*ZHp* ± s.d**.: mean and standard deviation for *ZHp* for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	**Start**	**Stop**	**# win**	**Fragment (kb)**	**u*ZHp* ± s.d.**	**GENE***
1	112390000	112440000	4	50	-4.47 ± 0.36	*EFHC2*
1	120930000	120960000	2	30	-4.55 ± 0.03	
1	185880000	185900000	1	20	-4.11	
1	32710000	32730000	1	20	-4.00	
1	89860000	89880000	1	20	-4.03	*TMPRSS7*
2	147370000	147390000	1	20	-4.42	*TSNARE1*
2	48280000	48300000	1	20	-4.13	*PDE1C*
4	20350000	20370000	1	20	-4.03	
4	20390000	20430000	4	40	-4.25 ± 0.06	
4	48210000	48240000	2	30	-4.16 ± 0.06	*ADGRL3*
4	48210000	48230000	1	20	-4.12	*gga-mir-1730*
4	75650000	75670000	1	20	-4.85	
5	21730000	21760000	2	30	-4.33 ± 0.07	*MIR129-2*
5	22750000	22800000	4	50	-4.74 ± 0.35	
5	41000000	41030000	2	30	-5.45 ± 0.36	*TSHR*
6	14580000	14600000	1	20	-4.07	*KCNMA1*
6	16240000	16290000	4	50	-4.46 ± 0.25	*VCL*
6	19150000	19170000	1	20	-4.92	
7	11670000	11700000	2	30	-4.16 ± 0.02	*FZD7*
7	19830000	19860000	2	30	-4.29 ± 0.08	*SCN2A*
7	20390000	20410000	1	20	-4.49	
7	29280000	29310000	2	30	-4.67 ± 0.14	*DPP10*
7	6950000	6970000	1	20	-4.18	*DIP2A*
8	24640000	24670000	2	30	-4.27 ± 0.14	*ZFYVE9*
9	23530000	23570000	3	40	-4.35 ± 0.26	*MBNL1*
10	18710000	18730000	1	20	-4.22	*MEGF11*
10	6320000	6340000	1	20	-4.36	*APBA2*
12	4170000	4190000	1	20	-4.52	
12	4790000	4830000	3	40	-4.55 ± 0.30	*ATG7*
12	7920000	7950000	2	30	-4.26 ± 0.06	*CACNA2D3*
24	6150000	6180000	2	30	-4.15 ± 0.20	*BCDO2*

Table S5.5.3. Candidate selective sweep regions detected in Ecotype 3 using the *Hp* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u*ZHp* ± s.d.**: mean and standard deviation for *ZHp* for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	Start	Stop	# win	Fragment (kb)	u*ZHp* ± s.d.	GENE*
1	117550000	117570000	1	20	-4.06	*IL1RAPL1*
1	32730000	32750000	1	20	-4.58	
1	32830000	32850000	1	20	-4.09	*TAFA2*
1	32940000	32960000	1	20	-4.29	
1	76440000	76470000	2	30	-4.30 ± 0.17	
1	83730000	83750000	1	20	-5.00	*DRD3*
2	118690000	118750000	3	60	-4.32 ± 0.23	*HNF4G*
2	118950000	118990000	3	40	-4.51 ± 0.12	
2	119410000	119440000	2	30	-4.09 ± 0.01	*ZFHX4*
2	130960000	130980000	1	20	-4.38	*HSF1*
2	141330000	141350000	1	20	-4.10	*EFR3A*
2	147100000	147160000	3	60	-4.37 ± 0.23	*TSNARE1*
2	48290000	48310000	1	20	-4.03	*PDE1C*
3	12500000	12530000	2	30	-4.31 ± 0.24	*CDC42BPA*
3	50260000	50290000	2	30	-4.34 ± 0.02	
3	50280000	50300000	1	20	-4.16	
3	51550000	51570000	1	20	-4.40	*ARID1B*
4	16400000	16420000	1	20	-4.15	*CUL4B*
4	20350000	20410000	5	60	-4.41 ± 0.25	*LRAT*
4	39150000	39220000	6	70	-5.28 ± 0.23	*TACR3*
4	47400000	47420000	1	20	-4.40	
4	48500000	48560000	5	60	-4.71 ± 0.24	
4	64540000	64560000	1	20	-4.07	*DLC1*
4	6650000	6680000	2	30	-4.23 ± 0.11	
4	75820000	75840000	1	20	-4.45	*NCAPG, LCORL*
5	12910000	12930000	1	20	-4.00	*PTPN5*
5	17910000	17930000	1	20	-4.10	
5	18180000	18200000	1	20	-4.37	*SHANK2*
5	21740000	21780000	3	40	-4.34 ± 0.22	*MIR129-2*
5	22750000	22790000	3	40	-4.64 ± 0.34	
5	23000000	23020000	1	20	-4.12	*C1QTNF4*, *FAM180B*
5	41000000	41030000	2	30	-6.04 ± 0.79	*TSHR*
5	41870000	41890000	1	20	-4.09	
5	49020000	49070000	4	50	-4.55 ± 0.39	
6	14580000	14610000	2	30	-4.06 ± 0.08	*KCNMA1*
6	19150000	19170000	1	20	-5.20	*ARHGAP22*
6	9980000	10000000	1	20	-4.13	*CHUK*
7	19820000	19850000	2	30	-4.31 ± 0.27	*SCN2A*
7	8270000	8310000	3	40	-4.27 ± 0.10	*TMEFF2*
8	10000000	10020000	2	20	-4.41 ± 0.12	
8	10310000	10330000	3	20	-4.51 ± 0.23	
8	420000	480000	4	60	-4.142± 0.09	*AMY2A*
8	510000	640000	11	130	-4.47 ± 0.14	
8	650000	700000	4	50	-4.31 ± 0.14	

8	9110000	9130000	1	20	-4.34	
8	9270000	9290000	1	20	-4.04	
10	6310000	6340000	2	30	-4.63 ± 0.22	*MPHOSPH10, APBA2*
11	13360000	13390000	2	30	-4.32 ± 0.16	
11	13790000	13810000	1	20	-4.21	*ADAMTS18, AP1G1, PHLPP2*
12	8130000	8150000	1	20	-4.05	*CACNA2D3*
14	8740000	8760000	1	20	-4.71	*SMG1*
18	10840000	10870000	2	30	-4.05 ± 0.06	*CASKIN2*
18	10880000	10900000	1	20	-4.02	*TSEN54, LLGL2*
22	990000	1020000	2	30	-4.73 ± 0.14	*DOCK5*
26	5260000	5280000	1	20	-4.08	*gga-mir-7454*
27	7800000	7830000	2	30	-4.63 ± 0.14	*TUBG1, CCR10, CNTNAP1*

Table S5.5.4. Candidate selective sweep regions detected in Ecotype 4 using the *Hp* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u*ZHp* ± s.d**.: mean and standard deviation for *ZHp* for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	**Start**	**Stop**	**# win**	**Fragment (kb)**	**u*ZHp* ± s.d.**	**GENE***
1	112400000	112430000	2	30	-4.30 ± 0.03	*EFHC2*
1	162460000	162480000	1	20	-4.05	
1	71630000	71650000	1	20	-4.15	
2	147140000	147160000	1	20	-4.80	*TSNARE1*
2	147300000	147320000	1	20	-4.50	*TSNARE1*
3	51610000	51630000	1	20	-4.00	*TMEM242*
3	61430000	61450000	1	20	-4.24	
3	77700000	77730000	2	30	-4.35 ± 0.24	*SNAP91, PRSS35*
3	85440000	85460000	1	20	-4.54	
4	81460000	81480000	1	20	-4.08	
5	22750000	22780000	2	30	-4.12 ± 0.08	
5	40990000	41040000	4	50	-4.99 ± 0.69	*TSHR, GTF2A1*
6	12300000	12320000	1	20	-4.54	*SGPL1*
7	17000000	17020000	1	20	-4.17	*OLA1*
7	19830000	19850000	1	20	-4.34	*SCN2A*
8	460000	480000	2	20	-4.19 ± 0.27	
8	560000	580000	1	20	-4.00	
8	600000	700000	9	100	-4.40 ± 0.26	
9	10420000	10460000	3	40	-5.22 ± 0.14	*RNF7, GRK7, ATP1B3*
11	1550000	1570000	1	20	-4.82	*HYDIN*
18	10850000	10870000	1	20	-4.24	*CASKIN2*
27	7860000	7960000	6	100	-4.22 ± 0.22	*CNTD1, BECN1, PSME3, G6PC, AARSD1, PTGES3L, RUNDC1, RPL27, VAT1, RND2*

Table S5.5.5. Candidate selective sweep regions detected in Ecotype 5 using the *Hp* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u*ZHp* ± s.d.**: mean and standard deviation for *ZHp* for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	Start	Stop	# win	Fragment (kb)	u*ZHp* ± s.d.	GENE*
1	162440000	162460000	1	20	-4.09	
1	163850000	163880000	2	30	-4.10 ± 0.07	*DIAPH3*
1	193490000	193510000	1	20	-4.34	*TENM4*
1	71880000	71900000	1	20	-4.18	*DUSP16*
1	72650000	72670000	1	20	-4.39	*MANSC4*
1	75580000	75600000	1	20	-4.11	*NANOG*
2	141290000	141310000	1	20	-4.27	*EFR3A*
2	141370000	141400000	2	30	-5.19 ± 0.10	*HHLA1*, *OC90*
2	147010000	147160000	12	150	-4.38 ± 0.31	*TSNARE1*
2	147370000	147410000	3	40	-4.11 ± 0.14	*TSNARE1*
3	51470000	51490000	1	20	-4.18	*ARID1B*
3	51610000	51630000	1	20	-4.17	*TMEM242*
3	51820000	51850000	2	30	-4.04 ± 0.05	*SNX9*
3	61550000	61590000	3	40	-4.31 ± 0.21	
3	85480000	85520000	3	40	-4.14 ± 0.07	
4	39200000	39220000	1	20	-4.42	*TACR3*
4	48230000	48260000	2	30	-4.29 ± 0.30	*ADGRL3*
5	22750000	22790000	3	40	-4.69 ± 0.36	
5	41000000	41040000	3	40	-5.33 ± 0.74	*TSHR*, *GTF2A1*
5	41180000	41200000	1	20	-4.08	
7	6950000	6970000	1	20	-4.13	*DIP2A*
9	11640000	11680000	3	40	-4.44 ± 0.35	*PLOD2*
11	18320000	18340000	1	20	-4.11	*ZNF469*
18	10770000	10790000	1	20	-4.66	*GRB2*
18	10880000	10900000	1	20	-4.38	*TSEN54*, *LLGL2*
18	40000	60000	1	20	-4.11	*ZNF302*
23	5660000	5680000	1	20	-4.87	*TRIT1*
27	7840000	7860000	1	20	-4.28	*EZH1*, *RAMP2*, *VPS25*
27	7880000	7930000	4	50	-4.38 ± 0.22	*PSME3*, *G6PC*, *AARSD1*, *PTGES3L*, *RUNDC1*, *RPL27*, *IFI35*
27	7940000	7970000	2	30	-4.05 ± 0.00	*VAT1*, *RND2*
30	940000	960000	1	20	-4.34	*ACTL6B*, *MEPCE*

Table S5.5.6. Candidate selective sweep regions detected in Ecotype 6 using the *Hp* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u*ZHp* ± s.d**.: mean and standard deviation for *ZHp* for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	Start	Stop	# win	Fragment (kb)	u*ZHp* ± s.d.	GENE*
1	117790000	117810000	1	20	-4.074	*IL1RAPL1*
1	132070000	132090000	1	20	-4.23	*OCA2*
1	176720000	176750000	2	30	-4.33 ± 0.01	*UBL3*
2	118700000	118720000	1	20	-4.04	*HNF4G*
2	119400000	119450000	4	50	-4.10 ± 0.06	*ZFHX4*
2	147140000	147160000	1	20	-4.17	*TSNARE1*
3	20100000	20120000	1	20	-4.03	*ESRRG*
3	50580000	50610000	2	30	-4.13 ± 0.13	*TIAM2*
3	69170000	69210000	3	40	-4.22 ± 0.28	
3	85440000	85460000	1	20	-4.01	
4	39160000	39190000	2	30	-4.13 ± 0.12	*TACR3*
4	42460000	42480000	1	20	-4.06	
4	77860000	77880000	1	20	-4.21	
4	81460000	81490000	2	30	-4.10 ± 0.06	
5	22750000	22780000	2	30	-4.18 ± 0.15	
5	3190000	3240000	4	50	-4.29 ± 0.20	*GAS2*
5	4010000	4030000	1	20	-4.05	*BDNF*
5	40990000	41040000	4	50	-5.10 ± 0.73	*TSHR*, *GTF2A1*
5	41200000	41240000	3	40	-4.09 ± 0.09	
5	59680000	59700000	1	20	-4.03	*LRFN5*
6	19150000	19170000	1	20	-4.90	
7	19840000	19870000	2	30	-4.77 ± 0.12	*SCN2A*
7	6950000	6970000	1	20	-4.79	*DIP2A*
8	10000000	10030000	2	30	-4.04 ± 0.01	
8	19380000	19410000	2	30	-4.28 ± 0.05	
8	9110000	9130000	1	20	-4.25	
9	10360000	10460000	8	100	-4.43 ± 0.26	*ZBTB38*, *RASA2*, *RNF7*, *GRK7*, *ATP1B3*
9	10550000	10570000	1	20	-4.23	*GK5*
10	6320000	6340000	1	20	-4.47	*APBA2*
12	4170000	4190000	1	20	-4.16	
12	8120000	8150000	2	30	-4.51 ± 0.21	*CACNA2D3*
15	520000	540000	1	20	-4.08	*HIC2*
20	13880000	1390000	1	20	-4.86	*SLC9A8*, *B4GALT5*
22	3670000	3690000	1	20	-4.66	*TMEM127*
23	5660000	5680000	1	20	-4.80	*TRIT1*
23	5730000	5750000	1	20	-4.03	*TCEB3*, *PITHD1*, *LYPLA2*, *GALE*, *HMGCL*, *FUCA1*
23	5850000	5890000	3	40	-4.02 ± 0.01	*GRHL3*
26	90000	140000	4	50	-4.20 ± 0.17	*CLPS*, *LHFPL5*, *SRPK1*, *SLC26A8*
26	190000	220000	2	30	-4.05 ± 0.02	*BRPF3*, *TBC1D22B*
26	250000	290000	3	40	-4.20 ± 0.12	*CACNA1S*, *KIF21B*
26	10000	110000	9	100	-4.21 ± 0.09	*MICAL1*, *SMPD2*, *TEAD3*, *TULP1*, *FKBP5*

27	7860000	7970000	10	110	-4.37 ± 0.13	*CNTD1, BECN1, PSME3, G6PC, AARSD1, PTGES3L, RUNDC1, RPL27, IFI35, VAT1, RND2*
27	8010000	8030000	1	20	-4.57	*MPP2*
33	6500000	6530000	2	30	-4.15 ± 0.03	*TMEM223, STX5, WDR74*
33	6550000	6570000	1	20	-4.5	*MTA2, EML3*

Table S5.5.7. Candidate selective sweep regions detected in Ecotype 7 using the *Hp* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u*ZHp* ± s.d**.: mean and standard deviation for *ZHp* for selected windows within a region. GENE: annotated gene(s) overlapping selected region).

Chr	**Start**	**Stop**	**# win**	**Fragment (kb)**	**u*ZHp* ± s.d.**	**GENE***
5	22750000	22780000	2	30	-4.03 ± 0.04	
5	41000000	41030000	2	30	-4.26 ± 0.19	*TSHR*

Table S5.5.8. Candidate selective sweep regions detected in Ecotype 8 using the *Hp* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u*ZHp* ± s.d**.: mean and standard deviation for *ZHp* for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	**Start**	**Stop**	**# win**	**Fragment (kb)**	**u*ZHp* ± s.d.**	**GENE***
1	112400000	112440000	3	40	-4.65 ± 0.36	*EFHC2*
2	141290000	141310000	1	20	-4.04	*EFR3A*
2	28130000	28150000	1	20	-4.08	
3	109890000	109910000	1	20	-4.03	*ADGRF5*
4	39200000	39220000	1	20	-4.51	*TACR3*
4	62730000	62770000	3	40	-4.44 ± 0.27	
5	22760000	22790000	2	30	-4.28 ± 0.01	
5	40990000	41030000	3	40	-4.53 ± 0.35	*TSHR*
9	12520000	12560000	4	40	-4.17 ± 0.02	*AGTR1*
12	4760000	4820000	5	60	-4.16 ± 0.07	*HRH1, ATG7*

Table S5.5.9. Candidate selective sweep regions detected in Ecotype 9 using the *Hp* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u*ZHp* ± s.d.**: mean and standard deviation for *ZHp* for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	Start	Stop	# win	Fragment (kb)	u*ZHp* ± s.d.	GENE*
1	120930000	120960000	2	30	-4.53 ± 0.02	
1	162440000	162480000	3	40	-4.40 ± 0.34	
1	193490000	193510000	1	20	-4.16	*TENM4*
1	25670000	25700000	2	30	-4.11 ± 0.07	
1	27290000	27330000	3	40	-4.29 ± 0.06	*DOCK4*
1	32700000	32750000	4	50	-4.87 ± 0.09	
1	75480000	75500000	1	20	-4.055	*FOXJ2*
1	75720000	75760000	2	40	-4.25 ± 0.23	*RIMKLB*
1	81050000	81070000	1	20	-4.22	*ATP1A1*
1	82470000	82490000	1	20	-4.14	*LSAMP*
2	147050000	147200000	12	150	-4.41 ± 0.26	*TSNARE1*
2	25650000	25670000	1	20	-4.08	
2	82030000	82060000	2	30	-4.19 ± 0.18	
2	82260000	82310000	4	50	-4.37 ± 0.18	*VSTM2A*
2	82620000	82640000	1	20	-4.69	
3	20090000	20120000	2	30	-4.60 ± 0.66	*ESRRG*
4	20350000	20400000	4	50	-4.402 ± 0.23	*LRAT*
4	56590000	56640000	4	50	-4.47 ± 0.05	*CAMK2D*
4	74600000	74620000	1	20	-4.14	
4	75820000	75840000	1	20	-4.89	*NCAPG, LCORL*
4	77860000	77880000	1	20	-4.22	
4	81460000	81480000	1	20	-4.18	
5	17910000	17940000	2	30	-4.22 ± 0.14	
5	18050000	18070000	1	20	-4.003	*ANO1*, *FADD*
5	19010000	19030000	1	20	-4.06	*EHF*
5	22760000	22790000	2	30	-4.31 ± 0.10	
5	2740000	2760000	1	20	-4.20	*NELL1*
5	2770000	2810000	2	40	-4.28 ± 0.26	
5	3030000	3080000	3	50	-4.19 ± 0.11	*SLC17A6*
5	3180000	3280000	7	100	-4.39 ± 0.24	*GAS2*
5	3360000	3400000	3	40	-4.28 ± 0.19	
5	3430000	3450000	1	20	-4.250	*ANO3*
5	3520000	3570000	3	50	-4.18 ± 0.14	*ANO3*
5	3710000	3730000	1	20	-4.42	*FIBIN*
5	3830000	3890000	3	60	-4.28 ± 0.18	*LGR4*
5	40520000	40570000	4	50	-4.70 ± 0.34	*NRXN3*
5	3990000	4030000	2	40	-4.09 ± 0.03	*BDNF*
5	41000000	41020000	1	20	-5.57	*TSHR*
5	48990000	49080000	7	90	-6.16 ± 0.14	*BEGAIN*
6	12290000	12320000	2	30	-4.65 ± 0.09	*MTTPL, SGPL1*
7	19840000	19870000	2	30	-4.28 ± 0.15	*SCN2A*
7	6950000	6970000	1	20	-4.11	*DIP2A*
8	19380000	19410000	2	30	-5.22 ± 0.51	*ZZZ3*
8	24640000	24670000	2	30	-4.19 ± 0.09	*ZFYVE9*

9	12520000	12560000	3	40	-4.04 ± 0.03	*AGTR1*
10	6300000	6340000	3	40	-4.25 ± 0.12	*MPHOSPH10, APBA2*
11	1420000	1440000	1	20	-4.007	*CNOT1*
11	1460000	1480000	1	20	-4.20	*CNOT1, SLC38A7, GOT2*
11	1550000	1570000	1	20	-5.35	*HYDIN*
12	4790000	4810000	1	20	-4.07	*ATG7*
12	8120000	8150000	2	30	-4.62 ± 0.39	*CACNA2D3*
15	3250000	3350000	9	100	-4.74 ± 0.13	*STX2*
15	4240000	4260000	1	20	-4.01	*TMEM132C*
18	10850000	10870000	1	20	-4.24	*CASKIN2*
20	13870000	13900000	2	30	-4.14 ± 0.58	*SLC9A8, B4GALT5*
23	5660000	5680000	1	20	-4.08	*TRIT1*
30	1430000	1450000	1	20	-4.28	
32	440000	470000	2	30	-4.78 ± 0.13	*SAE1*

Table S5.5.10. Candidate selective sweep regions detected in Ecotype 10 using the *Hp* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u*ZHp* ± s.d.**: mean and standard deviation for *ZHp* for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	**Start**	**Stop**	**# win**	**Fragment (kb)**	**u*ZHp* ± s.d.**	**GENE***
1	100330000	100350000	1	20	-4.28	
1	102640000	102660000	1	20	-4.51	
1	112400000	112430000	2	30	-4.24 ± 0.02	*EFHC2*
1	117520000	117550000	2	30	-4.19 ± 0.09	*IL1RAPL1*
1	162440000	162460000	1	20	-4.03	
1	172810000	172830000	1	20	-4.08	
1	32700000	32750000	4	50	-5.14 ± 0.45	
1	32940000	32960000	1	20	-4.29	
1	42400000	42430000	2	30	-4.53 ± 0.10	
1	97900000	97930000	2	30	-4.50 ± 0.58	*ROBO1*
1	97940000	97970000	2	30	-4.35 ± 0.19	*ROBO1*
1	97980000	98010000	2	30	-4.85 ± 0.39	*ROBO1*
1	98060000	98080000	1	20	-4.26	
1	98890000	98910000	1	20	-4.02	
2	123050000	123080000	3	30	-4.11 ± 0.05	*CPNE3, CNGB3*
2	141320000	141340000	1	20	-4.31	*EFR3A*
2	147080000	147120000	3	40	-4.82 ± 0.14	*TSNARE1*
2	147130000	147210000	7	80	-4.93 ± 0.79	*TSNARE1*
2	78320000	78340000	1	20	-4.18	
2	89440000	89460000	1	20	-4.31	
3	51450000	51480	2	30	-4.14 ± 0.02	*ARID1B*
3	69180	69210000	2	30	-5.28 ± 0.01	
4	24310000	24340000	2	30	-5.13 ± 0.13	
4	56580	56640000	5	60	-4.78 ± 0.10	*CAMK2D*
4	67000000	67030000	2	30	-4.25 ± 0.18	
4	67170000	67190000	1	20	-4.14	*GABRG1*
4	75170000	75190000	1	20	-4.04	*SLIT2*

4	75820000	75840000	1	20	-4.54	*NCAPG, LCORL*
4	77460000	77480	1	20	-4.67	*RAB28*
4	77610000	77630000	1	20	-4.86	
4	77860000	77880	1	20	-4.41	
4	7920000	7960000	3	40	-4.11 ± 0.07	
4	81460000	81480	1	20	-4.41	
4	81510000	81560000	4	50	-4.22 ± 0.12	
4	81460000	81480	1	20	-4.41	
5	14120000	14140000	1	20	-4.43	
5	17890000	17910000	1	20	-4.37	
5	19020000	19050000	2	30	-4.56 ± 0.14	*EHF*
5	21390000	21410000	1	20	-4.23	
5	22750000	22800	4	50	-4.88 ± 0.65	
5	31370000	31390000	1	20	-4.00	
5	41000000	41040000	3	40	-5.19 ± 0.96	*TSHR, GTF2A1*
5	4450000	4470000	1	20	-4.36	
5	47640000	47660000	1	20	-4.38	
5	47790000	47810000	1	20	-4.39	
5	47840000	47880	3	40	-4.71 ± 0.19	
6	15830000	15870000	3	40	-4.30 ± 0.25	*KAT6B*
6	19150000	19170000	1	20	-4.18	
8	10000000	10030000	2	30	-4.98 ± 0.14	
8	10320000	10350000	2	30	-4.44 ± 0.18	
8	9100000	9130000	2	30	-4.95 ± 0.35	
8	9270000	9300000	2	30	-4.52 ± 0.38	
8	9860000	9880	1	20	-4.23	
9	10340000	10360000	1	20	-4.02	*ZBTB38*
9	12530000	12550000	1	20	-4.12	*AGTR1*
9	23530000	23550000	1	20	-4.55	*MBNL1*
10	17640000	17680	3	40	-4.65 ± 0.49	*ADAMTS17*
10	6320000	6340000	1	20	-4.15	*APBA2*
11	18330000	18350000	1	20	-4.42	*ZNF469, ZFPM1*
12	8130000	8150000	1	20	-4.20	*CACNA2D3*
18	10850000	10870000	1	20	-4.21	*CASKIN2*
18	40000	60000	1	20	-4.54	*ZNF302*
19	1810000	1830000	1	20	-4.02	*AUTS2*
20	2180	2200000	1	20	-4.57	*E2F1, NECAB3*
22	1020000	1050000	2	30	-4.17 ± 0.10	*DOCK5*
22	990000	1010000	1	20	-4.25	*DOCK5*
24	6150000	6180	2	30	-5.02 ± 0.13	*BCO2, TEX12, IL18, SDHD*

Table S5.5.11. Candidate selective sweep regions detected in Ecotype 11 using the *Hp* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u*ZHp* ± s.d**.: mean and standard deviation for *ZHp* for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	**Start**	**Stop**	**# win**	**Fragment (kb)**	**u*ZHp* ± s.d.**	**GENE***
1	163730000	163790000	4	60	-4.22 ± 0.22	*DIAPH3*
2	48280	48310000	2	30	-4.17 ± 0.01	*PDE1C*
2	61030000	61060000	2	30	-4.38 ± 0.45	
2	78320000	78350000	2	30	-4.13 ± 0.08	
3	23810000	23840000	2	30	-4.47 ± 0.00	*EML4*
3	82190000	82220000	2	30	-4.87 ± 0.00	
4	27580	27610000	2	30	-4.11 ± 0.05	
4	30920000	30940000	1	20	-4.32	*ANAPC10, HHIP*
4	64550000	64570000	1	20	-4.20	*DLC1*
4	75820000	75840000	1	20	-4.33	*NCAPG, LCORL*
5	11710000	11730000	1	20	-4.08	
5	22750000	22790000	3	40	-4.53 ± 0.16	
5	41000000	41020000	1	20	-4.82	*TSHR*
6	11800	11820000	1	20	-4.23	
7	11820000	11840000	1	20	-4.57	*BMPR2*
7	19830000	19850000	1	20	-4.19	*SCN2A*
7	20930000	20960000	2	30	-4.51 ± 0.52	*KCNH7*
7	20980	21000000	1	20	-4.15	*KCNH7*

Table S5.5.12. Candidate selective sweep regions detected in Ecotype 12 using the *Hp* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u*ZHp* ± s.d.**: mean and standard deviation for *ZHp* for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	Start	Stop	# win	Fragment (kb)	u*ZHp* ± s.d.	GENE*
1	135140000	135170000	2	30	-4.78 ± 0.38	
1	180940000	180970000	2	30	-4.74 ± 0.91	*RDX*
1	58630000	58650000	1	20	-4.11	*SYT10*
2	119340000	119360000	1	20	-4.65	*ZFHX4*
2	147130000	147180	4	50	-5.03 ± 0.63	*TSNARE1*
2	31670000	31700000	2	30	-4.35 ± 0.04	*OSBPL3*
2	36820000	36850000	2	30	-4.26 ± 0.34	*UBE2E2*
2	41460000	41490000	2	30	-4.79 ± 0.88	*CAPN7*
2	69300000	69330000	2	30	-4.32 ± 0.06	*CDH6*
2	88150000	88170000	1	20	-4.04	*GABBR2*
3	106170000	106190000	1	20	-4.07	*HMBOX1*
3	20100000	20120000	1	20	-4.35	*ESRRG*
3	50280	50300000	1	20	-4.09	
3	51650000	51670000	1	20	-4.30	*ZDHHC14*
4	28660000	28690000	2	30	-4.24 ± 0.19	
4	39150000	39180	2	30	-4.21 ± 0.01	*TACR3*
4	48240000	48260000	1	20	-4.16	*ADGRL3*
4	6980	7000000	1	20	-4.42	
5	40990000	41040000	4	50	-4.67 ± 0.68	*TSHR, GTF2A1*
5	41180	41200000	1	20	-4.09	
6	11800	11820000	1	20	-4.39	
6	9720000	9740000	1	20	-4.12	*ANK3*
7	13540000	13560000	1	20	-4.30	*RAPH1*
7	19820000	19850000	2	30	-4.11 ± 0.14	*SCN2A*
7	24360000	24380	1	20	-4.03	
7	32510000	32610000	9	100	-4.23 ± 0.08	
8	1680	1700000	1	20	-4.30	*STXBP3*
9	10330000	10460000	12	130	-4.71 ± 0.5	*ZBTB38, RASA2, RNF7, GRK7, ATP1B3*
10	17670000	17690000	1	20	-4.01	*ADAMTS17*
12	4800	4830000	2	30	-4.19 ± 0.00	*ATG7*
12	8120000	8150000	2	30	-5.07 ± 0.46	*CACNA2D3*
12	8180	8200000	1	20	-4.55	
14	4960000	4980	1	20	-4.15	*USP22*
24	6150000	6180	2	30	-4.34 ± 0.44	*BCO2, TEX12, IL18, SDHD*

Table S5.6.1. Candidate selective sweep regions detected in Ecotype 1 using the *iHS* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u(|*iHS*|) ± s.d.**: mean and standard deviation for |*iHS*| for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	Start	Stop	# win	Fragment (kb)	u(|*iHS*|) ± s.d.	GENE*
1	112410000	112430000	1	20	2.63	*EFHC2*
1	118590000	118620000	2	30	2.11 ± 0.01	
1	162440000	162460000	1	20	2.94	
2	109290000	109310000	1	20	2.39	
2	109360000	109400000	3	40	2.90 ± 0.13	
2	147290000	147320000	2	30	2.56 ± 0.06	*TSNARE1*
2	55450000	55490000	3	40	2.76 ± 0.20	*SLC12A7*
2	82790000	82810000	1	20	2.16	
2	82850000	82940000	8	90	2.39 ± 0.08	
3	106170000	106210000	3	40	2.85 ± 0.21	*HMBOX1*
3	106250000	106270000	1	20	2.77	*KIF13B*
3	72070000	72140000	6	700	2.56 ± 0.19	*gga-mir-12262*
3	72300000	72370000	6	700	2.47 ± 0.05	
3	72380000	72540000	13	60	2.65 ± 0.20	*MMS22L, KLHL32*
3	73170000	73220000	4	50	3.48 ± 0.33	*MANEA*
3	73290000	73400000	8	110	2.82 ± 0.31	
4	15460000	15500000	3	40	2.33 ± 0.03	*TENM1*
4	34140000	34160000	1	20	2.30	*INTS10*, *HGSNAT*, *POMK*
4	39410000	39430000	1	20	2.48	
4	45640000	45660000	1	20	2.46	*PKD2*, *SPP1*
4	89960000	8999000	2	30	2.68 ± 0.20	*EXOC6B*
5	31330000	31350000	1	20	2.46	
5	31770000	31790000	1	20	2.69	
5	34280000	34300000	1	20	2.56	
7	29280000	29310000	2	30	2.11 ± 0.02	*DPP10*
7	32470000	3249000	1	20	2.41	
7	32730000	32750000	1	20	3.27	*ARHGAP15*
8	24640000	24660000	1	20	2.36	*ZFYVE9*
10	6260000	6290000	2	30	2.48 ± 0.06	*MPHOSPH10, APBA2*
12	8520000	8540000	1	20	2.67	*ERC2*
14	4960000	4980000	1	20	2.77	*USP22*
15	3220000	3250000	2	30	2.59 ± 0.07	*STX2*
15	3280000	3320000	3	40	2.51 ± 0.04	*STX2*
15	4180000	4210000	2	30	2.41 ± 0.004	*TMEM132C*
15	4230000	4270000	3	40	2.62 ± 0.05	*TMEM132C*
19	9230000	9290000	3	60	2.66 ± 0.07	*NF1, WSB1*
19	9320000	9370000	4	50	2.50 ± 0.08	*NOS2*
25	250000	270000	1	20	2.73	*SCNM1, TMOD4, VPS72, PIP5K1A*
27	7930000	7960000	2	30	2.59 ± 0.05	*VAT1, RND2*
28	4530000	4580000	3	50	3.05 ± 0.15	*RAB8A*, *TPM4*

30	270000	320000	4	50	3.11 ± 0.44	*BCAP31*, *TAZ*, *WDR13*, *WAS*, *TBC1D25*

Table S5.6.2. Candidate selective sweep regions detected in Ecotype 2 using the *iHS* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u(|*iHS*|) ± s.d.**: mean and standard deviation for |*iHS*| for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	**Start**	**Stop**	**# win**	**Fragment (kb)**	**u(|*iHS*|) ± s.d.**	**GENE***
1	102550000	102570000	1	20	2.08	
1	104470000	104550000	7	80	2.30 ± 0.03	
1	104960000	10499000	2	30	2.64 ± 0.06	
1	111430000	111450000	1	20	2.37	
1	11239000	112440000	3	50	2.49 ± 0.11	*EFHC2*
1	173180000	173220000	3	40	2.23 ± 0.07	*TRPC4*
1	173230000	173280000	4	50	2.29 ± 0.06	*TRPC4*
1	17599000	176030000	3	40	2.24 ± 0.03	*FRY*
1	18919000	189220000	2	30	2.17 ± 0.03	*TYR*
1	189240000	189300000	5	60	2.33 ± 0.03	*GRM5*
1	193460000	193500000	3	40	2.31 ± 0.006	*TENM4*
1	51030000	5109000	5	60	2.71 ± 0.17	*PICK1*, *SOX10*, *POLR2F*, *C22orf23*, *MICALL1*, *ANKRD54*
1	61270000	61330000	5	60	2.26 ± 0.01	*CACNA1C*
1	62130000	62150000	1	20	2.16	*AKR1E2*
1	69960000	70000000	3	40	2.61 ± 0.01	*PHF21B*
1	71630000	71670000	3	40	2.20 ± 0.05	*LRP6*
1	89850000	89910000	5	60	2.67 ± 0.12	*TMPRSS7*
2	100610000	100660000	4	50	2.52 ± 0.07	*DLGAP1*
2	108100000	108150000	4	50	2.43 ± 0.06	
2	109050000	109110000	5	60	2.35 ± 0.10	*SNTG1*
2	117740000	117770000	2	30	2.42 ± 0.16	*RDH10*
2	147150000	147170000	1	20	2.16	*TSNARE1*
2	14729000	147320000	2	30	2.18 ± 0.007	*TSNARE1*
2	44530000	44550000	1	20	2.31	*FBXL2*
2	47250000	4729000	3	40	2.28 ± 0.08	
2	59740000	59850000	8	110	2.46 ± 0.14	
2	61000000	61060000	5	60	2.42 ± 0.109	
2	88520000	88540000	1	20	2.20	*CTNNAL1*
2	91470000	91530000	5	60	2.56 ± 0.03	*ZNF407*
3	22380000	22400000	1	20	2.18	*TRAF5*, *RCOR3*
3	61410000	61450000	3	40	2.26 ± 0.03	
3	77660000	77730000	6	70	2.31 ± 0.07	*SNAP91*, *PRSS35*
5	12880000	12930000	4	50	2.27 ± 0.07	*PTPN5*
5	27510000	27550000	3	40	2.34 ± 0.01	
6	17670000	17700000	2	30	2.57 ± 0.04	*GBF1*
6	17710000	17760000	4	50	2.67 ± 0.05	*GBF1*
7	25020000	25040000	1	20	2.27	

8	5930000	5960000	2	30	2.40 ± 0.05	*KLHL20*, *CENPL*, *DARS2*
10	10660000	10680000	1	20	2.20	*FBN1*
13	7240000	7260000	1	20	2.23	
17	10380000	10570000	6	70	2.14 ± 0.06	*LMX1B*, *ZBTB43*
17	469000	4740000	4	50	2.29 ± 0.02	
17	9130000	9210000	7	80	2.28 ± 0.07	*RABGAP1*, *GPR21*, *STRBP*
17	9240000	9510000	2	20	2.38 ± 0.12	*STRBP*, *DENND1A*
17	9330000	9360000	2	30	2.56 ± 0.00	*U4*
17	9450000	9470000	1	20	2.26	
17	949000	9550000	5	60	2.33 ± 0.04	*LHX2*
19	329000	3310000	1	20	2.21	*UBE2G1*
25	220000	340000	11	120	2.51 ± 0.07	*PRUNE1*, *CDC42SE1*, *MLLT11*, *GABPB2*, *SEMA6C*, *SCNM1*, *TMOD4*, *VPS72*, *PIP5K1A*, *PSMD4*, *ZNF687*, *PI4KB*, *RFX5*
30	350000	380000	2	30	2.29 ± 0.04	*MECP2*
31	930000	950000	1	20	2.18	
33	3810000	3830000	1	20	2.42	*DCTN2*
33	4120000	4140000	1	20	2.80	

Table S5.63. Candidate selective sweep regions detected in Ecotype 3 using the *iHS* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u(|*iHS*|) ± s.d.**: mean and standard deviation for |*iHS*| for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	Start	Stop	# win	Fragment (kb)	u(|*iHS*|) ± s.d.	GENE*
1	1310000	1340000	2	30	2.18 ± 0.01	*IRF5*, *TNPO3*, *SMO*
1	178260000	178280	1	20	2.22	*ATP8A2*
1	187530000	187550000	1	20	2.33	*MTNR1B*
1	32700000	32730000	2	30	2.17 ± 0.00	
1	32810000	32830000	1	20	2.23	*TAFA2*
1	32850000	32870000	1	20	2.16	*TAFA2*
1	32940000	32960000	1	20	2.02	
2	10929000	109310000	1	20	2.15	
2	109360000	109400000	3	40	2.48 ± 0.08	
2	126220000	126240000	1	20	2.09	
2	147030000	14709000	5	60	2.24 ± 0.03	*TSNARE1*
2	147120000	147150000	2	30	2.09 ± 0.00	*TSNARE1*
2	30320000	30350000	2	30	2.71 ± 0.08	
5	12910000	13020000	10	110	2.33 ± 0.10	*PTPN5*, *PTPRJ*
5	1340000	1360000	1	20	2.11	*LUZP2*
5	13820000	13840000	1	20	2.36	*CD81*
5	21740000	22800000	3	60	2.31 ± 0.22	*MIR129-2*
6	12910000	12960000	4	50	2.39 ± 0.06	
6	13030000	13050000	1	20	2.24	*ZCCHC24*
6	9980000	10000000	1	20	2.66	*CHUK*
8	69000	710000	1	20	2.50	
9	17830000	1799000	15	160	2.41 ± 0.09	
10	17660000	17710000	4	50	2.31 ± 0.03	*ADAMTS17*
11	18320000	18340000	1	20	2.62	*ZNF469*
12	1319000	13220000	2	30	2.31 ± 0.05	*PTPRG*
14	4920000	4980	5	60	2.21 ± 0.03	*NATD1*, *TMEM11*, *DHRS7B*, *USP22*
14	5030000	5060000	2	30	2.50 ± 0.05	*USP22*, *TNFRSF13B*
27	7850000	7960000	9	110	2.28 ± 0.13	*RAMP2*, *VPS25*, *CNTD1*, *BECN1*, *PSME3*, *G6PC*, *AARSD1*, *PTGES3L*, *RUNDC1*, *RPL27*, *IFI35*, *VAT1*, *RND2*
30	1680000	1710000	2	30	2.19 ± 0.00	
30	260000	300000	3	40	2.21 ± 0.02	*BCAP31*, *TAZ*, *WDR13*, *WAS*, *TBC1D25*
33	4110000	4140000	2	30	2.29 ± 0.007	*MARS*
33	6110000	6130000	1	20	2.26	

Table S5.6.4. Candidate selective sweep regions detected in Ecotype 4 using the *iHS* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **(|*iHS*|) ± s.d.**: mean and standard deviation for |*iHS*| for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	**Start**	**Stop**	**# win**	**Fragment (kb)**	**u(|*iHS*|) ± s.d.**	**GENE***
1	163750000	163780	2	30	2.07 ± 0.01	*DIAPH3*
1	163820000	163840000	1	20	2.08	*DIAPH3*
1	189240000	189320000	7	80	2.27 ± 0.08	*GRM5*
1	58820000	58840000	1	20	2.10	*PKP2*
2	31710000	31730000	1	20	2.08	*OSBPL3*
2	60460000	6049000	2	30	2.10 ± 0.00	*KIF13A*
3	12470000	12530000	5	60	2.16 ± 0.03	*CDC42BPA*
3	41410000	41560000	12	150	2.30 ± 0.08	*THBS2*
4	39180000	39220000	3	40	2.13 ± 0.01	*TACR3*
4	39830000	39850000	1	20	2.25	*CDKN2AIP*
4	39860000	39910000	3	50	2.43 ± 0.02	*WWC2*
4	73720000	73740000	1	20	2.13	*DHX15*
5	38240000	38270000	2	30	2.24 ± 0.00	*FCF1, YLPM1*
5	55480000	55510000	2	30	2.18 ± 0.01	*ACTR10*
7	21460000	21480000	1	20	2.14	*TANK*
7	29240000	29310000	6	70	2.59 ± 0.10	*DPP10*
9	11660000	11680000	1	20	2.28	*PLOD2*
9	14330000	14370000	3	40	2.33 ± 0.07	
9	1439000	14410000	1	20	2.25	
9	14450000	14480000	2	30	2.074 ± 0.03	
13	17320000	17340000	2	20	2.10 ± 0.00	*FNIP1*
15	12530000	12560000	2	30	2.29 ± 0.08	
15	7000000	7030000	2	30	2.18 ± 0.02	*MYO1H*
20	1140000	1220000	7	80	2.15 ± 0.01	*ERGIC3, CEP250*
26	9000	130000	3	40	2.12 ± 0.04	*CLPS, LHFPL5, SRPK1*
26	150000	170000	2	20	2.17 ± 0.03	*MAPK14, MAPK13*
26	30000	50000	1	20	2.14	*MICAL1, SMPD2, TEAD3*
26	60000	110000	4	50	2.14 ± 0.02	*TEAD3, TULP1, FKBP5*
30	270000	300000	2	30	2.32 ± 0.01	*BCAP31, TAZ, WDR13, WAS, TBC1D25*

Table S5.6.5. Candidate selective sweep regions detected in Ecotype 5 using the *iHS* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u(|*iHS*|) ± s.d**.: mean and standard deviation for |*iHS*| for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	Start	Stop	# win	Fragment (kb)	u(|*iHS*|) ± s.d.	GENE*
1	120930000	120960000	2	30	2.22 ± 0.00	
1	1449000	14510000	1	20	2.16	
1	14640000	14700000	4	60	2.48 ± 0.10	*PRKAR2B*, *HBP1*, *COG5*
1	196050000	196080000	2	30	2.34 ± 0.01	*PPME1*
1	3090000	30930000	2	30	2.54 ± 0.03	*ARID2*
1	32040000	32160000	10	120	2.42 ± 0.08	
1	32810000	3289000	6	80	2.19 ± 0.01	*TAFA2*
1	32940000	32970000	2	30	2.23 ± 0.01	
1	33660000	3369000	2	30	2.51 ± 0.06	*SRGAP1*
1	440000	500000	5	60	2.74 ± 0.14	
1	58810000	58840000	2	30	2.18 ± 0.01	*PKP2*
1	61310000	61330000	1	20	2.39	*CACNA1C*
2	123110000	123130000	1	20	2.19	*CNGB3*
2	124730000	124760000	2	30	2.32 ± 0.03	*SLC26A7*
2	128260000	12829000	2	30	2.37± 0.09	*FBXO43*
2	129230000	129250000	1	20	2.18	*BAALC*
2	137440000	137460000	1	20	2.20	
2	13790000	137940000	3	40	2.68 ± 0.07	*DERL1*, *TBC1D31*
2	143700000	143730000	2	30	2.40 ± 0.11	
2	14549000	145670000	17	180	2.58 ± 0.207	*TRAPPC9*, *CHRAC1*, *gga-mir-6572*, *AGO2*
2	14690000	146950000	4	50	2.42 ± 0.04	*TSNARE1*
2	147010000	147060000	3	50	2.72 ± 0.02	*TSNARE1*
2	147110000	147140000	2	30	2.74 ±0.01	*TSNARE1*
2	147150000	147170000	1	20	2.20	*TSNARE1*
2	147370000	147410000	3	40	2.27 ± 0.02	*TSNARE1*
2	149240000	149260000	1	20	2.34	
2	55450000	55530000	7	80	2.64 ± 0.11	*SLC12A7*
2	5569000	55820000	10	130	2.50 ± 0.22	*SLC9A3*, *LRRC14B*
3	1249000	12510000	2	20	2.27 ± 0.05	*CDC42BPA*
4	23180000	23230000	4	50	2.34 ± 0.101	
4	27500000	27620000	6	120	2.21 ± 0.02	
4	27770000	27810000	3	40	2.42 ± 0.02	
4	27840000	27880000	3	40	2.46 ± 0.06	
4	27910000	27940000	2	30	2.38 ± 0.05	
4	28070000	28100000	2	30	2.25 ± 0.03	
4	28200000	2829000	8	90	2.41 ± 0.04	
4	28460000	28550000	8	90	2.30 ± 0.05	
4	90360000	90380000	2	20	2.38 ± 0.009	
5	28310000	28360000	4	50	2.42 ± 0.08	*EXD2*, *DCAF5*
5	2839000	28420000	2	30	2.30 ± 0.103	*DCAF5*

5	28470000	35540000	6	70	2.21 ± 0.06	*ACTN1, NPAS3*
6	10250000	10270000	1	20	2.34	*PTEN*
6	11660000	11700000	3	40	2.29 ± 0.03	
6	11810000	11830000	1	20	2.34	
6	1229000	12330000	3	40	2.27 ± 0.03	*MTTPL, SGPL1, PCBD1*
6	12360000	12430000	6	70	2.50 ± 0.07	
6	12520000	12540000	1	20	2.20	*CDH23*
6	9980000	10000000	1	20	2.21	*CHUK*
7	11820000	11860000	3	40	2.35 ± 0.04	*BMPR2*
7	36600000	36620000	1	20	2.45	*BAZ2B*
8	15600000	15620000	1	20	2.22	*PKN2*
8	670000	730000	5	60	2.29 ± 0.09	
9	23980000	24010000	2	30	2.19 ± 0.01	*WWTR1, TM4SF4*
10	6040000	6060000	1	20	2.27	*KLF13*
10	6310000	6340000	2	30	2.36 ± 0.06	*MPHOSPH10, APBA2*
11	16300000	16320000	2	20	2.26 ± 0.06	*CDH13*
12	239000	2410000	1	20	2.24	*DOCK3*
15	329000	3310000	1	20	2.26	*STX2*
15	4120000	4260000	9	140	2.45 ± 0.12	*TMEM132C*
15	4380000	4430000	3	50	2.32 ± 0.31	
18	3120000	3160000	3	40	2.69 ± 0.11	*B3GNTL1*
19	5830000	5850000	1	20	2.38	*FOXN1*
20	1100000	1120000	1	20	2.21	*CPNE1*
20	2260000	2280000	1	20	2.62	*SNTA1*
21	5160000	5180000	1	20	2.45	*PRDM2*
24	120000	140000	1	20	2.41	*VPS11, SIK2*
26	130000	180000	4	50	2.64 ± 0.03	*SLC26A8, MAPK14*
26	70000	100000	2	30	2.45 ± 0.03	*TULP1, FKBP5*
33	3170000	319000	1	20	2.16	*SCN8A*
33	3270000	329000	2	20	2.26 ± 0.02	*ACVRL1, ACVR1B, GRASP*
33	3310000	3350000	3	40	2.35 ± 0.08	*NR4A1, ATG101, FMNL3, TMBIM6*
33	6110000	6130000	1	20	2.84	
33	6310000	6340000	2	30	2.53 ± 0.09	*SIPA1*
33	6640000	6670000	2	30	2.73 ± 0.16	*BANF1*

Table S5.6.6. Candidate selective sweep regions detected in Ecotype 6 using the *iHS* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u(|*iHS*|) ± s.d.**: mean and standard deviation for |*iHS*| for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	**Start**	**Stop**	**# win**	**Fragment (kb)**	**u(|*iHS*|) ± s.d.**	**GENE***
1	11779000	117830000	3	40	2.38 ± 0.14	*IL1RAPL1*
1	149930000	149950000	1	20	2.16	
2	11949000	119510000	1	20	2.64	*PEX2*
2	65550000	65570000	1	20	2.68	*RPP40*
3	38540000	38570000	2	30	3.05 ± 0.12	*TARBP1*, *COA6*, *SLC35F3*
3	38980000	3900000	1	20	2.45	*PCNX2*, *U1*
3	85880000	8590000	1	20	2.21	
4	2849000	28520000	2	30	2.19 ± 0.08	
4	28530000	28550000	1	20	2.22	
4	30870000	3089000	1	20	2.48	*ANAPC10*, *HHIP*
4	40940000	40960000	1	20	2.33	
4	41000000	41020000	1	20	2.38	
4	41050000	4109000	3	40	2.22 ± 0.09	
4	41470000	4149000	1	20	2.51	
5	390000	3920000	1	20	2.45	*LGR4*
5	3970000	399000	1	20	2.42	*BDNF*
5	4120000	4180000	5	60	2.34 ± 0.23	*KIF18A*, *METTL15P1*
6	36170000	3619000	1	20	2.30	*INPP5A*
6	36240000	36300000	4	60	2.27 ± 0.06	*OPNVA*, *TCTN3*, *ALDH18A1*
7	32730000	3279000	5	60	2.96 ± 0.24	*ARHGAP15*
20	8240000	8260000	1	20	2.67	*NTSR1*
23	5730000	5750000	1	20	2.12	*TCEB3*, *PITHD1*, *LYPLA2*, *GALE*, *HMGCL*
26	60000	80000	1	20	2.16	*TEAD3*, *TULP1*, *FKBP5*

Table S5.6.7. Candidate selective sweep regions detected in Ecotype 7 using the *iHS* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u(|*iHS*|) ± s.d.**: mean and standard deviation for |*iHS*| for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	Start	Stop	# win	Fragment (kb)	u(|*iHS*|) ± s.d.	GENE*
1	10890000	109030000	10	130	2.12 ± 0.05	*ERG*
1	109100000	109120000	1	20	2.13	
1	109680000	109820000	11	140	2.10 ± 0.04	*DSCAM*
1	139070000	13909000	1	20	2.04	*MCF2L*
1	146760000	146780000	3	20	2.19 ± 0.05	*MBNL2*
1	189200000	189340000	10	140	2.15 ± 0.04	*TYR*, *GRM5*
1	197480000	197600000	11	120	2.37 ± 0.09	*TPP1*, *gga-mir-1600*, *LCMT2*
1	23040000	2309000	4	50	2.23 ± 0.03	*PTPRZ1*
1	23140000	23170000	2	30	2.24 ± 0.000	*PTPRZ1*
1	450000	510000	5	60	2.65 ± 0.09	
2	108110000	108140000	2	30	2.06 ± 0.00	
2	124510000	124540000	2	30	2.42 ± 0.15	*TMEM64*
2	20270000	20330000	5	60	2.30 ± 0.05	*MINDY3*
2	20740000	20760000	1	20	2.17	*ABCB1*
2	41470000	4149000	1	20	2.40	
2	43530000	43550000	1	20	2.13	
4	15440000	15500000	5	60	2.17 ± 0.02	*TENM1*
4	75560000	7559000	2	30	2.18 ± 0.03	
4	75670000	75710000	3	40	2.11 ± 0.06	
4	75830000	75850000	1	20	2.05	*NCAPG*, *LCORL*
4	75870000	7589000	1	20	2.05	*NCAPG*
5	11710000	11760000	4	50	2.42 ± 0.09	
5	11770000	11860000	8	90	2.48 ± 0.09	*C5H11orf58*, *PLEKHA7*
5	35230000	35250000	1	20	2.18	*AKAP6*
5	35470000	35670000	19	200	2.45 ± 0.07	*NPAS3*
7	31500000	31550000	4	50	2.31 ± 0.08	
7	34320000	34340000	1	20	2.26	
8	22850000	22930000	7	80	2.20 ± 0.02	
8	24770000	24800000	2	30	2.14 ± 0.03	*TUT4*, *GPX7*
9	14350000	14370000	1	20	2.11	
9	14430000	14470000	3	40	2.15 ± 0.01	
10	17620000	17710000	7	90	2.36 ± 0.08	*ADAMTS17*
10	18140000	18170000	2	30	2.22 ± 0.01	*ANKDD1A*, *SPG21*, *CLPX*
12	20230000	20270000	2	40	2.10 ± 0.045	*FGD5*
17	7460000	7480000	1	20	2.17	*BRD3*
17	819000	8210000	1	20	2.20	*SNAPC4*, *CARD9*
18	369000	3730000	3	40	2.34 ± 0.06	*CD7*
18	3750000	3780000	2	30	2.34 ± 0.02	*SEPTIN9*
20	1180000	1210000	2	30	2.32 ± 0.03	*ERGIC3*
21	3340000	3370000	2	30	2.38 ± 0.02	*SPSB1*

23	5070000	5100000	2	30	2.34 ± 0.13	*NFYC*, *gga-mir-30e*, *gga-mir-30c-1*, *KCNQ4*
26	9000	320000	20	230	2.55 ± 0.22	*CLPS*, *LHFPL5*, *SRPK1*, *SLC26A8*, *MAPK14*, *MAPK13*, *BRPF3*, *TBC1D22B*, *PIM1*, *CACNA1S*, *KIF21B*
26	20000	40000	1	20	2.61	*MICAL1*
26	60000	110000	4	50	2.82 ± 0.01	*TEAD3*, *TULP1*, *FKBP5*

Table S5.6.8. Candidate selective sweep regions detected in Ecotype 8 using the *iHS* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u(|*iHS*|) ± s.d.**: mean and standard deviation for |*iHS*| for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	**Start**	**Stop**	**# win**	**Fragment (kb)**	**u(|*iHS*|) ± s.d.**	**GENE***
1	110120000	110150000	3	30	2.28 ± 0.108	
1	11019000	110220000	2	30	2.24 ± 0.002	*BACE2*
1	110240000	110260000	1	20	2.16	*MX1*
1	112350000	112400000	4	50	2.20 ± 0.03	*FUNDC1*, *EFHC2*
1	112420000	112520000	6	100	2.23 ± 0.08	*EFHC2*, *U6*, *NDP*
1	120910000	121970000	5	60	2.21 ± 0.101	*CDKL5*
1	128340000	128360000	2	20	2.34 ± 0.01	*NLGN4Y*
1	57260000	57280000	1	20	2.21	*AGK*
1	58810000	58860000	4	50	2.21 ± 0.05	*PKP2*
1	65400000	65420000	1	20	2.21	*BCAT1*
2	138240000	13829000	4	50	2.15 ± 0.04	*FER1L6*
2	13909000	139140000	4	50	2.69 ± 0.05	
2	145400000	14549000	8	90	2.42 ± 0.05	*TRAPPC9*
2	43530000	43560000	2	30	2.33 ± 0.104	
2	8360000	8380000	1	20	2.17	
5	11710000	11730000	1	20	2.36	
5	1290000	12960000	5	60	2.28 ± 0.04	*PTPN5*
5	16940000	16960000	1	20	2.31	*FADS2*
5	27510000	27550000	3	40	2.34 ± 0.03	*MAP3K9*
5	28460000	28530000	6	70	2.46 ± 0.101	*ACTN1*, *SRSF5A*
5	31220000	31280000	5	60	2.26 ± 0.08	
5	38240000	38270000	2	30	2.40 ± 0.01	*FCF1*, *YLPM1*
5	39580000	39610000	2	30	2.13 ± 0.04	*SNW1*, *ADCK1*
6	16180000	16200000	2	20	2.22 ± 0.001	*ADK*
6	21550000	21610000	5	60	2.15 ± 0.02	*EXOC6*
6	7340000	7380000	3	40	2.13 ± 0.02	*DNAJC12*, *CTNNA3*
6	9980000	10000000	1	20	2.44	*CHUK*
8	2110000	2180000	6	70	2.22 ± 0.09	
8	23580000	23630000	4	50	2.40 ± 0.06	*BEND5*
8	23660000	23680000	1	20	2.13	*BEND5*

8	2369000	23720000	2	30	2.34 ± 0.005	*BEND5*
8	309000	3150000	5	60	2.38 ± 0.06	*KCNT2*
8	3260000	3340000	7	80	2.40 ± 0.06	
11	16980	17000000	1	20	2.41	*HNF4beta*
12	11780000	11800000	1	20	2.13	*SETD5*
12	7710000	7750000	3	40	2.43 ± 0.09	*CACNA2D3*
12	7760000	7840000	7	80	2.34 ± 0.12	*CACNA2D3*
12	7850000	7870000	1	20	2.23	*CACNA2D3*
12	7880000	7910000	2	30	2.31 ± 0.02	*CACNA2D3*
12	8520000	8540000	1	20	2.16	*ERC2*
15	10380000	10400000	1	20	2.67	*SCARF2*
15	10730000	10760000	2	30	2.17 ± 0.02	
15	8550000	8570000	1	20	2.43	*CRKL*
17	8260000	829000	2	30	2.22 ± 0.01	*GPSM1*
19	1770000	1840000	6	70	2.32 ± 0.04	*AUTS2*
19	2510000	259000	7	80	2.41 ± 0.07	
19	9220000	9350000	8	130	2.19 ± 0.13	*NF1*, *WSB1*, *NOS2*
20	350000	470000	10	120	2.24 ± 0.07	*DLGAP4*
21	3340000	3370000	2	30	2.30 ± 0.04	*SPSB1*
21	4110000	4200000	8	90	2.34 ± 0.06	*SRM*, *EXOSC10*, *MTOR*, *ANGPTL7*
21	4350000	4370000	1	20	2.36	*FBXO42*
21	4720000	4740000	1	20	2.46	*CAPZB*
21	4800000	4820000	1	20	2.24	*TMCO4*
33	4120000	4140000	1	20	2.62	*MARS*
33	5660000	569000	2	30	2.16 ± 0.03	*DIP2B*
33	5740000	5760000	1	20	2.14	*LIMA1*, *CERS5*

Table S5.6.9. Candidate selective sweep regions detected in Ecotype 9 using the *iHS* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u(|*iHS*|) ± s.d.**: mean and standard deviation for |*iHS*| for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	Start	Stop	# win	Fragment (kb)	u(|*iHS*|) ± s.d.	GENE*
1	139040000	139100000	5	60	2.52 ± 0.1003	*MCF2L*
1	25650000	25700000	4	50	2.20 ± 0.01	
1	6590000	65950000	4	50	2.49 ± 0.03	*SOX5*
1	66450000	66480000	2	30	2.323 ± 0.03	*ETNK1*
1	69740000	69760000	1	20	2.23	*PRR5*
1	71880000	71920000	3	40	2.50 ± 0.05	*DUSP16*
1	72010000	72030000	1	20	2.41	
1	7369000	73720000	2	30	2.21 ± 0.02	*NDUFA9*
1	73730000	73750000	1	20	2.34	
1	73830000	73930000	9	100	2.35 ± 0.11	
1	73960000	7399000	2	30	2.45 ± 0.01	*KCNA5*
1	74010000	74030000	1	20	2.65	
1	74110000	74140000	2	30	2.72 ± 0.002	

1	7419000	74210000	1	20	2.57	
1	74220000	74300000	7	80	2.60 ± 0.03	*NTF3*
1	74330000	7469000	34	360	2.58 ± 0.12	*ANO2*, *VWF*
1	74700000	74800000	8	100	2.69 ± 0.03	*CD9*
1	74860000	74930000	6	70	2.77 ± 0.05	*PRMT8*
1	75030000	75050000	21	20	2.70 ± 0.16	*TSPAN9*
1	75200000	75360000	14	160	2.67 ± 0.07	*TEAD4*, *TULP3*, *RHNO1*, *FOXM1*
1	75400000	75430000	2	30	2.39 ± 0.06	*FKBP4*
1	76440000	76470000	2	30	2.79 ± 0.06	
1	77140000	77180000	3	40	2.89 ± 0.04	
1	78330000	78350000	1	20	2.22	*CLCN1*, *CASP2*
1	89870000	8989000	1	20	2.40	*TMPRSS7*
2	147080000	147150000	5	70	2.24 ± 0.03	*TSNARE1*
2	41470000	4149000	1	20	2.85	
2	4369000	43710000	1	20	2.22	*TRAK1*
2	44530000	44550000	1	20	2.56	*FBXL2*
2	5650000	5670000	1	20	2.40	*SCN5A*
2	76160000	76180000	1	20	2.35	*ANKH*
2	82030000	82070000	3	40	2.35 ± 0.03	
2	82620000	82700000	5	80	2.38 ± 0.08	
2	8279000	82840000	4	50	2.47 ± 0.02	
2	82850000	82940000	8	90	2.42 ± 0.03	
3	5020000	5040000	1	20	2.51	
4	76280000	76310000	2	30	2.44 ± 0.01	*LDB2*
4	81460000	81560000	8	100	2.59 ± 0.11	
4	81600000	81620000	1	20	2.34	
4	9029000	90320000	2	30	2.28 ± 0.01	
4	90360000	9039000	1	30	2.31	
5	16710000	16810000	8	100	2.47 ± 0.05	*PPP6R3*, *GAL*, *TESMIN*
5	2529000	25310000	1	20	2.32	*MAPKBP1*
5	35400000	35450000	4	50	2.29 ± 0.03	*NPAS3*
5	35460000	35550000	8	90	2.42 ± 0.07	*NPAS3*
5	4119000	41220000	2	30	2.25 ± 0.04	
5	48700000	48820000	11	120	2.80 ± 0.14	*YY2*, *SLC25A29*, *WARS*, *WDR25*
5	48850000	48870000	1	20	2.69	*WDR25*
5	4879000	48820000	2	30	2.86 ± 0.009	
5	48850000	48880000	2	30	2.65 ± 0.05	*BEGAIN*
5	4889000	48920000	2	30	2.69 ± 0.01	*BEGAIN*
5	4909000	49110000	1	20	2.67	
5	49140000	49170000	2	30	2.69 ± 0.01	*DLK1*
5	4919000	49210000	1	20	2.65	
5	49310000	49350000	3	40	2.54 ± 0.06	
5	49360000	49550000	18	190	2.60 ± 0.12	*PPP2R5C*
6	2730000	2760000	2	30	2.57 ± 0.07	
6	279000	2860000	6	70	2.45 ± 0.02	
9	11640000	1169000	4	50	2.44 ± 0.05	*PLOD2*
9	12520000	12560000	3	40	2.29 ± 0.01	*AGTR1*
9	14360000	14380000	1	20	2.75	
9	14400000	14470000	7	70	2.53 ± 0.07	
9	14500000	14520000	1	20	2.33	*LPP*
9	1639000	16460000	6	70	2.42 ± 0.01	*MCF2L2*, *B3GNT5*

12	20110000	20130000	1	20	2.19	*TRH*, *RBSN*
12	2019000	20260000	6	70	2.27 ± 0.01	*FGD5*
12	239000	2410000	1	20	2.26	*DOCK3*
16	169000	1710000	1	20	2.77	
19	9240000	9260000	1	20	2.27	*WSB1*
23	4350000	4400000	4	50	3.12 ± 0.22	*AGO1*, *AGO4*, *CLSPN*, *C1orf216*
25	210000	250000	3	40	2.29 ± 0.014	*SETDB1*, *CERS2*, *MINDY1*, *PRUNE1*, *CDC42SE1*, *MLLT11*, *GABPB2*, *SEMA6C*
25	39000	460000	6	70	2.64 ± 0.09	*CGN*, *gga-mir-6620*, *TUFT1*, *SNX27*, *CELF3*, *MRPL9*, *TDRKH*, *RORC*
26	280000	310000	2	30	2.26 ± 0.045	*CACNA1S*, *KIF21B*
27	2510000	2550000	3	40	2.29 ± 0.002	*gga-mir-6644-2*
27	2700000	2730000	2	30	2.48 ± 0.11	
27	7850000	7930000	7	80	2.26 ± 0.03	*RAMP2*, *VPS25*, *CNTD1*, *BECN1*, *PSME3*, *G6PC*, *AARSD1*, *PTGES3L*, *RUNDC1*, *RUNDC1*, *IFI35*
32	420000	450000	3	30	2.28 ± 0.01	*SAE1*
32	520000	540000	1	20	2.55	*OPA3*
33	4110000	4130000	1	20	2.68	*MARS*
33	6110000	6130000	1	20	2.28	

Table S5.6.10. Candidate selective sweep regions detected in Ecotype10 using the *iHS* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u(|*iHS*|) ± s.d.**: mean and standard deviation for |*iHS*| for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	**Start**	**Stop**	**# win**	**Fragment (kb)**	**u(|*iHS*|) ± s.d.**	**GENE***
1	12570000	1259000	1	20	2.57	*PHTF2*
1	132340000	132360000	1	20	2.41	*GABRG3*
1	132930000	132960000	1	30	2.49	
1	133680000	133700000	1	20	2.30	*MGAT4A*
1	133820000	133860000	3	40	2.51 ± 0.11	*LYG2*, *TXNDC9*
1	176710000	176730000	1	20	2.22	*UBL3*
1	98880000	9890000	2	20	2.02 ± 0.00	
11	1420000	1480000	5	60	2.14 ± 0.04	*CNOT1*, *SNORA50A*, *SLC38A7*, *GOT2*
14	10330000	10350000	1	20	2.11	*USP7*
2	78280000	78310000	2	30	2.16 ± 0.006	*38777*
2	78320000	78340000	1	20	2.03	
2	79230000	79250000	1	20	2.07	*ADCY2*
3	51460000	51500000	3	40	2.35 ± 0.008	*ARID1B*

3	54930000	54970000	3	40	2.14 ± 0.009	*IL22RA2*, *IL20RA*
3	61450000	61480000	2	30	2.06 ± 0.01	
3	61550000	61620000	6	70	2.18 ± 0.05	
4	359000	3610000	1	20	2.36	*HS6ST2*
4	3620000	3640000	1	20	2.36	*HS6ST2*
4	90360000	90380000	1	20	2.02	
8	3260000	3330000	6	70	2.21 ± 0.01	
9	16430000	1649000	5	60	2.10 ± 0.01	*MCF2L2*, *B3GNT5*
9	2229000	22340000	4	40	2.57 ± 0.13	*KPNA4*, *gga-mir-1692*, *SCARNA7*, *TRIM59*, *SMC4*

Table S5.6.11. Candidate selective sweep regions detected in Ecotype11 using the *iHS* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u(|*iHS*|) ± s.d.**: mean and standard deviation for |*iHS*| for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	Start	Stop	# win	Fragment (kb)	u(|*iHS*|) ± s.d.	GENE*
1	156650000	156680	2	30	3.19 ± 0.01	
1	15699000	157010000	1	20	3.08	*KLF12*
2	100970000	10099000	1	20	2.67	*MYOM1*
2	101030000	101060000	2	30	2.67 ± 0.004	*EMILIN2*
2	101100000	101160000	4	60	2.99 ± 0.24	*SMCHD1*
2	101510000	101540000	2	30	2.93 ± 0.06	
2	101930000	101960000	2	30	3.65 ± 0.18	*YES1*, *CLUL1*
2	102020000	102050000	2	30	2.73 ± 0.08	*COLEC12*
2	102650000	102670000	1	20	3.09	*GATA6*
2	102820000	102840000	1	20	2.79	*RBBP8*
2	102870000	10290000	2	30	3.24 ± 0.08	
2	104180000	104200000	1	20	2.97	*TAF4B*
2	104220000	104270000	2	50	3.22 ± 0.46	*TAF4B*, *KCTD1*
2	104750000	104770000	1	20	2.67	
2	109920000	109940000	1	20	3.03	*RB1CC1*
2	109960000	109980000	1	20	2.67	*RB1CC1*
2	110230000	110280000	3	50	2.86 ± 0.16	*ATP6V1H*
2	110310000	110330000	1	20	3.08	*RGS20*
2	110400000	110420000	1	20	2.78	
2	111120000	111140000	1	20	2.69	*LYN*
2	112860000	112910000	2	50	2.79 ± 0.16	*RAB2A*, *gga-mir-1557*
2	114700000	114720000	1	20	3.00	
2	89080000	89100000	1	20	2.74	*MSANTD3*
2	89150000	89170000	1	20	2.65	*TEX10*, *INVS*
2	89220000	89250000	2	30	2.79 ± 0.14	*INVS*
2	89360000	89380000	1	20	2.67	*NR4A3*
2	91320000	91360000	3	40	2.94 ± 0.13	*TSHZ1*
2	91720000	91750000	2	30	3.09 ± 0.26	*ZNF407*

2	9939000	99410000	1	20	3.14	*PTPRM*
4	51830000	51850000	1	20	2.75	
4	5189000	51940000	4	50	2.91 ± 0.104	
4	8229000	82330000	3	40	2.68 ± 0.01	*ADD1*
4	82430000	82460000	2	30	2.67 ± 0.03	*TNIP2*
4	82800000	82830000	2	30	3.22 ± 0.005	*gga-mir-7467*
4	82860000	8289000	2	30	2.76 ± 0.06	*MXD4*, *HAUS3*
4	8319000	83210000	1	20	2.82	*NELFA*
4	83330000	83350000	1	20	2.65	*LETM1*
4	83600000	83630000	2	30	2.90 ± 0.04	
4	8379000	83810000	1	20	2.69	
4	83870000	8389000	1	20	2.90	*TACC3*
4	83930000	83950000	1	20	2.94	*SLBP*
4	83960000	8399000	2	30	2.83 ± 0.09	*FAM53A*
4	84250000	84270000	1	20	2.76	
4	84410000	84430000	1	20	2.71	*NKX1-1*
4	84630000	84650000	1	20	2.67	*CTBP1*
4	85100000	85150000	4	50	3.16 ± 0.12	*FGFRL1*
4	86040000	87050000	8	110	3.12 ± 0.25	
4	88500000	88540000	3	40	2.67 ± 0.02	*RNF24*, *gga-mir-103-2*, *MAVS*
4	88720000	88770000	4	50	2.72 ± 0.08	*DDRGK1*, *gga-mir-146c*
4	88950000	88980000	2	30	2.96 ± 0.09	*SLC4A11*
4	89030000	89060000	2	30	2.67 ± 0.01	
4	89250000	89270000	1	20	2.67	*ATRN*
4	89310000	89330000	1	20	2.67	*ATRN*, *GFRA4*
4	89760000	89780000	1	20	2.83	*SFXN5*
4	89810000	89880000	6	70	3.26 ± 0.33	*SFXN5*, *EMX1*, *SPR*
4	9030000	9060000	2	30	3.07 ± 0.17	*EXOC6B*, *gga-mir-12276*
4	9080000	90110000	2	30	2.75 ± 0.01	*EXOC6B*
5	30180000	30200000	1	20	2.70	
5	30220000	30240000	1	20	2.64	*KATNBL1*
5	30830000	30870000	3	40	2.86 ± 0.09	*ARHGAP11B*
5	30880000	30940000	4	60	2.91 ± 0.24	*RASGRP1*, *FAM98B*
5	30960000	31010000	4	50	3.15 ± 0.2	*SPRED1*
5	3199000	32020000	2	30	2.81 ± 0.12	
5	32060000	32080000	1	20	2.68	
5	32380000	32400000	1	20	2.64	*ZNF770*
5	36120000	36150000	2	30	2.79 ± 0.17	*EAPP*, *SNX32*
5	36330000	36420000	8	90	3.18 ± 0.2	*PSMA6*, *NFKBIA*
5	37360000	37380000	1	20	2.64	*FOXA1*
5	3779000	37850000	5	60	3.05 ± 0.24	*SEC23A*, *GEMIN2*, *TRAPPC6A*, *PNN*
5	3809000	38110000	1	20	2.70	*VRTN*, *SYNDIG1L*
5	38300000	38330000	2	30	3.01 ± 0.09	*DLST*, *RPS6KL1*, *PGF*
5	38430000	38480000	4	50	3.03 ± 0.09	*FOS*, *JDP2*
5	7180000	7200000	1	20	2.76	*HIPK3*
5	749000	7510000	1	20	2.68	*SYT9*
5	8260000	8280000	1	20	2.69	*MICAL2*
8	5670000	5700000	2	30	2.84 ± 0.02	*PBX1*

10	3380000	3420000	2	40	2.71 ± 0.07	*HMG20A*
10	3570000	3610000	2	40	2.99 ± 0.2	*TSPAN3*, *PSTPIP1*
10	3640000	369000	4	50	2.90 ± 0.11	*PSTPIP1*, *RCN2*, *SCAPER*
10	3770000	3860000	7	90	3.22 ± 0.19	*SCAPER*, *ISL2*, *ETFA*, *TMEM266*
10	3870000	3950000	6	80	3.07 ± 0.24	*TMEM266*, *NRG4*, *FBXO22*
10	4020000	4040000	1	20	2.86	
10	4050000	4070000	1	20	3.89	*CHRNB4*
10	409000	4110000	1	20	3.77	*CHRNA5*, *PSMA4*, *HYKK*
10	4120000	4150000	2	30	2.81 ± 0.22	*IREB2*, *CRABP-I*
10	4170000	4200000	2	30	3.00 ± 0.23	*WDR61*, *DNAJA4*, *ACSBG1*
10	4260000	429000	2	30	2.78 ± 0.03	
10	439000	4420000	2	30	2.89 ± 0.09	*HERC1*
10	449000	4570000	6	80	3.02 ± 0.1	*USP3*, *CA12*, *RAB8B*
10	5300000	5330000	2	30	2.85 ± 0.1	*RORA*
10	5380000	5400000	1	20	2.69	*RORA*
10	5410000	5430000	1	20	2.91	*RORA*
10	5470000	549000	1	20	2.71	*RORA*
10	5810000	5840000	2	30	2.71 ± 0.07	*BNIP2*
10	7170000	7210000	3	40	2.88 ± 0.17	*FAM81A*, *MYO1E*
10	8080000	8120000	3	40	2.80 ± 0.07	*MNS1*, *TEX9*
10	8210000	829000	6	80	2.88 ± 0.16	*NEDD4*, *PRTG*
10	8400000	8450000	2	50	3.15 ± 0.19	*DYX1C1*, *CCPG1*, *PIGB*, *PIGBOS1*, *RAB27A*
10	8460000	849000	2	30	2.80 ± 0.17	*RAB27A*, *RSL24D1*
10	859000	869000	9	100000	3.12 ± 0.24	
10	8730000	8880000	11	150	2.93 ± 0.2	
10	9120000	9170000	4	50	3.19 ± 0.02	*FAM214A*
10	9180000	929000	9	110	3.30 ± 0.29	*ARPP19*, *MYO5A*
10	9320000	9360000	3	40	3.07 ± 0.17	*GNB5*, *BCL2L10*
10	9370000	939000	1	20	3.21	*MAPK4*
10	9480000	9500000	1	20	2.93	*TMOD3*, *TMOD2*
10	9560000	9650000	6	90	3.17 ± 0.25	*SCG3*, *AP4E1*, *TNFAIP8L3*
10	9680000	9700000	1	20	3.08	*CYP19A1*
10	9720000	9740000	1	20	2.75	
10	9800000	9830000	2	30	2.75 ± 0.11	
10	9880000	9940000	4	60	3.05 ± 0.31	
10	9950000	9970000	1	20	3.01	
10	10210000	1029000	7	80	2.89 ± 0.09	*SEMA6D*
10	10320000	10350000	2	30	2.87 ± 0.32	*SEMA6D*
10	10540000	10560000	1	20	2.65	*SLC12A1*
10	10820000	10840000	1	20	2.94	
10	11110000	11130000	1	20	2.68	*C15orf48*, *gga-mir-147*, *SLC30A4*
10	11500000	11530000	2	30	2.85 ± 0.15	*PDE8A*, *gga-mir-3540*
10	11660000	11710000	3	50	3.45 ± 0.54	*TM6SF1*, *HDGFL3*

10	11700000	11830000	9	130	3.01 ± 0.36	*HDGFL3*, *BNC1*
10	11840000	11920000	3	80	2.73 ± 0.05	*SH3GL3*
10	12460000	1249000	2	30	2.76 ± 0.07	
10	12510000	12550000	3	40	3.05 ± 0.19	*TMC3*, *STARD5*, *IL16*
10	12750000	1279000	3	40	2.93 ± 0.11	
10	12940000	13050000	7	110	3.49 ± 0.47	
10	13070000	13100000	2	30	3.11 ± 0.01	*ZFAND6*, *BCL2A1*, *MTHFS*
10	13230000	13250000	1	20	2.89	
10	13270000	1339000	9	120	2.95 ± 0.16	*KIF7*, *TICRR*, *RHCG*, *gga-mir-9-3*, *POLG*
10	1349000	13510000	1	20	2.64	*ACAN*
10	13530000	13610000	5	80	3.43 ± 0.54	*ACAN*, *ISG20*, *gga-mir-1720*, *gga-mir-7-2*
10	13650000	13670000	1	20	3.40	*NTRK3*
10	13710000	13770000	5	60	3.35 ± 0.25	*NTRK3*
10	13780000	13820000	3	40	3.39 ± 0.43	*NTRK3*
10	13880000	1390000	1	20	2.90	
10	13920000	13970000	3	50	2.71 ± 0.08	
10	14060000	14080000	1	20	2.99	*AGBL1*
10	14250000	1429000	3	40	3.59 ± 0.16	*AGBL1*
10	14340000	1439000	3	50	3.31 ± 0.5	
10	14400000	14480000	6	80	3.23 ± 0.17	*KLHL25*
10	14540000	14560000	1	20	2.83	
14	2110000	2150000	3	40	2.87 ± 0.1	*FAM20C*
14	2630000	2650000	1	20	2.69	*COX19*, *CYP2AC7*
14	279000	2810000	1	20	2.79	
14	2820000	2840000	1	20	3.50	
14	3180000	3210000	3	30	2.76 ± 0.11	*MAD1L1*
14	4530000	4560000	2	30	3.12 ± 0.01	*FBXL18*
14	5300000	5320000	1	20	2.75	*SREBF1*, *gga-mir-33-2*, *TOM1L2*
14	6200000	6230000	1	30	2.67	
14	6250000	6270000	1	20	2.80	
14	689000	6920000	2	30	2.96 ± 0.16	*BRICD5*, *PGP*, *E4F1*, *ZDHHC4*, *GRID2IP*
14	7140000	7210000	6	70	2.82 ± 0.15	*CACNG3*, *RBBP6*
17	5920000	5950000	2	30	2.88 ± 0.08	*NTMT1*, *ASB6*
17	5980000	6010000	2	30	2.96 ± 0.09	*PTGES*, *TOR1B*, *TOR1A*, *C9orf78*, *USP20*
17	6020000	6040000	1	20	2.64	*USP20*, *FNBP1*
17	6150000	6200000	4	50	3.07 ± 0.25	*GPR107*, *NCS1*, *ASS1*
17	6340000	6360000	1	20	2.76	*FIBCD1*
17	6400000	6420000	1	20	2.76	*NUP214*
17	6620000	6660000	3	40	2.96 ± 0.07	*MED27*
17	6720000	6750000	2	30	2.91 ± 0.06	*NTNG2*
17	7030000	7060000	2	30	2.69 ± 0.03	*GFI1B*

17	7070000	7120000	4	50	3.07 ± 0.12	*CEL*, *GTF3C5*
17	7220000	7260000	4	40	2.86 ± 0.12	*CACFD1*, *SLC2A6*, *TMEM8C*, *ADAMTSL2*
17	7340000	7360000	1	20	2.81	*SARDH*, *VAV2*
17	7540000	7600000	3	60	2.78 ± 0.06	*RXRA*
17	7630000	7660000	2	30	2.89 ± 0.03	*RXRA*
17	7750000	7780000	2	30	2.95 ± 0.03	*COL5A1*
17	7820000	7850000	2	30	2.79 ± 0.11	*COL5A1*
17	7880000	7920000	3	40	2.77 ± 0.05	*OLFM1*
17	8030000	8070000	3	40	2.78 ± 0.04	*DIPK1B*, *AGPAT2*, *EGFL7*, *gga-mir-126*
17	809000	8110000	1	20	2.95	*NOTCH1*
17	8560000	8580000	1	20	2.70	*PSMD5*, *B3GNTL*
20	4040000	4060000	1	20	2.96	
20	4160000	4180000	1	20	2.67	
20	4220000	4240000	1	20	2.73	
27	6870000	689000	1	20	2.64	*LASP1*, *FBXO47*
27	7100000	7130000	2	30	2.65 ± 0.01	*THRA*, *ORMDL3*, *GSDMA*, *PSMD3*, *CSF3*
27	7450000	7470000	1	20	2.65	*EIF1*
28	2700000	2720000	1	20	2.68	*SIRT6*, *CREB3L3*, *MAP2K2*
28	820000	840000	1	20	2.74	*RANBP3*
28	910000	930000	1	20	2.93	*37316*, *RAB11B*

Table S5.6.12. Candidate selective sweep regions detected in Ecotype12 using the *iHS* method. (**Chr:** chromosome. **# win**: number of analyzed windows within the region. **Fragment (kb)**: fragment size for the sweep regions in kilobase pair. **u(|*iHS*|) ± s.d.**: mean and standard deviation for |*iHS*| for selected windows within a region. GENE: annotated gene(s) overlapping selected region). * Red highlighted genes were found uniquely in this ecotype.

Chr	**Start**	**Stop**	**# win**	**Fragment (kb)**	**u(|*iHS*|) ± s.d.**	**GENE***
1	45640000	45660000	1	20	2.90	*NTN4*
1	87060000	87080000	1	20	2.58	
2	102650000	102670000	1	20	2.85	*GATA6*
2	107670000	10769000	1	20	2.52	
2	137860000	13789000	2	30	2.76 ± 0.15	*ZHX2*
2	55320000	55340000	1	20	3.06	
2	78220000	78240000	1	20	2.79	*ANKRD33B*
2	78980000	7900000	1	20	2.59	
2	79300000	79320000	1	20	2.52	*ADCY2*
2	97200000	97220000	1	20	2.67	*PRELID3A*, *AFG3L2*
2	98860000	9889000	2	30	2.83 ± 0.007	*RAB12*

3	11920000	11970000	4	50	2.86 ± 0.18	*CNRIP1*, *PLEK*, *FBXO48*, *APLF*
3	12040000	12060000	1	20	2.59	
3	12180000	12240000	5	60	2.68 ± 0.1	*PCSK2*
3	16640000	16670000	2	30	2.55 ± 0.05	*SRSF7*, *MORN2*
3	16870000	1690000	2	30	2.69 ± 0.06	*GPCPD1*
3	31560000	31620000	5	60	2.72 ± 0.36	*TTC27*, *BIRC6*
3	3219000	32260000	4	70	2.60 ± 0.13	*STRN*, *VIT*
3	3379000	33950000	6	160	2.61 ± 0.11	*MEMO1*, *AHCTF1*
3	72080000	72140000	4	60	2.17 ± 0.03	
4	15440000	15530000	8	90	2.40 ± 0.05	*TENM1*
4	15440000	15460000	1	20	2.45	
4	39150000	39180000	2	30	2.04 ± 0.02	*TACR3*
4	81460000	81510000	2	50	2.17 ± 0.02	
4	81600000	81620000	1	20	2.26	
5	19460000	19620000	15	160	2.99 ± 0.46	*LDLRAD3*
5	59000	610000	1	20	2.78	
6	28530000	28600000	6	70	2.20 ± 0.02	*NHLRC2*
7	30940000	30970000	2	30	2.38 ± 0.03	*THSD7B*
7	31130000	31150000	1	20	2.91	*THSD7B*
7	33250000	33280000	2	30	3.21 ± 0.02	*GTDC1*
7	33380000	33400000	1	20	2.87	*ZEB2*
7	34920000	34950000	2	30	3.04 ± 0.28	*MMADHC*
7	34960000	34980000	1	20	2.87	
8	619000	6220000	2	30	2.45 ± 0.15	*STX10*, *XPR1*
8	6340000	6370000	2	30	3.10 ± 0.008	*ACBD6*, *LHX4*, *QSOX1*
8	6450000	6500000	4	50	2.57 ± 0.05	*TOR1AIP2*, *FAM163A*
9	21700000	21720000	1	20	2.41	
11	10130000	10150000	1	20	2.55	*GPATCH1*
11	10310000	10350000	3	40	2.70 ± 0.121	*CEBPG*, *PEPD*
11	9260000	9280000	1	20	2.87	*TSHZ3*
15	269000	2710000	1	20	2.14	*GALNT9*
19	140000	170000	2	30	2.25 ± 0.03	*DDX25*
21	6570000	6680000	3	110	2.37 ± 0.08	*USP48*, *RAP1GAP*, *EIF4G3*
21	6760000	6780000	1	20	2.54	*DDOST*
22	169000	1770000	6	80	2.68 ± 0.07	*XPO7*, *DOK2*, *GFRA2*

Table S5.7. Candidate selection signature regions (SRs) from ecotype analyses overlapping with extreme population groups analyses (chapter 4).

SR_ID_ecotype	Size (kb)	detection method	detected ecotypes	SR chapter 4	environmental parameter	Overlapping genes
chr1:65400000_65420000	20	*iHS*	E8	chr1:65410000_65430000	Landuse	*BCAT1*
chr1:110240000_110260000	20	*iHS*	E8	chr1:110220000_110270000	Landuse	*MX1*
chr1:112390000_112440000	50	*ZHp and iHS*	E2	chr1:112400000_112430000	Landuse	*EFHC2*
chr1:112400000_112430000	30	*ZHp*	E4, E10	chr1:112400000_112430000	Landuse	*EFHC2*
chr1:112400000_112440000	40	*ZHp*	E8	chr1:112400000_112430000	Landuse	*EFHC2*
chr1:112410000_112430000	20	*iHS*	E1	chr1:112400000_112430000	Landuse	*EFHC2*
chr1:112420000_112520000	100	*iHS*	E8	chr1:112400000_112430000	Landuse	*EFHC2, U6, NDP*
chr1:112420000_112520000	100	*iHS*	E8	chr1:112490000_112520000	Landuse	*EFHC2, U6, NDP*
chr1:128340000_128360000	20	*iHS*	E8	chr1:128350000_128470000	Landuse	
chr1:128460000_128510000	40	*ZHp*	E5, E12	chr1:128350000_128470000	Landuse	
chr1:128460000_128520000	60	*ZHp*	E4	chr1:128350000_128470000	Landuse	
chr1:128460000_128540000	80	*ZHp*	E10	chr1:128350000_128470000	Landuse	
chr1:128460000_128520000	60	*ZHp*	E4	chr1:128510000_128560000	Landuse	
chr1:128460000_128540000	80	*ZHp*	E10	chr1:128510000_128560000	Landuse	
chr1:128530000_128810000	280	*iHS*	E5	chr1:128510000_128560000	Landuse	
chr2:82260000_82310000	50	*ZHp*	E9	chr2:82270000_82290000	Landuse	
chr2:82340000_82630000	290	*ZHp*	E3	chr2:82330000_82360000	Landuse	
chr2:82330000_82640000	290	*ZHp*	E9	chr2:82330000_82360000	Landuse	
chr2:82340000_82630000	290	*ZHp*	E3	chr2:82370000_82390000	Landuse	
chr2:82330000_82640000	290	*ZHp*	E9	chr2:82370000_82390000	Landuse	
chr2:82340000_82630000	290	*ZHp*	E3	chr2:82580000_82600000	Landuse	

chr2:82330000_82640000	290	*ZHp*	E9	chr2:82580000_82600000	Landuse	
chr2:82620000_82640000	20	*ZHp*	E9	chr2:82630000_82770000	Landuse	
chr2:82620000_82700000	80	*iHS*	E9	chr2:82630000_82770000	Landuse	
chr2:82330000_82640000	290	*ZHp*	E9	chr2:82630000_82770000	Landuse	
chr2:82790000_82810000	20	*iHS*	E1	chr2:82790000_82840000	Landuse	
chr2:82790000_82840000	50	*iHS*	E9	chr2:82790000_82840000	Landuse	*ENSGALG00000049962 (LncRNA)*
chr2:82850000_82940000	90	*iHS*	E1, E9	chr2:82850000_82940000	Landuse	*ENSGALG00000049962 (LncRNA), ENSGALG00000051615 (LncRNA)*
chr4:75670000_75710000	40	*iHS*	E7	chr4:75690000_75810000	Landuse	
chr5:35470000_35670000	200	*iHS*	E7	chr5:35600000_35660000	Landuse	*NPAS3*
chr5:48990000_49080000	90	*ZHp*	E9	chr5:48990000_49080000	Landuse	*BEGAIN, ENSGALG00000051728 (LncRNA), ENSGALG00000053427 (LncRNA)*
chr5:49020000_49070000	50	*ZHp*	E3	chr5:48990000_49080000	Landuse	
chr7:31500000_31550000	50	*iHS*	E7	chr7:31520000_31550000	Landuse	
chr8:3260000_3340000	80	*iHS*	E8	chr8:3330000_3350000	Landuse	
chr28:4530000_4580000	50	*iHS*	E1	chr28:4520000_4570000	Landuse	*RAB8A, TPM4*
chr1:440000_500000	60	*iHS*	E5	chr1:430000_450000	minTemp	
chr1:440000_500000	60	*iHS*	E5	chr1:460000_510000	minTemp	
chr1:450000_510000	60	*iHS*	E7	chr1:460000_510000	minTemp	
chr1:109100000_109120000	20	*iHS*	E7	chr1:109000000_109120000	minTemp	
chr1:108900000_109030000	130	*iHS*	E7	chr1:109000000_109120000	minTemp	*ERG*
chr1:109680000_109820000	140	*iHS*	E7	chr1:109760000_109800000	minTemp	*DSCAM*
chr1:109680000_109820000	140	*iHS*	E7	chr1:109810000_109830000	minTemp	*DSCAM*

chr2:43530000_43550000	20	*iHS*	E7	chr2:43430000_43550000	minTemp	
chr2:43530000_43560000	30	*iHS*	E8	chr2:43430000_43550000	minTemp	
chr5:16940000_16960000	20	*iHS*	E8	chr5:16940000_16980000	minTemp	*FADS2*
chr5:28470000_35540000	70	*iHS*	E5	chr5:34720000_34740000	minTemp	*ACTN1, NPAS3*
chr6:36170000_36190000	20	*iHS*	E6	chr6:36170000_36220000	minTemp	*INPP5A*
chr10:18140000_18170000	30	*iHS*	E7	chr10:18140000_18180000	minTemp	*ANKDD1A, SPG21, CLPX*
chr18:3690000_3730000	40	*iHS*	E7	chr18:3690000_3770000	minTemp	*CD7*
chr18:3750000_3780000	30	*iHS*	E7	chr18:3690000_3770000	minTemp	*SEPTIN9*
chr21:3340000_3370000	30	*iHS*	E7, E8	chr21:3340000_3380000	minTemp	*SPSB1*
chr1:65400000_65420000	20	*iHS*	E8	chr1:65410000_65440000	precDQ	*BCAT1*
chr1:109680000_109820000	140	*iHS*	E7	chr1:109700000_109730000	precDQ	*DSCAM*
chr1:109680000_109820000	140	*iHS*	E7	chr1:109740000_109790000	precDQ	*DSCAM*
chr1:112350000_112400000	50	*iHS*	E8	chr1:112350000_112370000	precDQ	*FUNDC1, EFHC2*
chr2:28130000_28150000	20	*ZHp*	E8	chr2:28120000_28210000	precDQ	
chr2:145490000_145670000	180	*iHS*	E5	chr2:145640000_145670000	precDQ	*TRAPPC9, CHRAC1, gga-mir-6572, AGO2*
chr3:20090000_20120000	30	*ZHp*	E9	chr3:20090000_20110000	precDQ	*ESRRG*
chr3:20100000_20120000	20	*ZHp*	E1, E6, E12	chr3:20090000_20110000	precDQ	*ESRRG*
chr3:72070000_72140000	70	*iHS*	E1	chr3:71840000_72120000	precDQ	*gga-mir-12262*
chr3:72080000_72140000	60	*iHS*	E12	chr3:71840000_72120000	precDQ	
chr3:72070000_72140000	70	*iHS*	E1	chr3:72130000_72150000	precDQ	
chr3:72080000_72140000	60	*iHS*	E12	chr3:72130000_72150000	precDQ	
chr3:72250000_72280000	20	*iHS*	E1	chr3:72220000_72340000	precDQ	
chr3:72300000_72370000	70	*iHS*	E1	chr3:72220000_72340000	precDQ	
chr3:72380000_72540000	160	*iHS*	E1	chr3:72360000_72540000	precDQ	
chr3:72300000_72370000	70	*iHS*	E1	chr3:72360000_72540000	precDQ	
chr3:73170000_73220000	50	*iHS*	E1	chr3:73120000_73250000	precDQ	*MANEA*

chr3:73290000_73400000	110	*iHS*	E1	chr3:73300000_73390000	precDQ	
chr4:27840000_27880000	40	*iHS*	E5	chr4:27820000_27880000	precDQ	
chr4:28200000_28290000	90	*iHS*	E5	chr4:28210000_28230000	precDQ	
chr4:28200000_28290000	90	*iHS*	E5	chr4:28270000_28300000	precDQ	
chr4:28460000_28550000	90	*iHS*	E5	chr4:28490000_28520000	precDQ	
chr4:28490000_28520000	30	*iHS*	E6	chr4:28490000_28520000	precDQ	*ENSGALG00000048582 (LncRNA), ENSGALG00000054220 (LncRNA)*
chr4:74600000_74620000	20	*ZHp*	E9	chr4:74550000_74620000	precDQ	
chr4:89960000_89990000	30	*iHS*	E1	chr4:89950000_89980000	precDQ	*EXOC6B*
chr6:12520000_12540000	20	*iHS*	E5	chr6:12510000_12530000	precDQ	*CDH23*
chr9:12520000_12560000	40	*ZHp and/or iHS*	E1, E8, E9	chr9:12490000_12570000	precDQ	*AGTR1*
chr9:12530000_12550000	20	*ZHp*	E10	chr9:12490000_12570000	precDQ	*AGTR1*
chr17:8260000_8290000	30	*iHS*	E8	chr17:8280000_8300000	precDQ	*GPSM1*
chr17:9450000_9470000	20	*iHS*	E2	chr17:9440000_9520000	precDQ	
chr17:9490000_9550000	60	*iHS*	E2	chr17:9440000_9520000	precDQ	*LHX2*
chr17:9240000_9510000	270	*iHS*	E2	chr17:9440000_9520000	precDQ	*STRBP, DENND1A*
chr19:1810000_1830000	20	*ZHp*	E10	chr19:1810000_1850000	precDQ	*AUTS2*
chr19:1770000_1840000	70	*iHS*	E8	chr19:1810000_1850000	precDQ	*AUTS2*
chr20:1100000_1120000	20	*iHS*	E5	chr20:1090000_1110000	precDQ	*CPNE1*
chr23:4350000_4400000	50	*iHS*	E9	chr23:4350000_4400000	precDQ	*AGO1, AGO4, CLSPN, C1orf216*
chr1:440000_500000	60	*iHS*	E5	chr1:450000_490000	precSeasonality	
chr1:450000_510000	60	*iHS*	E7	chr1:450000_490000	precSeasonality	
chr1:185880000_185900000	20	*ZHp*	E2	chr1:185880000_185920000	precSeasonality	
chr2:28130000_28150000	20	*ZHp*	E8	chr2:28140000_28200000	precSeasonality	
chr2:104220000_104270000	50	*iHS*	E11	chr2:104260000_104320000	precSeasonality	*TAF4B, KCTD1*

chr2:146900000_146950000	50	*iHS*	E5	chr2:146890000_146920000	precSeasonality	*TSNARE1*
chr3:20090000_20120000	30	*ZHp*	E9	chr3:20080000_20110000	precSeasonality	*ESRRG*
chr3:20100000_20120000	20	*ZHp*	E1, E6, E12	chr3:20080000_20110000	precSeasonality	*ESRRG*
chr3:85880000_85900000	20	*iHS*	E6	chr3:85880000_85900000	precSeasonality	*ENSGALG00000047770 (miRNA)*
chr4:28200000_28290000	90	*iHS*	E5	chr4:28270000_28360000	precSeasonality	
chr4:75650000_75670000	20	*ZHp*	E2	chr4:75660000_75780000	precSeasonality	
chr4:75670000_75710000	40	*iHS*	E7	chr4:75660000_75780000	precSeasonality	
chr4:77520000_77550000	30	*ZHp and iHS*	E9, E10	chr4:77500000_77530000	precSeasonality	
chr4:89960000_89990000	30	*iHS*	E1	chr4:89960000_89980000	precSeasonality	*EXOC6B*
chr4:90280000_90420000	140	*iHS*	E9	chr4:90310000_90340000	precSeasonality	
chr4:90290000_90320000	30	*iHS*	E9	chr4:90310000_90340000	precSeasonality	
chr5:21740000_22800000	60	*iHS*	E3	chr5:21780000_21800000	precSeasonality	*MIR129-2*
chr5:28470000_35540000	70	*iHS*	E5	chr5:28910000_28970000	precSeasonality	*ACTN1, NPAS3*
chr5:35400000_35450000	50	*iHS*	E9	chr5:35390000_35410000	precSeasonality	*NPAS3*
chr5:28470000_35540000	70	*iHS*	E5	chr5:35390000_35410000	precSeasonality	*ACTN1, NPAS3*
chr5:40520000_40570000	50	*ZHp*	E9	chr5:40460000_40570000	precSeasonality	*NRXN3*
chr5:48790000_48820000	30	*iHS*	E9	chr5:48710000_49000000	precSeasonality	
chr5:48850000_48870000	20	*iHS*	E9	chr5:48710000_49000000	precSeasonality	*WDR25*
chr5:48850000_48880000	30	*iHS*	E9	chr5:48710000_49000000	precSeasonality	*BEGAIN*
chr5:48890000_48920000	30	*iHS*	E9	chr5:48710000_49000000	precSeasonality	*BEGAIN*
chr5:48700000_48820000	120	*iHS*	E9	chr5:48710000_49000000	precSeasonality	*YY2, SLC25A29, WARS, WDR25*
chr5:48990000_49080000	90	*ZHp*	E9	chr5:48710000_49000000	precSeasonality	*BEGAIN, ENSGALG00000051728 (LncRNA), ENSGALG00000053427 (LncRNA)*

chr5:49090000_49110000	20	*iHS*	E9	chr5:49070000_49190000	precSeasonality	
chr5:48990000_49080000	90	*ZHp*	E9	chr5:49070000_49190000	precSeasonality	*BEGAIN, ENSGALG00000051728 (LncRNA), ENSGALG00000053427 (LncRNA)*
chr5:49140000_49170000	30	*iHS*	E9	chr5:49070000_49190000	precSeasonality	*DLK1*
chr6:12520000_12540000	20	*iHS*	E5	chr6:12510000_12620000	precSeasonality	*CDH23*
chr9:12520000_12560000	40	*ZHp and/or iHS*	E1, E8, E9	chr9:12510000_12570000	precSeasonality	*AGTR1*
chr9:12530000_12550000	20	*ZHp*	E10	chr9:12510000_12570000	precSeasonality	*AGTR1*
chr17:9490000_9550000	60	*iHS*	E2	chr17:9480000_9510000	precSeasonality	
chr17:9240000_9510000	270	*iHS*	E2	chr17:9480000_9510000	precSeasonality	*STRBP, DENND1A*
chr19:1770000_1840000	70	*iHS*	E8	chr19:1750000_1790000	precSeasonality	*AUTS2*
chr20:1100000_1120000	20	*iHS*	E5	chr20:1090000_1110000	precSeasonality	*CPNE1*
chr23:4350000_4400000	50	*iHS*	E9	chr23:4350000_4400000	precSeasonality	*AGO1, AGO4, CLSPN, C1orf216*
chr25:210000_250000	40	*iHS*	E9	chr25:240000_260000	precSeasonality	*SETDB1, CERS2, MINDY1, PRUNE1, CDC42SE1, MLLT11, GABPB2, SEMA6C*
chr25:220000_340000	120	*iHS*	E2	chr25:240000_260000	precSeasonality	*PRUNE1, CDC42SE1, MLLT11, GABPB2, SEMA6C, SCNM1, TMOD4, VPS72, PIP5K1A, PSMD4, ZNF687, PI4KB, RFX5*
chr25:250000_270000	20	*iHS*	E1	chr25:240000_260000	precSeasonality	*SCNM1, TMOD4, VPS72, PIP5K1A*
chr28:4530000_4580000	50	*iHS*	E1	chr28:4540000_4580000	precSeasonality	*RAB8A, TPM4*
chr31:930000_950000	20	*iHS*	E2	chr31:920000_950000	precSeasonality	
chr1:98990000_99020000	20	*ZHp*	E10	chr1:99010000_99040000	precWQ	
chr2:65550000_65570000	20	*iHS*	E6	chr2:65560000_65580000	precWQ	*RPP40*

chr4:39410000_39430000	20	*iHS*	E1	chr4:39410000_39440000	precWQ	
chr4:89960000_89990000	30	*iHS*	E1	chr4:89960000_89990000	precWQ	*EXOC6B*
chr5:28470000_35540000	70	*iHS*	E5	chr5:29730000_29750000	precWQ	*ACTN1, NPAS3*
chr5:55480000_55510000	30	*iHS*	E4	chr5:55470000_55500000	precWQ	*ACTR10*
chr8:6450000_6500000	50	*iHS*	E12	chr8:6440000_6460000	precWQ	*TOR1AIP2, FAM163A*
chr10:13650000_13670000	20	*iHS*	E11	chr10:13630000_13730000	precWQ	
chr10:13710000_13770000	60	*iHS*	E11	chr10:13630000_13730000	precWQ	*NTRK3*
chr11:16300000_16320000	20	*iHS*	E5	chr11:16300000_16320000	precWQ	*CDH13*
chr15:10730000_10760000	30	*iHS*	E8	chr15:10710000_10750000	precWQ	
chr1:197480000_197600000	120	*iHS*	E7	chr1:197480000_197580000	soilOrgC	*TPP1, gga-mir-1600, LCMT2*
chr2:126400000_126480000	80	*iHS*	E10	chr2:126380000_126450000	soilOrgC	
chr2:137900000_137940000	40	*iHS*	E5	chr2:137920000_137940000	soilOrgC	*DERL1, TBC1D31*
chr2:147010000_147060000	50	*iHS*	E5	chr2:147030000_147090000	soilOrgC	*TSNARE1*
chr2:147080000_147120000	40	*ZHp*	E10	chr2:147030000_147090000	soilOrgC	*TSNARE1*
chr2:147080000_147150000	70	*iHS*	E9	chr2:147030000_147090000	soilOrgC	*TSNARE1*
chr2:147010000_147160000	150	*ZHp*	E5	chr2:147030000_147090000	soilOrgC	*TSNARE1*
chr2:147030000_147090000	60	*iHS*	E3	chr2:147030000_147090000	soilOrgC	*TSNARE1*
chr2:147050000_147200000	150	*ZHp*	E9	chr2:147030000_147090000	soilOrgC	*TSNARE1*
chr4:40940000_40960000	20	*iHS*	E6	chr4:40930000_40980000	soilOrgC	
chr4:75670000_75710000	40	*iHS*	E7	chr4:75670000_75710000	soilOrgC	
chr7:29240000_29310000	70	*iHS*	E4	chr7:29270000_29290000	soilOrgC	*DPP10*
chr7:29280000_29310000	30	*ZHp or iHS*	E1, E2	chr7:29270000_29290000	soilOrgC	*DPP10*
chr7:31500000_31550000	50	*iHS*	E7	chr7:31520000_31560000	soilOrgC	
chr10:13650000_13670000	20	*iHS*	E11	chr10:13660000_13690000	soilOrgC	*NTRK3*
chr33:6470000_6570000	100	*ZHp*	E1	chr33:6470000_6500000	soilOrgC	*DRAP1, RELA, TMEM223, STX5, WDR74, SNORD30, SNORD22, TAF6L, MTA2, EML3*

Table S5.8. Mean value for the six selected agro-climatic variables per ecotype. (Variables were used for ecotype definition and genomic-environmental association analysis).

Ecotype	minTemp Minimum temperature (C°) in the coldest month	precSeasonality Precipitation Seasonality (mm/m^2)	precWQbio16 Precipitation in the wettest quarter (mm/m^2)	precDQ Precipitation in the driest quarter (mm/m^2)	SoilOrgC Soil organic carbon content (g/Kg at depth 0.00 m)	LandUse (%)
E1	5.5	87.42	1050.17	79.97	130.11	32.85
E2	6.9	105.30	970.30	24.75	139.45	11.62
E3	9.8	71.42	428.32	37.89	109.05	12.18
E4	6.2	90.93	758.27	59.20	140.33	38.28
E5	8.7	50.26	505.37	116.42	71.89	25.02
E6	11.6	82.61	362.20	42.78	113.07	32.24
E7	1.8	105.75	770.10	46.10	127.40	37.53
E8	11.4	103.67	1089.28	11.56	123.50	1.33
E9	7.3	141.00	458.79	9.42	113.89	38.20
E10	10.5	136.90	480.70	19.60	111.00	29.19
E11	10.4	84.82	529.27	72.73	124.64	38.14
E12	11.2	121.60	767.50	11.30	105.70	34.22

Table S5.9. Identified outlier SNPs by applying RDA on Hp and iHS SNPs of candidate regions. (SD > 3 and p-value = 0.0027 two-tailed cut-off) (**axis:** significant axes (p-value ≤ 0.1 – 0.01). **loading:** SNP value in the ordination of PCA space. **minTemp:** minimum temperature in the coldest month (bio6). **precSeason:** precipitation seasonality (bio15). **precWQ:** precipitation in the wettest quarter (bio16). **precDQ:** precipitation in the driest quarter (bio17). **SoilOrgC:** soil organic carbon content. **Predictor:** most strongly correlated predictor with each candidate SNP. **Correlation:** an absolute value of the strongest correlation between SNP and predictor). Highlighted SNPs with r ≥ 0.28 – 0.3 were further investigated.

			Correlations of each candidate SNP with environmental predictors							
axis	**SNP position (bp)**	**loading**	**minTemp**	**precSeason**	**precWQ**	**precDQ**	**SoilOrgC**	**LandUse**	**predictor**	**correlation**
1	chr1_109700810	0.07	0.04	-0.18	-0.01	0.22	-0.23	-0.10	SoilOrgC	0.23
1	chr1_112390303	0.07	-0.01	0.20	0.18	0.02	-0.06	0.02	precSeason	0.20
1	chr1_118502383	0.07	0.08	-0.15	-0.20	0.20	-0.33	0.08	SoilOrgC	0.33
1	chr1_118539067	0.07	0.09	-0.13	-0.21	0.19	-0.33	0.08	SoilOrgC	0.33
1	chr1_132947694	0.07	0.07	-0.15	-0.29	0.17	-0.35	0.03	SoilOrgC	0.35
1	chr1_132948242	-0.07	0.03	0.15	0.08	-0.26	0.27	-0.06	SoilOrgC	0.27
1	chr1_132950294	-0.08	-0.07	0.24	0.32	-0.21	0.33	-0.02	SoilOrgC	0.33
1	chr1_132951411	0.08	0.08	-0.25	-0.24	0.23	-0.34	0.02	SoilOrgC	0.34
1	chr1_150103310	0.07	-0.02	-0.11	-0.11	0.15	-0.19	0.08	SoilOrgC	0.19
1	chr1_150128380	0.07	-0.02	-0.11	-0.11	0.15	-0.19	0.08	SoilOrgC	0.19
1	chr1_156616848	0.08	0.05	-0.11	-0.20	0.11	-0.25	0.07	SoilOrgC	0.25
1	chr1_176714724	-0.08	0.10	-0.14	-0.01	-0.05	0.03	-0.15	LandUse	0.15
1	chr1_176731485	0.08	-0.07	0.11	0.04	-0.01	0.05	0.09	precSeason	0.11
1	chr1_189278605	-0.07	0.03	-0.03	0.04	-0.06	0.10	-0.06	SoilOrgC	0.10
1	chr1_23065714	0.07	0.08	-0.18	-0.10	0.22	-0.30	-0.02	SoilOrgC	0.30
1	chr1_23140159	0.08	0.00	-0.16	-0.11	0.26	-0.31	0.01	SoilOrgC	0.31
1	chr1_23141490	0.07	0.12	-0.25	-0.26	0.25	-0.40	-0.05	SoilOrgC	0.40
1	chr1_23147646	0.07	0.09	-0.10	-0.11	0.14	-0.29	-0.08	SoilOrgC	0.29
1	chr1_23164664	0.08	0.07	-0.11	-0.19	0.19	-0.33	0.04	SoilOrgC	0.33
1	chr1_33675583	0.07	0.21	-0.14	-0.27	0.06	-0.29	-0.02	SoilOrgC	0.29
1	chr10_12770763	0.07	0.11	-0.06	-0.12	0.14	-0.19	0.08	SoilOrgC	0.19
1	chr10_12779492	0.08	0.17	-0.11	-0.21	0.14	-0.29	0.04	SoilOrgC	0.29
1	chr10_13015093	0.07	-0.07	-0.06	-0.08	0.14	-0.18	0.04	SoilOrgC	0.18
1	chr10_13073553	-0.08	-0.06	0.12	0.09	-0.15	0.25	0.03	SoilOrgC	0.25
1	chr10_13667613	0.08	0.07	-0.18	-0.12	0.16	-0.22	-0.02	SoilOrgC	0.22
1	chr10_14416135	-0.07	-0.15	-0.20	0.00	0.13	0.08	0.11	precSeason	0.20
1	chr10_17664883	-0.08	-0.10	-0.03	0.22	-0.04	0.21	-0.12	precWQ	0.22
1	chr10_5385384	-0.07	-0.05	0.13	0.07	-0.15	0.22	0.03	SoilOrgC	0.22

1	chr10_5396155	-0.09	-0.08	0.05	0.02	-0.11	0.22	0.02	SoilOrgC	0.22
1	chr10_7170539	0.08	0.07	-0.19	-0.10	0.22	-0.26	0.00	SoilOrgC	0.26
1	chr10_7190022	0.07	0.08	-0.17	-0.19	0.24	-0.33	0.09	SoilOrgC	0.33
1	chr11_1444895	-0.08	0.08	-0.22	-0.25	-0.07	0.04	-0.02	precWQ	0.25
1	chr14_2833882	0.07	0.22	-0.14	-0.18	0.11	-0.25	0.05	SoilOrgC	0.25
1	chr14_2844786	0.07	0.23	-0.26	-0.20	0.09	-0.23	-0.09	precSeason	0.26
1	chr14_4925611	-0.08	0.16	-0.16	-0.17	-0.10	0.15	0.03	precWQ	0.17
1	chr14_5684953	0.08	0.08	-0.16	-0.16	0.18	-0.26	0.00	SoilOrgC	0.26
1	chr14_6200065	0.09	0.10	-0.12	-0.13	0.16	-0.32	0.01	SoilOrgC	0.32
1	chr14_6200159	0.09	0.08	-0.14	-0.11	0.18	-0.32	-0.01	SoilOrgC	0.32
1	chr14_6200496	0.08	0.09	-0.16	-0.11	0.15	-0.30	-0.07	SoilOrgC	0.30
1	chr14_6201610	-0.09	-0.21	0.06	0.23	-0.08	0.36	0.01	SoilOrgC	0.36
1	chr14_6202614	0.08	0.18	-0.01	-0.13	0.07	-0.19	0.01	SoilOrgC	0.19
1	chr14_6204028	-0.08	-0.17	-0.09	0.13	-0.04	0.17	-0.14	SoilOrgC	0.17
1	chr14_6205574	-0.08	-0.13	-0.02	0.10	-0.06	0.30	0.01	SoilOrgC	0.30
1	chr14_6209248	0.08	0.12	-0.05	-0.09	0.11	-0.30	-0.04	SoilOrgC	0.30
1	chr14_6212235	0.07	0.00	0.06	-0.10	0.11	-0.12	0.14	LandUse	0.14
1	chr14_6905850	0.07	0.12	-0.08	-0.20	0.13	-0.20	0.03	SoilOrgC	0.20
1	chr14_7165714	-0.07	-0.08	-0.04	0.17	-0.06	0.26	-0.02	SoilOrgC	0.26
1	chr14_7170636	-0.08	-0.13	-0.09	0.18	-0.03	0.27	-0.03	SoilOrgC	0.27
1	chr17_5927565	0.07	0.07	-0.12	-0.11	0.18	-0.23	0.01	SoilOrgC	0.23
1	chr17_5931285	0.09	0.11	-0.21	-0.05	0.25	-0.24	-0.06	precDQ	0.25
1	chr17_6740515	0.07	0.07	-0.38	-0.13	0.36	-0.37	-0.07	precSeason	0.38
1	chr17_7847916	0.07	0.13	0.08	-0.21	0.03	-0.21	0.16	precWQ	0.21
1	chr17_8064416	-0.08	-0.10	0.05	0.21	0.00	0.22	-0.05	SoilOrgC	0.22
1	chr19_1794269	0.07	0.03	-0.18	-0.03	0.21	-0.27	0.02	SoilOrgC	0.27
1	chr19_1816989	0.08	-0.05	-0.26	-0.02	0.34	-0.29	0.03	precDQ	0.34
1	chr19_1819556	0.07	0.11	-0.26	-0.25	0.23	-0.37	0.05	SoilOrgC	0.37
1	chr19_1830033	0.07	0.06	-0.29	-0.16	0.27	-0.34	0.01	SoilOrgC	0.34
1	chr2_102660808	-0.07	0.00	-0.12	0.14	-0.05	0.15	-0.10	SoilOrgC	0.15
1	chr2_102820856	0.07	-0.02	0.00	0.06	0.11	-0.14	-0.10	SoilOrgC	0.14
1	chr2_107682455	-0.08	0.05	-0.14	0.11	-0.07	0.17	-0.19	LandUse	0.19
1	chr2_107689781	0.08	0.01	0.08	0.02	0.10	-0.19	0.04	SoilOrgC	0.19
1	chr2_108120063	-0.07	0.03	-0.15	-0.09	0.05	-0.01	0.04	precSeason	0.15
1	chr2_108120625	-0.08	-0.03	-0.04	0.02	0.03	0.13	0.08	SoilOrgC	0.13
1	chr2_126412659	0.07	0.06	-0.22	-0.17	0.21	-0.27	0.03	SoilOrgC	0.27

1	chr2_126471892	0.07	0.09	-0.16	-0.15	0.22	-0.30	0.06	SoilOrgC	0.30
1	chr2_134301101	0.07	0.07	0.01	-0.23	0.10	-0.29	0.08	SoilOrgC	0.29
1	chr2_134301318	0.07	0.07	0.01	-0.23	0.10	-0.29	0.08	SoilOrgC	0.29
1	chr2_134301357	0.07	0.07	0.01	-0.23	0.10	-0.29	0.08	SoilOrgC	0.29
1	chr2_134327192	0.07	0.07	0.01	-0.23	0.10	-0.29	0.08	SoilOrgC	0.29
1	chr2_137552718	0.07	0.03	0.01	-0.17	0.09	-0.18	0.04	SoilOrgC	0.18
1	chr2_140301053	-0.07	-0.11	0.13	0.37	-0.18	0.30	-0.13	precWQ	0.37
1	chr2_140313253	-0.07	-0.11	0.11	0.36	-0.17	0.30	-0.12	precWQ	0.36
1	chr2_140367628	-0.07	-0.06	0.27	0.30	-0.19	0.26	-0.04	precWQ	0.30
1	chr2_140379965	0.07	0.12	-0.04	-0.31	0.06	-0.25	0.03	precWQ	0.31
1	chr2_140386472	-0.07	-0.07	0.09	0.27	-0.06	0.23	0.06	precWQ	0.27
1	chr2_88521218	0.07	-0.05	-0.05	0.02	0.07	0.04	0.12	LandUse	0.12
1	chr2_88533917	-0.07	-0.02	0.06	0.03	-0.12	0.06	-0.09	precDQ	0.12
1	chr2_88539680	0.08	-0.04	-0.08	-0.01	0.09	0.03	0.11	LandUse	0.11
1	chr2_91150684	-0.09	-0.02	0.03	-0.05	-0.09	0.05	0.01	precDQ	0.09
1	chr2_91151231	-0.08	-0.08	0.01	-0.04	-0.11	0.05	-0.02	precDQ	0.11
1	chr2_91177588	-0.07	0.00	-0.01	-0.09	-0.07	0.04	0.02	precWQ	0.09
1	chr2_91303700	-0.07	-0.01	-0.03	0.04	-0.06	0.07	-0.02	SoilOrgC	0.07
1	chr2_91326459	-0.07	0.08	0.14	-0.06	-0.18	0.02	0.02	precDQ	0.18
1	chr2_91337155	0.08	0.04	-0.05	-0.20	0.07	-0.20	0.12	SoilOrgC	0.20
1	chr2_91347863	-0.07	-0.01	0.07	0.05	-0.11	0.10	-0.09	precDQ	0.11
1	chr2_91354794	-0.07	0.01	0.18	0.17	-0.18	0.18	-0.12	precSeason	0.18
1	chr2_91738133	0.07	0.11	-0.12	-0.27	0.06	-0.19	0.13	precWQ	0.27
1	chr2_99486336	0.07	0.14	-0.10	-0.11	0.11	-0.20	0.03	SoilOrgC	0.20
1	chr20_4173643	0.07	0.13	-0.18	-0.31	0.14	-0.29	0.03	precWQ	0.31
1	chr20_4176065	0.08	0.19	-0.15	-0.28	0.07	-0.28	-0.03	precWQ	0.28
1	chr20_4183810	0.07	0.12	-0.39	-0.25	0.28	-0.39	-0.05	precSeason	0.39
1	chr21_6574007	0.07	0.10	-0.23	-0.15	0.15	-0.15	-0.03	precSeason	0.23
1	chr3_12494125	-0.08	-0.10	-0.01	0.13	-0.13	0.26	-0.02	SoilOrgC	0.26
1	chr3_16645865	0.07	0.17	0.21	-0.14	-0.08	-0.12	0.11	precSeason	0.21
1	chr3_16891572	0.07	0.09	-0.34	-0.10	0.32	-0.41	-0.12	SoilOrgC	0.41
1	chr3_38560061	0.08	0.12	-0.21	-0.24	0.18	-0.31	-0.06	SoilOrgC	0.31
1	chr3_50580677	-0.07	-0.04	0.00	0.15	-0.09	0.09	-0.21	LandUse	0.21
1	chr3_61410429	0.07	0.06	-0.03	-0.07	0.10	-0.18	0.00	SoilOrgC	0.18
1	chr3_61414185	0.08	0.09	-0.04	-0.15	0.08	-0.19	0.02	SoilOrgC	0.19
1	chr3_61443805	0.07	0.12	-0.07	-0.09	0.03	-0.17	-0.08	SoilOrgC	0.17

1	chr3_61459215	0.07	0.17	0.04	-0.25	-0.09	-0.17	0.07	precWQ	0.25
1	chr3_77671096	-0.07	-0.17	-0.09	0.23	0.02	0.24	-0.10	SoilOrgC	0.24
1	chr3_77680432	-0.07	-0.17	-0.12	0.26	0.13	0.12	-0.15	precWQ	0.26
1	chr4_23241575	0.07	0.12	-0.01	-0.17	0.09	-0.25	0.02	SoilOrgC	0.25
1	chr4_28679923	0.07	0.03	-0.07	-0.16	0.07	-0.07	0.14	precWQ	0.16
1	chr4_3600618	-0.07	-0.12	-0.08	-0.02	-0.02	0.10	0.01	minTemp	0.12
1	chr4_3627776	0.07	0.15	0.06	-0.33	-0.02	-0.21	0.13	precWQ	0.33
1	chr4_51836212	-0.07	-0.19	-0.02	0.13	0.02	0.17	0.10	minTemp	0.19
1	chr4_51936281	0.07	0.11	0.04	-0.25	0.02	-0.21	0.09	precWQ	0.25
1	chr4_88511966	-0.07	-0.07	-0.04	-0.05	-0.11	0.14	-0.02	SoilOrgC	0.14
1	chr4_88536311	-0.07	-0.28	0.10	0.28	0.01	0.23	-0.03	precWQ	0.28
1	chr4_89983070	0.08	0.16	-0.13	-0.25	0.07	-0.23	0.01	precWQ	0.25
1	chr4_90097954	0.08	0.14	-0.15	-0.17	0.09	-0.28	-0.06	SoilOrgC	0.28
1	chr5_12949778	0.07	0.03	-0.12	-0.05	0.20	-0.21	-0.02	SoilOrgC	0.21
1	chr5_19564424	0.08	0.01	-0.11	-0.14	0.19	-0.33	-0.03	SoilOrgC	0.33
1	chr5_28320494	-0.07	-0.17	0.21	0.15	-0.15	0.28	0.10	SoilOrgC	0.28
1	chr5_35557653	0.07	0.13	-0.29	-0.13	0.23	-0.36	-0.07	SoilOrgC	0.36
1	chr5_39594296	0.07	0.08	-0.09	-0.27	0.14	-0.33	0.03	SoilOrgC	0.33
1	chr5_48729625	0.10	0.22	-0.28	-0.22	0.20	-0.30	-0.06	SoilOrgC	0.30
1	chr5_49525047	0.07	0.34	-0.27	-0.34	0.05	-0.32	-0.08	minTemp	0.34
1	chr5_49530367	0.09	0.24	-0.16	-0.25	0.12	-0.30	0.04	SoilOrgC	0.30
1	chr5_49544730	-0.07	-0.26	0.25	0.27	-0.05	0.27	0.03	precWQ	0.27
1	chr5_8266920	-0.07	-0.08	0.01	0.07	0.02	0.01	0.00	minTemp	0.08
1	chr6_28550216	0.07	0.18	-0.19	-0.22	0.10	-0.31	-0.06	SoilOrgC	0.31
1	chr7_30965068	0.08	0.12	-0.27	-0.11	0.25	-0.38	-0.13	SoilOrgC	0.38
1	chr7_30967748	0.07	0.14	-0.21	-0.27	0.15	-0.32	0.00	SoilOrgC	0.32
1	chr7_32597927	0.07	0.00	-0.13	-0.08	0.24	-0.26	0.01	SoilOrgC	0.26
1	chr7_32604693	0.07	0.10	-0.16	-0.13	0.19	-0.26	0.03	SoilOrgC	0.26
1	chr7_34948717	0.08	0.03	-0.19	-0.10	0.26	-0.37	-0.08	SoilOrgC	0.37
1	chr8_6198062	-0.08	-0.11	0.30	0.20	-0.25	0.38	0.07	SoilOrgC	0.38
1	chr9_17890042	0.07	0.08	-0.03	-0.04	0.11	-0.26	0.01	SoilOrgC	0.26
1	chr9_17953063	-0.08	-0.06	0.17	0.19	-0.19	0.29	-0.02	SoilOrgC	0.29
1	chr9_17960192	-0.08	-0.06	0.14	0.24	-0.18	0.30	-0.12	SoilOrgC	0.30
1	chr9_17960492	-0.07	-0.03	0.03	0.02	-0.21	0.31	-0.06	SoilOrgC	0.31
2	chr1_110212779	0.07	0.15	0.12	-0.09	-0.22	0.01	0.01	precDQ	0.22
2	chr1_112416137	0.07	0.11	0.17	0.07	-0.20	-0.06	-0.19	precDQ	0.20

2	chr1_118390274	-0.08	-0.07	-0.28	-0.01	0.30	-0.14	0.08	precDQ	0.30
2	chr1_118390636	-0.07	-0.05	-0.21	-0.05	0.25	-0.17	0.11	precDQ	0.25
2	chr1_118392895	-0.07	-0.05	-0.21	-0.05	0.25	-0.17	0.11	precDQ	0.25
2	chr1_118396566	-0.07	-0.05	-0.22	-0.05	0.26	-0.17	0.10	precDQ	0.26
2	chr1_118408182	-0.08	-0.06	-0.28	-0.03	0.29	-0.12	0.09	precDQ	0.29
2	chr1_128340105	-0.07	-0.24	0.02	0.05	0.10	0.07	0.12	minTemp	0.24
2	chr1_128341996	0.07	0.20	0.05	0.03	-0.16	0.02	-0.16	minTemp	0.20
2	chr1_128342571	-0.07	-0.24	0.02	0.05	0.10	0.07	0.12	minTemp	0.24
2	chr1_128461033	-0.07	-0.20	0.03	0.06	0.13	0.10	0.16	minTemp	0.20
2	chr1_128518467	-0.07	-0.20	0.03	0.06	0.13	0.10	0.16	minTemp	0.20
2	chr1_128548372	-0.07	-0.24	0.03	0.06	0.13	0.12	0.18	minTemp	0.24
2	chr1_128693862	-0.07	-0.23	-0.07	0.09	0.13	0.05	0.06	minTemp	0.23
2	chr1_132953316	0.07	0.20	0.12	-0.05	-0.25	0.07	-0.06	precDQ	0.25
2	chr1_132953669	0.07	0.31	0.01	-0.24	-0.17	-0.15	-0.14	minTemp	0.31
2	chr1_132954120	0.07	0.13	0.04	-0.08	-0.17	0.00	-0.13	precDQ	0.17
2	chr1_132954906	0.08	0.22	0.15	-0.17	-0.26	-0.04	-0.12	precDQ	0.26
2	chr1_143186941	-0.07	-0.06	0.02	0.05	0.19	-0.01	0.15	precDQ	0.19
2	chr1_192211061	0.07	0.09	0.17	0.02	-0.20	0.03	-0.10	precDQ	0.20
2	chr1_197483358	-0.07	-0.11	-0.20	0.10	0.37	-0.13	0.05	precDQ	0.37
2	chr1_98060042	-0.07	-0.20	0.03	0.13	0.13	0.06	0.13	minTemp	0.20
2	chr1_98060182	-0.07	-0.20	0.03	0.13	0.13	0.06	0.13	minTemp	0.20
2	chr1_98810822	-0.07	-0.20	0.01	0.12	0.13	0.07	0.04	minTemp	0.20
2	chr1_98810963	-0.07	-0.20	0.01	0.12	0.13	0.07	0.04	minTemp	0.20
2	chr1_98834610	-0.08	-0.21	0.03	0.14	0.14	0.07	0.07	minTemp	0.21
2	chr1_98880031	-0.08	-0.22	0.05	0.16	0.13	0.09	0.06	minTemp	0.22
2	chr1_98883766	-0.08	-0.22	0.05	0.16	0.13	0.09	0.06	minTemp	0.22
2	chr1_98909806	-0.09	-0.23	0.04	0.20	0.15	0.11	0.07	minTemp	0.23
2	chr1_98920429	-0.07	-0.16	0.04	0.19	0.11	0.09	0.03	precWQ	0.19
2	chr1_98920797	-0.09	-0.23	0.04	0.20	0.15	0.11	0.07	minTemp	0.23
2	chr1_98921867	-0.09	-0.23	0.04	0.20	0.15	0.11	0.07	minTemp	0.23
2	chr1_98924646	-0.08	-0.23	0.02	0.19	0.16	0.11	0.07	minTemp	0.23
2	chr1_98955453	-0.09	-0.23	0.01	0.20	0.17	0.11	0.07	minTemp	0.23
2	chr1_98990920	-0.07	-0.19	0.07	0.15	0.08	0.07	0.06	minTemp	0.19
2	chr1_98991028	-0.08	-0.21	0.03	0.20	0.12	0.11	0.06	minTemp	0.21
2	chr1_98991409	-0.08	-0.21	0.01	0.18	0.15	0.11	0.07	minTemp	0.21
2	chr1_98998007	-0.08	-0.19	0.00	0.15	0.13	0.12	0.08	minTemp	0.19

2	chr1_99003736	-0.08	-0.21	0.03	0.19	0.13	0.11	0.05	minTemp	0.21
2	chr1_99015095	-0.07	-0.14	-0.03	0.14	0.12	0.10	0.03	precWQ	0.14
2	chr10_13599109	-0.07	-0.30	0.19	0.35	0.02	0.21	0.07	precWQ	0.35
2	chr10_4126468	-0.08	-0.10	-0.16	0.10	0.20	0.05	0.09	precDQ	0.20
2	chr10_4188149	0.07	0.05	0.09	-0.12	-0.17	-0.06	-0.14	precDQ	0.17
2	chr10_5317612	-0.07	-0.19	-0.15	0.14	0.19	0.11	0.05	minTemp	0.19
2	chr10_6328263	-0.07	-0.08	-0.21	-0.02	0.25	-0.11	0.09	precDQ	0.25
2	chr10_9680312	0.08	0.21	0.09	-0.10	-0.15	-0.16	0.00	minTemp	0.21
2	chr10_9680456	0.08	0.21	0.06	-0.10	-0.14	-0.15	-0.04	minTemp	0.21
2	chr11_10332226	0.07	0.19	0.07	-0.06	-0.22	-0.07	-0.17	precDQ	0.22
2	chr12_11794158	-0.07	-0.14	-0.12	0.15	0.21	0.09	0.01	precDQ	0.21
2	chr14_2823172	-0.07	-0.04	-0.02	0.07	0.20	-0.13	0.05	precDQ	0.20
2	chr14_2834643	0.07	0.36	0.04	-0.23	-0.15	-0.12	-0.08	minTemp	0.36
2	chr14_2839546	0.08	0.33	-0.13	-0.17	-0.11	-0.20	-0.24	minTemp	0.33
2	chr14_2841689	0.07	0.24	0.16	0.03	-0.24	0.00	-0.20	minTemp	0.24
2	chr17_4716964	-0.07	-0.17	0.04	0.04	0.12	0.01	0.12	minTemp	0.17
2	chr17_5927857	-0.08	-0.15	-0.14	0.17	0.23	0.07	0.03	precDQ	0.23
2	chr17_5931456	-0.07	-0.12	-0.23	0.05	0.29	-0.06	0.04	precDQ	0.29
2	chr17_6198259	-0.07	-0.07	-0.09	0.03	0.16	0.01	0.08	precDQ	0.16
2	chr17_6647388	0.07	0.22	0.01	-0.03	-0.14	-0.17	-0.12	minTemp	0.22
2	chr17_8086055	0.06	0.15	0.05	-0.08	-0.16	-0.01	-0.02	precDQ	0.16
2	chr17_9532678	0.07	0.24	-0.04	-0.28	-0.13	-0.22	-0.08	precWQ	0.28
2	chr19_2420166	0.07	0.07	0.13	-0.08	-0.25	-0.03	-0.05	precDQ	0.25
2	chr19_2435462	-0.07	-0.04	-0.14	-0.04	0.25	0.00	0.16	precDQ	0.25
2	chr19_2449547	-0.07	-0.13	-0.04	0.04	0.18	0.04	0.13	precDQ	0.18
2	chr19_2451005	-0.07	-0.12	-0.14	0.08	0.26	-0.03	0.04	precDQ	0.26
2	chr19_2539524	-0.07	-0.09	-0.11	0.05	0.29	-0.11	0.04	precDQ	0.29
2	chr19_2549553	-0.08	-0.13	-0.04	0.05	0.22	0.03	0.12	precDQ	0.22
2	chr19_2550056	0.07	0.05	-0.01	0.05	-0.12	-0.04	-0.17	LandUse	0.17
2	chr19_901472	-0.07	-0.06	-0.04	0.08	0.15	0.07	0.11	precDQ	0.15
2	chr2_104258829	-0.07	-0.16	-0.18	0.06	0.26	0.07	0.16	precDQ	0.26
2	chr2_104264716	0.08	0.18	0.15	-0.02	-0.28	-0.05	-0.13	precDQ	0.28
2	chr2_104264975	0.07	0.24	0.12	-0.28	-0.28	-0.10	0.01	precWQ	0.28
2	chr2_104268593	0.07	0.11	0.20	0.09	-0.22	0.01	-0.10	precDQ	0.22
2	chr2_109398511	0.07	0.18	0.10	-0.22	-0.18	-0.08	0.00	precWQ	0.22
2	chr2_110320431	-0.07	-0.14	-0.02	0.21	0.18	0.10	-0.06	precWQ	0.21

2	chr2_117742214	-0.08	-0.18	-0.03	0.23	0.19	0.07	0.06	precWQ	0.23
2	chr2_117749087	-0.07	-0.15	-0.05	0.42	0.15	0.13	-0.14	precWQ	0.42
2	chr2_117752300	-0.07	-0.14	-0.08	0.25	0.22	0.00	-0.01	precWQ	0.25
2	chr2_118863529	-0.07	-0.11	0.03	0.15	0.13	0.10	0.10	precWQ	0.15
2	chr2_123115169	0.07	0.15	0.06	-0.08	-0.19	-0.01	-0.11	precDQ	0.19
2	chr2_123116707	0.08	0.16	0.10	-0.08	-0.27	0.08	-0.11	precDQ	0.27
2	chr2_137593028	0.07	0.16	0.02	-0.08	-0.13	-0.06	-0.08	minTemp	0.16
2	chr2_137611144	0.07	0.22	0.07	-0.22	-0.10	-0.10	-0.01	precWQ	0.22
2	chr2_137622713	0.07	0.26	-0.07	-0.23	-0.02	-0.22	-0.05	minTemp	0.26
2	chr2_137624985	0.07	0.27	0.07	-0.21	-0.15	-0.07	-0.06	minTemp	0.27
2	chr2_138341731	-0.07	-0.14	-0.12	0.04	0.15	0.00	0.05	precDQ	0.15
2	chr2_55326637	-0.07	-0.09	-0.12	0.19	0.21	0.03	-0.03	precDQ	0.21
2	chr2_55460931	-0.07	-0.17	-0.22	0.23	0.27	0.09	0.07	precDQ	0.27
2	chr2_55475568	-0.07	-0.18	-0.17	0.22	0.24	0.09	0.10	precDQ	0.24
2	chr2_78280118	0.07	0.28	-0.10	-0.17	-0.11	-0.04	-0.10	minTemp	0.28
2	chr2_78280162	0.09	0.31	-0.01	-0.24	-0.17	-0.06	-0.07	minTemp	0.31
2	chr2_78280175	0.09	0.31	-0.01	-0.24	-0.17	-0.06	-0.07	minTemp	0.31
2	chr2_78280397	-0.08	-0.27	0.00	0.17	0.16	0.04	0.07	minTemp	0.27
2	chr2_78280431	-0.08	-0.27	0.00	0.17	0.16	0.04	0.07	minTemp	0.27
2	chr2_78307383	0.08	0.29	-0.04	-0.16	-0.12	-0.08	-0.13	minTemp	0.29
2	chr2_82801924	0.07	0.26	-0.03	-0.01	-0.11	-0.01	-0.17	minTemp	0.26
2	chr2_82808194	0.07	0.26	-0.03	-0.01	-0.11	-0.01	-0.17	minTemp	0.26
2	chr2_82851573	0.07	0.25	-0.03	0.02	-0.09	0.01	-0.16	minTemp	0.25
2	chr2_82874119	0.07	0.26	-0.02	0.02	-0.09	0.01	-0.16	minTemp	0.26
2	chr2_82913226	0.07	0.28	0.03	0.05	-0.16	0.03	-0.18	minTemp	0.28
2	chr2_99407858	-0.06	-0.26	-0.04	0.16	0.20	0.05	0.09	minTemp	0.26
2	chr20_1187011	-0.07	-0.12	-0.08	0.01	0.29	-0.20	0.05	precDQ	0.29
2	chr20_2199652	-0.07	-0.20	-0.15	0.06	0.27	0.05	0.09	precDQ	0.27
2	chr20_4177646	0.07	0.16	0.01	-0.11	-0.16	-0.08	-0.10	precDQ	0.16
2	chr21_4144568	-0.07	-0.05	-0.18	-0.05	0.18	-0.03	0.10	precDQ	0.18
2	chr21_4353990	0.06	0.17	0.00	-0.31	-0.10	-0.19	0.05	precWQ	0.31
2	chr23_4357346	-0.07	-0.14	-0.19	-0.01	0.21	-0.03	0.07	precDQ	0.21
2	chr25_289654	0.07	0.18	-0.22	-0.16	0.02	-0.15	-0.24	LandUse	0.24
2	chr3_11966134	0.07	0.18	0.17	0.08	-0.19	0.01	-0.16	precDQ	0.19
2	chr3_12020211	0.07	0.14	0.13	0.12	-0.21	0.06	-0.23	LandUse	0.23
2	chr3_12020253	0.07	0.14	0.13	0.12	-0.21	0.06	-0.23	LandUse	0.23

2	chr3_5039795	-0.09	-0.24	-0.07	0.15	0.24	0.10	0.15	precDQ	0.24
2	chr4_77541544	0.07	0.17	-0.19	-0.08	-0.10	-0.02	-0.24	LandUse	0.24
2	chr4_82439583	-0.07	-0.19	0.02	0.31	0.12	0.15	0.00	precWQ	0.31
2	chr4_82870734	-0.07	-0.21	-0.06	0.28	0.19	0.17	0.05	precWQ	0.28
2	chr4_83335619	-0.06	-0.17	-0.16	0.13	0.27	0.01	0.03	precDQ	0.27
2	chr4_88536330	0.07	0.31	0.02	-0.28	-0.23	-0.08	-0.01	minTemp	0.31
2	chr5_19062036	-0.07	-0.29	-0.02	0.16	0.17	0.09	0.15	minTemp	0.29
2	chr5_19568630	0.06	0.14	-0.04	-0.21	-0.12	-0.05	-0.11	precWQ	0.21
2	chr5_27549756	-0.08	-0.15	0.04	0.07	0.14	0.07	0.20	LandUse	0.20
2	chr5_32397932	-0.07	-0.23	0.15	0.13	0.08	0.04	0.19	minTemp	0.23
2	chr5_36239584	0.07	0.25	0.05	-0.11	-0.24	0.11	-0.07	minTemp	0.25
2	chr5_36359777	-0.07	-0.19	-0.05	0.09	0.19	0.03	0.10	minTemp	0.19
2	chr5_38101243	0.07	0.14	0.08	-0.17	-0.19	-0.06	-0.09	precDQ	0.19
2	chr6_10251021	0.07	0.11	0.24	-0.02	-0.25	-0.03	-0.05	precDQ	0.25
2	chr7_19860503	0.06	0.15	-0.09	-0.13	-0.09	-0.17	-0.15	SoilOrgC	0.17
2	chr7_20954316	-0.07	-0.12	0.10	0.20	0.08	0.15	0.13	precWQ	0.20
2	chr8_23623147	0.07	0.17	-0.07	-0.10	-0.05	-0.15	-0.11	minTemp	0.17
2	chr8_24798255	-0.07	-0.19	-0.17	0.21	0.30	0.07	0.03	precDQ	0.30
2	chr9_21716194	-0.07	-0.20	-0.12	0.11	0.24	0.09	0.10	precDQ	0.24
2	chr9_21717354	-0.07	-0.12	-0.05	0.11	0.15	0.19	0.16	SoilOrgC	0.19
3	chr1_121966647	0.06	0.04	0.11	-0.05	0.03	-0.14	0.03	SoilOrgC	0.14
3	chr1_128799314	0.06	-0.06	0.04	-0.06	0.09	-0.14	0.11	SoilOrgC	0.14
3	chr1_132955026	0.07	0.05	0.20	-0.04	-0.11	-0.02	0.03	precSeason	0.20
3	chr1_173272540	0.07	-0.10	0.03	0.02	-0.01	0.03	0.05	minTemp	0.10
3	chr1_45641988	-0.06	0.16	0.03	0.09	-0.17	0.27	-0.13	SoilOrgC	0.27
3	chr1_45642449	-0.07	0.08	-0.12	-0.13	0.04	-0.01	0.06	precWQ	0.13
3	chr1_57263184	-0.07	0.07	-0.06	-0.06	0.03	-0.01	-0.01	minTemp	0.07
3	chr1_57264767	-0.06	0.14	0.01	-0.23	-0.04	-0.10	-0.05	precWQ	0.23
3	chr10_11885633	0.07	-0.18	0.05	0.23	-0.02	0.13	-0.05	precWQ	0.23
3	chr10_11908930	-0.07	-0.06	-0.10	0.07	0.04	0.06	-0.11	LandUse	0.11
3	chr10_13361869	0.06	-0.14	0.26	0.07	-0.07	0.08	0.13	precSeason	0.26
3	chr10_4270778	-0.07	0.11	-0.07	-0.19	0.02	-0.13	-0.05	precWQ	0.19
3	chr10_4288252	-0.07	0.06	-0.02	-0.15	0.03	-0.10	0.08	precWQ	0.15
3	chr10_4508583	0.06	-0.06	0.12	0.08	-0.05	0.03	0.02	precSeason	0.12
3	chr10_4523065	0.06	-0.16	0.06	0.17	-0.05	0.14	0.01	precWQ	0.17
3	chr10_4523408	0.07	-0.02	0.06	-0.10	-0.12	-0.02	0.09	precDQ	0.12

3	chr10_4568036	-0.07	0.08	-0.11	-0.06	-0.06	0.10	-0.06	precSeason	0.11
3	chr10_8276922	-0.06	0.04	-0.26	-0.05	0.11	0.04	-0.04	precSeason	0.26
3	chr10_8281861	-0.07	0.00	-0.11	-0.06	0.05	0.03	-0.02	precSeason	0.11
3	chr10_9402417	0.07	-0.12	0.30	0.19	-0.12	0.10	0.02	precSeason	0.30
3	chr10_9409961	0.07	0.08	0.10	-0.13	-0.07	-0.13	0.07	precWQ	0.13
3	chr11_16972959	0.07	-0.05	0.24	0.33	-0.09	0.07	-0.13	precWQ	0.33
3	chr13_17335168	-0.07	-0.03	-0.10	0.05	0.06	0.08	-0.10	LandUse	0.10
3	chr14_4548200	0.06	-0.06	0.05	-0.04	0.06	-0.16	0.08	SoilOrgC	0.16
3	chr15_4120266	-0.07	0.09	-0.22	-0.18	0.09	0.00	0.00	precSeason	0.22
3	chr17_7239162	-0.07	0.07	-0.24	-0.01	0.05	-0.14	-0.25	LandUse	0.25
3	chr17_7364636	-0.07	0.04	-0.10	-0.02	0.12	0.01	0.07	precDQ	0.12
3	chr17_7367928	-0.07	0.10	-0.09	-0.15	0.06	-0.05	0.03	precWQ	0.15
3	chr17_7631736	0.07	0.05	0.13	0.08	-0.02	-0.02	0.06	precSeason	0.13
3	chr17_7633071	0.07	-0.08	0.04	0.10	0.12	-0.04	0.06	precDQ	0.12
3	chr17_7635727	0.08	0.00	0.06	0.03	0.09	-0.10	0.01	SoilOrgC	0.10
3	chr17_7637261	-0.07	-0.08	-0.04	0.07	-0.01	0.08	-0.06	SoilOrgC	0.08
3	chr17_8288451	-0.07	0.11	-0.08	-0.11	0.03	-0.06	-0.03	minTemp	0.11
3	chr19_2433049	-0.06	0.11	0.11	-0.04	-0.07	0.04	0.05	minTemp	0.11
3	chr19_3307010	0.06	-0.07	0.18	0.08	-0.06	0.08	0.08	precSeason	0.18
3	chr19_922652	0.06	-0.03	0.00	0.05	-0.01	-0.05	0.03	precWQ	0.05
3	chr19_924613	-0.07	0.04	-0.08	-0.04	0.00	0.05	-0.09	LandUse	0.09
3	chr2_102695913	-0.07	0.12	0.01	-0.12	-0.11	0.05	-0.08	minTemp	0.12
3	chr2_110243241	-0.06	-0.09	0.09	0.11	0.01	0.12	0.06	SoilOrgC	0.12
3	chr2_112808674	0.08	-0.01	0.19	0.04	-0.04	-0.01	0.14	precSeason	0.19
3	chr2_112903173	0.07	0.06	0.14	-0.06	-0.03	-0.09	0.12	precSeason	0.14
3	chr2_114706004	-0.06	-0.03	-0.22	-0.09	0.10	0.00	-0.07	precSeason	0.22
3	chr2_134309595	0.07	0.08	0.13	-0.16	-0.03	-0.20	0.00	SoilOrgC	0.20
3	chr2_134317069	0.07	0.06	0.14	-0.19	-0.01	-0.25	0.06	SoilOrgC	0.25
3	chr2_140323648	-0.07	0.03	-0.07	-0.21	0.09	-0.19	0.02	precWQ	0.21
3	chr2_140335533	-0.07	0.05	-0.06	-0.25	0.07	-0.19	0.00	precWQ	0.25
3	chr20_4184104	-0.07	0.10	0.00	-0.22	-0.05	-0.01	0.01	precWQ	0.22
3	chr25_285695	-0.07	0.12	-0.26	-0.22	0.05	-0.17	-0.24	precSeason	0.26
3	chr25_322674	0.06	-0.22	0.04	0.15	0.05	0.10	0.08	minTemp	0.22
3	chr3_100431237	0.07	-0.05	0.10	0.11	-0.09	0.02	0.07	precWQ	0.11
3	chr3_100435235	0.06	0.05	0.11	0.09	-0.07	-0.05	0.11	precSeason	0.11
3	chr3_16289174	0.07	0.17	0.12	-0.17	-0.12	-0.23	0.05	SoilOrgC	0.23

3	chr3_16289248	0.06	0.10	-0.05	-0.15	0.04	-0.22	0.05	SoilOrgC	0.22
3	chr3_31596452	-0.07	-0.01	-0.05	0.05	0.00	0.05	-0.14	LandUse	0.14
3	chr3_31608169	0.07	-0.14	0.09	0.07	0.06	0.05	0.18	LandUse	0.18
3	chr3_31617477	-0.07	0.25	-0.20	-0.19	-0.09	-0.03	-0.14	minTemp	0.25
3	chr3_32208733	0.06	-0.14	0.22	0.14	-0.11	0.09	0.04	precSeason	0.22
3	chr3_41555894	-0.07	0.06	-0.03	-0.05	0.05	0.04	0.02	minTemp	0.06
3	chr3_5024316	-0.07	0.10	-0.03	-0.21	0.01	-0.05	0.11	precWQ	0.21
3	chr3_69197178	-0.07	0.04	-0.25	-0.14	0.06	-0.09	-0.21	precSeason	0.25
3	chr3_72111929	-0.07	0.14	-0.09	-0.12	0.04	-0.11	-0.04	minTemp	0.14
3	chr3_72131787	-0.07	0.15	0.08	-0.17	0.04	-0.09	0.10	precWQ	0.17
3	chr3_72275989	-0.06	0.05	0.01	-0.06	-0.01	0.02	0.00	precWQ	0.06
3	chr3_85440453	0.07	0.07	0.04	-0.11	-0.07	-0.08	0.06	precWQ	0.11
3	chr3_85441224	0.07	0.07	0.04	-0.11	-0.07	-0.08	0.06	precWQ	0.11
3	chr3_85441441	0.07	0.07	0.04	-0.11	-0.07	-0.08	0.06	precWQ	0.11
3	chr3_85442013	0.07	0.05	0.06	-0.08	-0.06	-0.09	0.05	SoilOrgC	0.09
3	chr3_85444381	0.07	0.07	0.04	-0.11	-0.07	-0.08	0.06	precWQ	0.11
3	chr4_77531109	0.06	-0.01	0.16	-0.08	-0.17	0.08	0.09	precDQ	0.17
3	chr4_77531435	0.08	0.02	0.22	-0.04	-0.15	0.04	0.13	precSeason	0.22
3	chr4_88530878	0.06	-0.23	0.03	0.13	0.13	-0.03	0.04	minTemp	0.23
3	chr4_88531308	0.08	-0.24	0.13	0.09	0.08	-0.08	0.11	minTemp	0.24
3	chr4_88725345	-0.07	0.24	-0.23	-0.27	0.09	-0.08	0.06	precWQ	0.27
3	chr4_88729307	-0.06	0.23	-0.21	-0.14	0.00	-0.03	-0.12	minTemp	0.23
3	chr4_89034764	0.08	-0.03	-0.06	0.00	0.11	-0.16	0.03	SoilOrgC	0.16
3	chr4_89035031	0.08	-0.10	-0.01	0.09	0.04	0.00	0.00	minTemp	0.10
3	chr4_90333060	0.06	-0.15	0.16	0.08	-0.04	0.08	0.12	precSeason	0.16
3	chr4_90455438	-0.07	0.08	-0.08	-0.06	-0.06	-0.01	-0.20	LandUse	0.20
3	chr5_11690121	0.07	0.04	0.13	-0.06	-0.02	-0.07	-0.02	precSeason	0.13
3	chr5_16730855	0.08	-0.13	-0.02	0.03	0.05	0.01	0.07	minTemp	0.13
3	chr5_16743859	0.07	-0.14	0.10	0.05	-0.09	0.07	0.09	minTemp	0.14
3	chr5_19517440	0.07	-0.15	0.06	0.10	0.04	-0.08	-0.01	minTemp	0.15
3	chr5_36417319	-0.08	0.05	-0.24	-0.10	0.11	0.05	-0.03	precSeason	0.24
3	chr5_38454962	0.07	0.12	0.21	-0.19	-0.18	-0.10	0.05	precSeason	0.21
3	chr5_38465056	0.07	-0.06	0.15	-0.14	-0.04	-0.07	0.11	precSeason	0.15
3	chr5_4421038	0.07	0.00	0.14	0.14	-0.01	-0.04	-0.04	precWQ	0.14
3	chr5_48748802	-0.06	0.03	-0.15	-0.05	0.06	0.04	-0.12	precSeason	0.15
3	chr5_7508595	0.07	-0.18	0.16	0.31	-0.14	0.30	-0.03	precWQ	0.31

3	chr6_11688993	-0.07	-0.06	0.03	0.18	0.06	0.18	-0.02	precWQ	0.18
3	chr6_12414334	-0.08	0.12	-0.04	-0.12	0.06	-0.08	0.01	minTemp	0.12
3	chr6_18182501	-0.07	0.03	0.12	-0.14	-0.07	-0.02	0.11	precWQ	0.14
3	chr6_6811973	-0.07	0.08	-0.08	0.01	-0.02	0.07	-0.11	LandUse	0.11
3	chr6_6824738	-0.06	0.09	-0.04	0.01	-0.01	0.07	-0.07	minTemp	0.09
3	chr6_6831321	-0.07	0.11	-0.08	-0.06	-0.02	0.02	-0.04	minTemp	0.11
3	chr7_20959463	0.08	0.03	0.17	-0.05	-0.11	-0.03	-0.01	precSeason	0.17
3	chr7_20980495	0.08	0.05	0.21	-0.04	-0.12	-0.05	0.01	precSeason	0.21
3	chr7_20995969	0.08	0.03	0.19	-0.06	-0.10	-0.05	0.00	precSeason	0.19
3	chr7_34140184	0.06	-0.12	0.04	0.17	0.15	-0.17	-0.06	precWQ	0.17
3	chr7_34338728	-0.07	0.09	-0.08	-0.07	0.03	0.01	-0.08	minTemp	0.09
3	chr8_3133225	-0.07	-0.07	0.00	0.12	-0.02	0.14	-0.06	SoilOrgC	0.14
3	chr8_5681926	0.08	0.01	-0.02	-0.11	0.11	-0.22	0.13	SoilOrgC	0.22
3	chr8_5682410	0.06	-0.19	0.15	0.10	0.07	-0.02	0.18	minTemp	0.19
3	chr8_6217694	-0.07	-0.06	-0.06	0.14	-0.02	0.09	-0.12	precWQ	0.14
3	chr8_6368423	0.08	0.05	0.30	-0.09	-0.17	0.00	0.18	precSeason	0.30
3	chr8_6499160	0.07	-0.15	0.23	0.14	-0.06	0.04	0.10	precSeason	0.23
4	chr1_100341963	0.06	0.12	-0.06	-0.16	0.05	-0.25	0.01	SoilOrgC	0.25
4	chr1_172818620	0.06	0.11	-0.17	-0.10	0.13	-0.24	-0.12	SoilOrgC	0.24
4	chr1_32707669	0.06	0.05	0.13	0.13	-0.14	0.02	-0.18	LandUse	0.18
4	chr1_32777323	-0.06	-0.05	-0.12	-0.05	0.12	0.00	0.11	precDQ	0.12
4	chr1_32824711	0.06	0.02	0.13	0.13	-0.13	0.01	-0.16	LandUse	0.16
4	chr1_33669632	0.06	0.05	0.03	-0.07	-0.01	-0.11	-0.08	SoilOrgC	0.11
4	chr1_45644005	0.06	0.02	0.13	0.03	0.04	-0.07	-0.01	precSeason	0.13
4	chr10_10390359	0.06	0.23	0.02	-0.08	-0.14	0.08	-0.10	minTemp	0.23
4	chr10_13030303	-0.06	-0.13	-0.02	0.14	-0.03	0.11	0.00	precWQ	0.14
4	chr10_13577303	-0.06	-0.02	-0.09	0.00	-0.01	0.07	0.07	precSeason	0.09
4	chr10_13580811	0.06	0.21	0.05	-0.08	-0.09	-0.04	-0.11	minTemp	0.21
4	chr10_13742394	0.06	-0.06	0.00	-0.08	0.06	-0.06	-0.01	precWQ	0.08
4	chr10_13791576	-0.06	0.00	0.08	-0.05	-0.10	0.01	0.07	precDQ	0.10
4	chr10_14253503	0.06	-0.09	0.08	0.17	-0.03	0.12	-0.12	precWQ	0.17
4	chr10_14355919	0.06	-0.01	-0.03	0.10	0.07	0.03	-0.11	LandUse	0.11
4	chr10_14414067	-0.06	0.03	0.01	-0.05	-0.04	0.01	0.13	LandUse	0.13
4	chr10_14415294	0.06	-0.09	0.22	0.18	-0.10	0.11	-0.17	precSeason	0.22
4	chr10_14430487	-0.07	-0.03	0.01	0.01	-0.04	0.02	0.10	LandUse	0.10
4	chr10_14439806	0.06	-0.06	0.10	0.09	-0.02	-0.01	-0.11	LandUse	0.11

4	chr10_14462197	-0.06	0.04	-0.08	-0.14	0.01	0.00	0.19	LandUse	0.19
4	chr10_14475584	-0.06	-0.05	-0.05	-0.01	0.04	-0.04	0.08	LandUse	0.08
4	chr10_14628900	0.06	0.12	-0.18	-0.07	0.05	-0.06	-0.13	precSeason	0.18
4	chr10_8418666	-0.07	-0.06	-0.08	0.18	0.01	0.14	-0.03	precWQ	0.18
4	chr10_8446512	-0.08	-0.02	-0.18	0.05	0.09	0.02	0.06	precSeason	0.18
4	chr10_9041002	-0.06	0.06	-0.10	-0.09	0.05	0.00	0.17	LandUse	0.17
4	chr10_9044433	-0.06	0.02	-0.02	-0.07	0.00	-0.04	0.06	precWQ	0.07
4	chr10_9070584	0.06	0.06	-0.13	0.00	0.03	-0.05	-0.23	LandUse	0.23
4	chr10_9070642	0.06	0.06	-0.13	0.00	0.03	-0.05	-0.23	LandUse	0.23
4	chr10_9078157	0.06	0.07	-0.19	0.04	0.12	-0.03	-0.20	LandUse	0.20
4	chr10_9085449	0.07	0.07	-0.11	0.03	0.08	-0.07	-0.16	LandUse	0.16
4	chr11_16990425	-0.06	0.00	-0.05	-0.09	-0.04	0.11	0.16	LandUse	0.16
4	chr12_7901143	-0.07	-0.14	0.10	0.21	0.00	0.15	0.08	precWQ	0.21
4	chr14_2632555	0.06	0.08	-0.16	-0.20	0.21	-0.32	0.03	SoilOrgC	0.32
4	chr14_2637001	0.06	-0.02	-0.04	0.06	0.07	-0.04	-0.07	LandUse	0.07
4	chr14_2833375	-0.07	-0.15	0.13	0.03	-0.07	0.11	0.14	minTemp	0.15
4	chr14_2836838	-0.08	-0.13	0.02	0.02	0.10	0.10	0.25	LandUse	0.25
4	chr14_2837904	-0.07	-0.15	0.00	0.11	0.10	0.13	0.20	LandUse	0.20
4	chr14_2844020	0.06	0.26	-0.01	-0.01	-0.09	-0.06	-0.23	minTemp	0.26
4	chr14_3191934	0.06	0.03	0.03	0.07	0.01	-0.04	-0.11	LandUse	0.11
4	chr14_3200622	-0.06	0.08	0.04	-0.11	-0.07	-0.06	0.12	LandUse	0.12
4	chr15_8550227	-0.06	-0.10	-0.06	0.09	0.10	-0.02	0.06	precDQ	0.10
4	chr15_8551715	-0.06	-0.04	0.01	-0.05	0.03	0.01	0.18	LandUse	0.18
4	chr17_6033798	-0.06	0.03	0.00	-0.03	-0.02	0.05	0.14	LandUse	0.14
4	chr17_7030364	-0.06	-0.06	0.02	-0.15	0.00	0.04	0.24	LandUse	0.24
4	chr17_7056075	0.06	0.02	-0.10	0.20	0.09	-0.08	-0.31	LandUse	0.31
4	chr17_7071574	0.06	0.08	-0.03	0.02	0.01	-0.03	-0.18	LandUse	0.18
4	chr17_7239853	-0.06	-0.06	0.23	-0.01	-0.05	0.05	0.30	LandUse	0.30
4	chr17_7909817	-0.06	-0.04	0.04	0.12	-0.08	0.22	0.03	SoilOrgC	0.22
4	chr17_7916208	0.06	0.10	0.08	0.05	-0.02	-0.13	-0.13	SoilOrgC	0.13
4	chr17_8062409	-0.06	0.01	0.14	0.02	-0.13	-0.02	0.09	precSeason	0.14
4	chr2_101936788	0.06	0.13	-0.10	-0.14	0.07	-0.17	0.10	SoilOrgC	0.17
4	chr2_107677712	0.06	-0.12	0.07	0.16	0.05	0.09	-0.06	precWQ	0.16
4	chr2_117744943	0.06	-0.01	-0.07	0.22	0.09	-0.01	-0.19	precWQ	0.22
4	chr2_117755910	0.06	0.03	-0.03	0.28	-0.01	0.05	-0.25	precWQ	0.28
4	chr2_124534594	-0.06	-0.01	0.11	-0.08	-0.15	0.13	0.14	precDQ	0.15

4	chr2_138249521	-0.07	-0.03	0.04	-0.14	-0.02	0.04	0.22	LandUse	0.22
4	chr2_138261275	-0.06	-0.03	0.07	-0.14	-0.04	0.03	0.19	LandUse	0.19
4	chr2_138273548	-0.06	-0.06	-0.01	-0.05	0.00	0.04	0.14	LandUse	0.14
4	chr2_138279828	-0.07	0.02	0.02	-0.24	-0.09	0.02	0.20	precWQ	0.24
4	chr2_138285456	-0.08	-0.05	0.06	-0.14	-0.07	0.04	0.20	LandUse	0.20
4	chr2_138289606	-0.06	-0.12	-0.01	-0.04	0.03	0.05	0.14	LandUse	0.14
4	chr2_138289999	-0.08	-0.01	-0.05	-0.21	0.00	0.02	0.21	precWQ	0.21
4	chr2_139065120	-0.06	0.03	0.04	0.01	-0.06	0.03	0.07	LandUse	0.07
4	chr2_141370744	-0.07	-0.19	0.04	0.18	-0.04	0.23	0.02	SoilOrgC	0.23
4	chr2_141386781	-0.06	-0.16	0.02	0.10	0.01	0.12	0.12	minTemp	0.16
4	chr2_79010957	0.07	0.23	-0.13	-0.07	0.05	-0.10	-0.16	minTemp	0.23
4	chr2_79024416	0.08	0.20	-0.18	-0.05	0.15	-0.18	-0.16	minTemp	0.20
4	chr2_79029493	0.07	0.14	0.19	0.27	-0.19	0.16	-0.28	LandUse	0.28
4	chr2_82079608	0.06	0.01	-0.10	0.08	0.10	-0.16	-0.18	LandUse	0.18
4	chr2_82421131	0.06	0.04	-0.08	0.09	0.07	-0.18	-0.22	LandUse	0.22
4	chr2_82801925	0.06	0.30	-0.18	-0.11	-0.09	-0.07	-0.19	minTemp	0.30
4	chr2_82806510	0.06	0.29	-0.18	-0.10	-0.09	-0.07	-0.20	minTemp	0.29
4	chr2_82818189	0.06	0.30	-0.17	-0.11	-0.10	-0.07	-0.20	minTemp	0.30
4	chr2_82839244	0.06	0.30	-0.19	-0.10	-0.08	-0.06	-0.20	minTemp	0.30
4	chr2_82898529	0.06	0.03	-0.10	0.00	0.02	-0.11	-0.14	LandUse	0.14
4	chr2_82899418	0.06	0.28	-0.16	-0.08	-0.10	-0.04	-0.17	minTemp	0.28
4	chr2_88150231	-0.06	-0.13	-0.02	0.06	0.08	0.05	0.15	LandUse	0.15
4	chr2_99403133	-0.06	-0.08	0.05	-0.12	-0.07	0.08	0.15	LandUse	0.15
4	chr21_4350665	0.07	-0.01	-0.03	-0.07	0.01	-0.01	-0.08	LandUse	0.08
4	chr21_6586889	0.06	0.09	-0.06	-0.04	0.12	-0.21	-0.07	SoilOrgC	0.21
4	chr21_6765149	-0.06	-0.05	0.01	0.03	-0.02	0.06	0.08	LandUse	0.08
4	chr21_6768465	-0.08	-0.13	-0.01	0.12	-0.03	0.13	0.05	SoilOrgC	0.13
4	chr27_7122554	-0.06	-0.13	-0.06	-0.01	0.08	-0.06	0.12	minTemp	0.13
4	chr28_829783	0.06	-0.04	-0.07	0.17	0.10	-0.01	-0.14	precWQ	0.17
4	chr3_100431475	0.07	0.16	0.02	-0.05	0.02	-0.15	-0.04	minTemp	0.16
4	chr3_102789546	0.06	0.08	0.03	0.06	0.04	-0.04	-0.09	LandUse	0.09
4	chr3_102800999	0.06	0.08	0.05	0.06	0.03	-0.05	-0.11	LandUse	0.11
4	chr3_103130166	-0.06	-0.04	0.01	-0.08	-0.12	0.05	0.06	precDQ	0.12
4	chr3_103133548	-0.06	-0.04	0.00	-0.07	-0.11	0.05	0.06	precDQ	0.11
4	chr3_16666369	0.07	0.09	0.11	0.03	-0.10	0.01	-0.07	precSeason	0.11
4	chr3_16666427	0.06	0.12	0.21	-0.03	-0.10	-0.09	-0.02	precSeason	0.21

4	chr3_99996874	0.07	0.16	-0.03	0.02	0.02	-0.10	-0.17	LandUse	0.17
4	chr4_27805454	-0.06	-0.02	-0.09	-0.11	-0.01	0.05	0.13	LandUse	0.13
4	chr4_40940247	0.06	0.06	-0.09	-0.13	0.06	-0.02	-0.03	precWQ	0.13
4	chr4_41000225	-0.06	0.05	-0.04	0.06	-0.01	-0.02	-0.08	LandUse	0.08
4	chr4_41000230	-0.06	0.05	-0.04	0.06	-0.01	-0.02	-0.08	LandUse	0.08
4	chr4_41000236	-0.06	0.05	-0.04	0.06	-0.01	-0.02	-0.08	LandUse	0.08
4	chr4_41000242	-0.06	0.05	-0.04	0.06	-0.01	-0.02	-0.08	LandUse	0.08
4	chr4_41068796	-0.06	0.00	0.02	0.05	-0.05	-0.02	-0.02	precDQ	0.05
4	chr4_82451385	-0.06	-0.08	-0.03	-0.04	0.03	0.03	0.12	LandUse	0.12
4	chr4_83790761	-0.06	-0.07	-0.11	0.16	0.19	-0.03	0.01	precDQ	0.19
4	chr4_88189134	0.06	0.07	-0.13	0.00	0.10	-0.10	-0.16	LandUse	0.16
4	chr4_89868144	-0.07	-0.23	0.11	0.10	-0.01	0.21	0.17	minTemp	0.23
4	chr4_89870632	-0.06	-0.17	0.08	0.08	0.05	0.19	0.19	LandUse	0.19
4	chr5_19472668	-0.06	0.02	0.01	-0.07	-0.08	0.03	-0.01	precDQ	0.08
4	chr5_19496481	0.07	0.01	-0.09	-0.11	0.13	-0.19	-0.01	SoilOrgC	0.19
4	chr5_19541510	-0.06	-0.15	0.26	0.15	-0.16	0.23	0.12	precSeason	0.26
4	chr5_19548802	0.06	0.04	-0.03	-0.02	0.03	0.01	-0.03	minTemp	0.04
4	chr5_19550705	0.06	0.08	-0.15	-0.11	0.13	-0.10	0.03	precSeason	0.15
4	chr5_19552051	-0.08	-0.09	0.06	0.12	-0.02	0.09	0.06	precWQ	0.12
4	chr5_19554343	0.06	0.14	-0.31	-0.27	0.13	-0.24	-0.16	precSeason	0.31
4	chr5_19573070	-0.06	-0.11	0.20	0.19	-0.09	0.13	0.10	precSeason	0.20
4	chr5_19601002	-0.07	-0.13	0.06	0.12	0.01	0.10	0.16	LandUse	0.16
4	chr5_30830034	-0.07	-0.11	0.18	0.10	-0.09	0.15	0.11	precSeason	0.18
4	chr5_30830987	0.06	0.01	-0.04	0.15	0.00	0.00	-0.18	LandUse	0.18
4	chr5_30836679	-0.08	-0.25	-0.05	0.02	0.03	0.24	0.19	minTemp	0.25
4	chr5_30853881	-0.06	-0.11	0.07	0.08	-0.02	0.07	0.06	minTemp	0.11
4	chr5_30865734	0.06	0.06	-0.05	-0.02	0.01	-0.06	-0.10	LandUse	0.10
4	chr5_30866179	0.06	0.14	-0.03	-0.13	-0.01	-0.07	-0.02	minTemp	0.14
4	chr5_38318256	-0.06	-0.03	-0.09	-0.04	0.07	-0.04	0.05	precSeason	0.09
4	chr5_38441756	-0.06	0.02	0.03	-0.13	-0.07	-0.07	0.09	precWQ	0.13
4	chr5_48708049	-0.06	0.07	0.00	-0.01	-0.14	0.05	-0.02	precDQ	0.14
4	chr8_3107239	-0.07	-0.14	0.02	0.09	0.00	0.07	0.06	minTemp	0.14
4	chr8_3283167	-0.06	-0.16	0.04	0.18	0.00	0.19	0.05	SoilOrgC	0.19
4	chr8_3300364	-0.06	-0.14	0.04	0.14	0.00	0.15	0.06	SoilOrgC	0.15
4	chr8_3318996	-0.06	-0.16	0.10	0.19	-0.05	0.19	0.01	precWQ	0.19
4	chr8_3338714	-0.06	-0.18	0.03	0.11	-0.03	0.08	0.01	minTemp	0.18

4	chr8_6342169	-0.07	-0.01	0.15	-0.05	-0.05	0.11	0.20	LandUse	0.20
4	chr8_6481351	0.06	0.12	0.21	0.09	-0.15	0.03	-0.17	precSeason	0.21
5	chr1_1329783	-0.06	0.07	-0.09	-0.03	0.03	-0.04	0.07	precSeason	0.09
5	chr1_1332114	0.06	-0.04	-0.05	0.02	0.08	-0.01	-0.11	LandUse	0.11
5	chr1_133839237	-0.06	0.06	-0.11	0.18	0.03	0.01	-0.08	precWQ	0.18
5	chr1_133845303	0.07	0.09	0.08	-0.05	-0.15	-0.06	-0.20	LandUse	0.20
5	chr1_146781584	-0.06	0.18	0.01	-0.20	0.06	-0.23	0.14	SoilOrgC	0.23
5	chr1_177460567	0.06	-0.24	0.10	0.09	0.02	0.08	-0.03	minTemp	0.24
5	chr1_485373	-0.06	0.16	0.10	-0.13	-0.07	-0.14	0.09	minTemp	0.16
5	chr1_69960245	0.06	-0.05	-0.07	0.00	-0.03	0.09	-0.12	LandUse	0.12
5	chr1_69999819	-0.08	0.08	-0.05	0.01	0.11	-0.10	0.15	LandUse	0.15
5	chr1_71668364	-0.06	0.22	-0.31	-0.21	0.16	-0.23	0.08	precSeason	0.31
5	chr10_11737471	-0.06	0.04	-0.05	0.06	0.08	0.00	0.10	LandUse	0.10
5	chr10_11748993	0.06	-0.01	0.00	0.03	-0.09	0.15	-0.15	LandUse	0.15
5	chr10_11756804	-0.06	0.04	-0.09	0.04	0.11	-0.04	0.09	precDQ	0.11
5	chr10_12773142	0.06	-0.11	0.04	0.09	-0.01	0.13	-0.03	SoilOrgC	0.13
5	chr10_13017263	-0.06	0.09	0.00	-0.14	0.02	-0.19	0.15	SoilOrgC	0.19
5	chr10_13032537	-0.06	0.17	-0.01	-0.17	-0.03	-0.10	0.16	minTemp	0.17
5	chr10_13085516	0.08	0.02	-0.09	-0.14	-0.06	0.05	-0.09	precWQ	0.14
5	chr10_13085896	0.07	0.04	-0.14	-0.14	-0.01	0.02	-0.09	precSeason	0.14
5	chr10_13968283	-0.06	-0.07	-0.05	0.01	0.15	-0.01	0.18	LandUse	0.18
5	chr10_14066105	-0.06	-0.02	0.02	-0.08	-0.02	-0.04	0.16	LandUse	0.16
5	chr10_14068597	-0.06	-0.11	0.00	0.00	0.05	-0.05	0.12	LandUse	0.12
5	chr10_14344058	0.06	-0.18	0.17	0.15	-0.10	0.22	-0.05	SoilOrgC	0.22
5	chr10_14347193	-0.07	0.17	-0.25	-0.10	0.23	-0.25	0.05	SoilOrgC	0.25
5	chr10_14577341	-0.06	-0.17	-0.05	0.05	0.09	-0.02	0.16	minTemp	0.17
5	chr10_4096505	0.07	-0.03	0.10	0.12	-0.18	0.17	-0.24	LandUse	0.24
5	chr10_8107262	0.06	-0.08	-0.12	0.11	0.13	0.07	-0.12	precDQ	0.13
5	chr10_9073389	0.06	0.06	-0.05	-0.01	-0.01	-0.10	-0.23	LandUse	0.23
5	chr10_9074234	-0.06	0.13	-0.12	-0.03	0.06	0.02	0.06	minTemp	0.13
5	chr10_9591892	-0.06	-0.01	-0.06	-0.12	0.08	-0.15	0.20	LandUse	0.20
5	chr10_9596232	-0.06	0.14	0.03	-0.15	0.00	-0.12	0.18	LandUse	0.18
5	chr10_9602198	0.06	-0.10	-0.09	0.16	0.00	0.14	-0.23	LandUse	0.23
5	chr10_9610040	0.06	-0.17	-0.14	0.06	0.09	0.06	-0.14	minTemp	0.17
5	chr10_9929014	-0.08	0.07	-0.19	-0.17	0.16	-0.18	0.18	precSeason	0.19
5	chr10_9989994	0.06	-0.04	0.10	0.20	-0.09	0.13	-0.25	LandUse	0.25

5	chr11_19801888	-0.06	0.02	0.03	0.06	0.03	-0.06	0.09	LandUse	0.09
5	chr13_15622363	0.06	0.04	0.19	-0.12	-0.16	-0.07	-0.04	precSeason	0.19
5	chr14_5309155	0.06	-0.22	0.08	0.08	0.03	0.02	-0.08	minTemp	0.22
5	chr14_5310564	0.06	-0.10	0.10	-0.05	-0.02	-0.07	-0.07	minTemp	0.10
5	chr14_5314625	0.06	-0.10	0.09	-0.09	-0.02	-0.12	-0.07	SoilOrgC	0.12
5	chr14_5315870	-0.06	-0.01	-0.03	0.16	0.01	0.15	0.12	precWQ	0.16
5	chr14_5319263	0.06	-0.19	0.18	0.19	-0.06	0.12	-0.13	precWQ	0.19
5	chr15_10462608	0.06	-0.08	-0.01	0.23	0.01	0.14	-0.20	precWQ	0.23
5	chr17_5923013	0.06	0.11	-0.11	-0.22	-0.06	-0.01	-0.09	precWQ	0.22
5	chr17_7223488	0.07	0.03	0.05	0.12	-0.03	0.14	-0.15	LandUse	0.15
5	chr17_7230401	0.07	-0.05	0.13	0.18	-0.10	0.19	-0.18	SoilOrgC	0.19
5	chr17_7260060	-0.06	-0.03	0.08	0.09	0.07	0.07	0.21	LandUse	0.21
5	chr17_7763240	-0.06	0.11	-0.05	0.02	-0.01	-0.12	0.01	SoilOrgC	0.12
5	chr17_7770840	-0.06	0.12	-0.17	-0.09	0.17	-0.25	0.02	SoilOrgC	0.25
5	chr17_7771595	0.06	-0.12	0.18	-0.06	-0.10	0.06	0.03	precSeason	0.18
5	chr19_1828639	-0.06	-0.04	-0.02	0.06	0.12	0.00	0.14	LandUse	0.14
5	chr2_102036673	-0.06	0.10	-0.01	-0.04	0.00	-0.09	0.08	minTemp	0.10
5	chr2_107670586	0.06	-0.02	-0.03	-0.09	-0.04	0.00	-0.13	LandUse	0.13
5	chr2_140384276	-0.07	-0.08	0.18	0.23	-0.05	0.15	0.14	precWQ	0.23
5	chr2_31675935	-0.06	-0.16	0.15	0.16	0.06	0.03	0.22	LandUse	0.22
5	chr2_31687703	-0.06	-0.16	0.18	0.15	0.04	0.03	0.23	LandUse	0.23
5	chr2_31710186	0.06	0.16	-0.16	-0.14	-0.05	-0.05	-0.26	LandUse	0.26
5	chr2_31711772	0.06	0.13	-0.16	-0.11	-0.04	-0.03	-0.24	LandUse	0.24
5	chr2_31712297	0.07	0.13	-0.15	-0.03	-0.05	0.02	-0.28	LandUse	0.28
5	chr2_31716689	0.06	0.15	-0.11	-0.23	-0.10	-0.07	-0.17	precWQ	0.23
5	chr2_31721970	0.06	0.15	-0.21	-0.16	0.00	-0.05	-0.19	precSeason	0.21
5	chr2_31723496	0.07	0.01	0.05	-0.07	-0.04	0.00	-0.14	LandUse	0.14
5	chr2_31725655	0.06	0.15	0.07	0.03	-0.15	0.03	-0.30	LandUse	0.30
5	chr21_3367906	-0.06	0.11	0.08	-0.07	-0.08	-0.18	0.06	SoilOrgC	0.18
5	chr21_4166986	0.06	-0.15	0.02	-0.02	0.01	0.00	0.00	minTemp	0.15
5	chr23_5741035	0.06	0.01	0.04	0.02	-0.06	0.11	-0.12	LandUse	0.12
5	chr23_5742024	0.06	0.01	0.04	0.02	-0.06	0.11	-0.12	LandUse	0.12
5	chr23_5854584	0.06	0.01	0.04	0.02	-0.06	0.11	-0.12	LandUse	0.12
5	chr23_5858495	0.06	0.01	0.04	0.02	-0.06	0.11	-0.12	LandUse	0.12
5	chr23_5870583	0.06	0.01	0.04	0.02	-0.06	0.11	-0.12	LandUse	0.12
5	chr23_5875592	0.06	0.01	0.04	0.02	-0.06	0.11	-0.12	LandUse	0.12

5	chr3_106170444	0.06	-0.15	0.04	0.14	0.02	0.11	-0.11	minTemp	0.15
5	chr3_106183584	0.06	-0.13	-0.03	0.16	0.08	0.09	-0.11	precWQ	0.16
5	chr3_51626845	0.06	-0.08	0.14	0.09	-0.09	0.06	-0.14	precSeason	0.14
5	chr3_72070901	-0.06	0.02	0.01	0.16	-0.05	0.10	0.06	precWQ	0.16
5	chr3_72074558	-0.06	0.00	0.05	0.21	-0.08	0.18	0.08	precWQ	0.21
5	chr3_72465046	-0.06	0.01	-0.04	-0.04	0.07	-0.06	0.17	LandUse	0.17
5	chr3_72534190	0.06	0.01	-0.07	-0.01	-0.03	-0.06	-0.21	LandUse	0.21
5	chr4_20363461	-0.06	-0.02	-0.06	0.04	-0.01	0.08	0.13	LandUse	0.13
5	chr4_20364134	-0.06	0.00	-0.08	0.06	0.05	-0.06	0.05	precSeason	0.08
5	chr4_20401325	-0.06	0.15	-0.17	-0.24	0.00	-0.09	0.20	precWQ	0.24
5	chr4_20405768	-0.06	0.15	-0.17	-0.24	0.00	-0.09	0.20	precWQ	0.24
5	chr4_20414170	-0.06	0.15	-0.17	-0.24	0.00	-0.09	0.20	precWQ	0.24
5	chr4_20426307	-0.06	0.02	-0.04	-0.09	0.04	-0.09	0.18	LandUse	0.18
5	chr4_51833641	-0.06	-0.05	0.05	-0.06	0.08	-0.03	0.22	LandUse	0.22
5	chr4_76289426	-0.07	0.15	0.09	-0.02	0.05	-0.06	0.12	minTemp	0.15
5	chr4_77478540	-0.06	0.13	0.02	-0.01	0.02	-0.02	0.08	minTemp	0.13
5	chr4_82876710	-0.07	0.09	0.15	-0.03	-0.07	-0.04	0.15	LandUse	0.15
5	chr4_83947352	-0.06	0.06	-0.25	-0.13	0.28	-0.17	0.12	precDQ	0.28
5	chr4_84638948	-0.06	-0.09	-0.18	0.08	0.20	-0.12	0.05	precDQ	0.20
5	chr4_84643310	0.06	0.04	0.31	0.00	-0.26	0.08	-0.06	precSeason	0.31
5	chr4_84645518	-0.06	0.04	-0.25	-0.07	0.28	-0.21	0.07	precDQ	0.28
5	chr4_86620030	-0.06	0.08	-0.03	-0.29	-0.04	-0.18	0.21	precWQ	0.29
5	chr4_86636826	0.06	0.04	-0.13	0.08	0.05	0.16	-0.13	SoilOrgC	0.16
5	chr4_86637617	0.06	0.16	-0.23	-0.06	0.05	0.06	-0.14	precSeason	0.23
5	chr4_88172497	-0.06	0.21	-0.28	-0.32	0.14	-0.21	0.14	precWQ	0.32
5	chr4_88173584	-0.06	0.13	-0.24	-0.24	0.08	-0.09	0.15	precWQ	0.24
5	chr4_88176070	-0.06	0.25	-0.15	-0.26	0.04	-0.18	0.09	precWQ	0.26
5	chr4_88508395	-0.07	-0.10	-0.03	0.13	0.24	-0.08	0.16	precDQ	0.24
5	chr4_88515508	-0.06	-0.12	-0.03	0.28	0.18	-0.04	0.01	precWQ	0.28
5	chr4_88515727	-0.06	-0.02	-0.12	0.11	0.20	-0.17	0.00	precDQ	0.20
5	chr4_88516552	0.06	0.03	0.06	-0.04	-0.11	0.06	-0.07	precDQ	0.11
5	chr4_88520091	0.06	-0.03	-0.15	-0.04	0.12	0.03	-0.10	precSeason	0.15
5	chr4_88964291	0.06	-0.01	0.02	-0.06	-0.10	0.06	-0.06	precDQ	0.10
5	chr4_88974495	-0.06	0.13	-0.25	-0.20	0.14	-0.17	0.09	precSeason	0.25
5	chr5_11770052	0.06	-0.02	0.14	-0.09	-0.05	-0.02	-0.03	precSeason	0.14
5	chr5_19491957	0.06	-0.10	0.09	0.06	-0.07	0.18	-0.08	SoilOrgC	0.18

5	chr5_27541569	-0.07	-0.04	-0.12	-0.02	0.15	-0.10	0.14	precDQ	0.15
5	chr5_38442644	0.06	0.00	0.13	-0.03	-0.10	0.02	-0.11	precSeason	0.13
5	chr5_599713	-0.06	0.08	-0.15	-0.11	0.17	-0.20	0.06	SoilOrgC	0.20
5	chr5_600083	-0.06	0.13	-0.29	-0.23	0.14	-0.18	0.10	precSeason	0.29
5	chr6_19152621	-0.06	0.03	0.03	-0.08	0.06	-0.14	0.13	SoilOrgC	0.14
5	chr6_19198674	-0.06	0.03	0.06	-0.07	0.02	-0.07	0.14	LandUse	0.14
5	chr7_32487480	0.06	-0.01	0.07	-0.05	-0.01	0.04	-0.02	precSeason	0.07
5	chr7_32511488	0.06	-0.05	-0.03	-0.10	0.06	-0.08	-0.10	precWQ	0.10
5	chr7_32512010	-0.07	0.08	-0.01	-0.05	-0.02	-0.08	0.11	LandUse	0.11
5	chr7_32566048	-0.06	-0.07	-0.06	0.15	0.03	0.02	0.07	precWQ	0.15
5	chr7_32579990	-0.06	-0.02	0.05	0.22	-0.06	0.15	0.07	precWQ	0.22
5	chr7_34191572	0.06	0.00	0.22	0.00	-0.15	0.10	-0.03	precSeason	0.22

Table S5.10. Allele frequencies by ecotype for selected outlier SNPs (RDA correlation ≥ 0.3 -0.28) and correlation scores with mean environmental values. (**Mean value:** corresponds to the average value of the environmental predictor for each ecotype. **r:** Pearson correlation value between outlier SNP and mean value of environmental predictor for each ecotype. **Correlation trend:** best-adjusted trendline to describe correlation (pol (2nd): quadratic polynomial function, pol (3rd): cubic polynomial function . **RDA cor:** RDA correlation between outlier SNP and predictor. *Allele frequencies above 0.7 are indicated in bold style).

		Allele frequencies by ecotype*														
		E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	r	Correlation trend	RDA cor
Outlier SNPs (bio6) Minimum Temperature in the coldest month	Mean value (bio6)	55.64	69.80	98.53	61.67	87.32	116.1	18.25	114.5	73.53	105.5	104.1	112.9			
	chr14_2834643_C_T	0.63	**0.70**	**0.79**	**0.72**	**0.74**	**0.88**	0.40	**0.81**	**0.89**	**0.80**	0.65	**0.90**	0.85	pol (3rd)	0.36
	chr5_49525047_T_C	0.01	0.00	0.11	0.00	0.16	0.26	0.00	0.03	0.00	0.00	0.00	0.00	0.39	pol (3rd)	0.34
	chr14_2839546_G_A	0.32	0.38	0.63	0.25	0.55	0.51	0.18	0.67	0.50	0.40	0.40	0.55	0.81	pol (3rd)	0.33
	chr1_132953669_A_G	0.33	0.58	0.66	0.63	0.66	0.67	0.28	0.67	**0.71**	**0.70**	0.45	0.55	0.75	pol (3rd)	0.31
	chr2_78280162_C_T	0.50	0.63	**0.87**	**0.75**	0.66	**0.82**	0.55	0.83	**0.79**	**0.70**	**0.95**	**0.80**	0.76	pol (3rd)	0.31
	chr2_78280175_T_C	0.50	0.63	**0.87**	**0.75**	0.66	**0.82**	0.55	0.83	**0.79**	**0.70**	**0.95**	**0.80**	0.76	pol (3rd)	0.31
	chr4_88536330_G_A	0.35	0.40	0.63	0.34	0.29	0.66	0.23	0.33	0.55	0.40	0.30	0.60	0.96	pol (3rd)	0.31
Outlier SNPs (bio15) Precipitation seasonality	Mean value (bio15)	87.41	105	71.42	90.9	50.3	82.61	105.8	103.7	141	136.9	84.81	121.6			
	chr20_4183810_G_A	0.04	0.00	0.11	0.03	0.34	0.12	0.03	0.00	0.03	0.00	0.00	0.00	0.91	pol (2nd)	0.39
	chr17_6740515_C_T	0.03	0.00	0.11	0.09	0.45	0.02	0.00	0.03	0.00	0.05	0.00	0.00	0.22	pol (2nd)	0.38
	chr4_84643310_T_C	0.06	0.20	0.13	0.31	0.03	0.16	0.20	0.33	0.32	0.50	0.05	0.15	0.71	lineal	0.31
	chr1_71668364_C_T	0.19	0.03	0.05	0.06	0.42	0.38	0.03	0.11	0.11	0.10	0.05	0.20	0.32	pol (2nd)	0.31
	chr5_19554343_A_T	0.32	0.35	**0.87**	0.31	**0.82**	0.55	0.55	0.58	0.39	0.40	**0.75**	0.40	0.63	pol (2nd)	0.31
	chr10_9402417_C_T	0.56	0.48	0.39	0.53	0.42	0.26	0.45	0.61	**0.76**	**0.75**	0.40	0.70	0.87	pol (2nd)	0.30

		Allele frequencies by ecotype*														
		E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	r	Correlation trend	RDA cor
Outlier SNPs (bio16) Precipitation in the wettest quarter	Mean value (bio16)	1050	970	428.3	758	505	362.2	770.1	1089	458.8	480.7	529.3	767.5			
	chr2_117749087_T_C	0.29	0.20	0.00	0.03	0.08	0.01	0.03	0.19	0.00	0.00	0.00	0.10	0.65	pol (2nd)	0.42
	chr2_140301053_C_T	**0.74**	**0.95**	0.47	**0.72**	0.39	0.56	**0.83**	**0.94**	0.53	0.65	**0.85**	**0.80**	0.60	pol (2nd)	0.37
	chr2_140313253_T_C	**0.74**	**0.95**	0.47	**0.72**	0.39	0.56	**0.83**	**0.94**	0.53	0.60	**0.85**	**0.80**	0.60	pol (2nd)	0.36
	chr10_13599109_C_G	0.60	0.48	0.08	0.44	0.26	0.17	0.53	0.47	0.42	0.40	0.30	0.50	0.59	pol (2nd)	0.35
	chr11_16972959_C_G	**0.85**	**0.90**	0.50	0.63	**0.76**	0.43	0.53	**0.92**	**0.89**	**0.85**	**0.75**	**0.90**	0.49	lineal	0.33
	chr4_3627776_T_C	0.03	0.08	0.08	0.03	0.21	0.30	0.10	0.00	0.26	0.20	0.15	0.00	0.57	pol (2nd)	0.33
	chr4_88172497_C_T	0.40	0.20	0.61	0.63	0.61	0.68	0.40	0.33	0.45	0.15	**0.75**	0.65	0.57	pol (2nd)	0.32
	chr2_140379965_A_G	0.19	0.05	0.47	0.06	0.42	0.32	0.15	0.11	0.42	0.40	0.15	0.10	0.56	pol (2nd)	0.31
	chr21_4353990_C_G	**0.71**	**0.78**	**0.92**	**0.97**	**0.97**	**0.96**	**0.98**	**0.94**	**0.95**	**1.00**	**1.00**	**0.95**	0.56	pol (2nd)	0.31
	chr20_4173643_A_T	0.00	0.00	0.08	0.06	0.21	0.13	0.03	0.00	0.11	0.00	0.00	0.00	0.60	lineal	0.31
	chr4_82439583_T_C	0.36	0.38	0.05	0.09	0.13	0.07	0.13	0.08	0.08	0.15	0.15	0.20	0.60	lineal	0.31
	chr5_7508595_C_T	**0.97**	**1.00**	**0.84**	**1.00**	**0.79**	**0.77**	**0.93**	**0.97**	**0.92**	**1.00**	**0.95**	**1.00**	0.55	pol (2nd)	0.31
	chr2_140367628_A_G	**0.99**	**0.98**	**0.84**	**1.00**	0.68	**0.79**	**0.88**	**1.00**	**0.92**	**1.00**	**1.00**	**1.00**	0.55	pol (2nd)	0.30
Outlier SNPs (bio17) Precipitation in the driest quarter	Mean value (bio17)	79.97	24.8	37.89	59.2	116	42.78	46.1	11.55	9.42	19.6	72.72	11.3			
	chr1_197483358_A_G	0.17	0.05	0.00	0.00	0.26	0.02	0.03	0.00	0.03	0.00	0.05	0.05	0.61	pol (2nd)	0.37
	chr19_1816989_G_A	0.15	0.10	0.00	0.03	0.39	0.10	0.08	0.00	0.08	0.00	0.05	0.05	0.59	pol (2nd)	0.34
	chr1_118390274_T_C	0.46	0.23	0.16	0.13	0.58	0.37	0.18	0.17	0.29	0.10	0.10	0.15	0.54	pol (2nd)	0.30
	chr8_24798255_C_T	0.29	0.23	0.11	0.22	0.26	0.02	0.05	0.00	0.05	0.05	0.10	0.05	0.70	lineal	0.30
	chr20_1187011_A_C	0.11	0.08	0.03	0.13	0.29	0.00	0.05	0.03	0.11	0.20	0.10	0.00	0.54	pol (2nd)	0.29
	chr20_1187011_A_C	0.47	0.23	0.18	0.13	0.58	0.38	0.18	0.17	0.32	0.10	0.15	0.10	0.54	pol (2nd)	0.29

		Allele frequencies by ecotype*														
		E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	r	Correlation trend	RDA cor
Outlier SNPs Soil organic carbon content	Mean value (Soil)	130.1	139	109.1	140	71.9	113.1	127.4	123.5	113.9	111	124.6	105.7			
	chr3_16891572_C_T	0.06	0.10	0.11	0.06	0.58	0.10	0.00	0.08	0.00	0.10	0.00	0.05	0.64	pol (2nd)	0.41
	chr1_23141490_G_A	0.18	0.18	0.37	0.25	**0.76**	0.32	0.10	0.11	0.24	0.30	0.05	0.20	0.64	pol (2nd)	0.40
	chr8_6198062_A_G	**0.96**	**1.00**	**0.87**	**1.00**	**0.71**	**0.93**	**0.98**	**1.00**	**1.00**	**0.90**	**1.00**	**0.95**	0.62	pol (2nd)	0.38
	chr7_30965068_C_T	0.01	0.08	0.03	0.03	0.37	0.07	0.00	0.06	0.00	0.05	0.00	0.05	0.61	pol (2nd)	0.38
	chr19_1819556_A_G	0.04	0.05	0.11	0.22	0.50	0.24	0.08	0.06	0.16	0.00	0.05	0.15	0.61	pol (2nd)	0.37
	chr7_34948717_G_A	0.04	0.15	0.08	0.03	0.53	0.09	0.03	0.06	0.11	0.10	0.00	0.15	0.61	pol (2nd)	0.37
	chr14_6201610_C_T	**0.76**	0.65	0.66	**0.72**	0.29	0.59	**0.88**	**0.81**	0.50	0.65	**0.75**	0.60	0.60	pol (2nd)	0.36
	chr5_35557653_G_A	0.01	0.03	0.03	0.06	0.32	0.07	0.00	0.06	0.03	0.00	0.00	0.05	0.60	pol (2nd)	0.36
	chr1_132947694_G_A	0.06	0.08	0.26	0.22	0.47	0.24	0.20	0.03	0.13	0.25	0.00	0.25	0.59	pol (2nd)	0.35
	chr19_1830033_T_A	0.10	0.05	0.13	0.09	0.37	0.20	0.08	0.03	0.05	0.00	0.05	0.10	0.58	pol (2nd)	0.34
	chr1_132951411_G_A	0.10	0.08	0.13	0.09	0.45	0.22	0.10	0.03	0.13	0.10	0.00	0.10	0.58	pol (2nd)	0.34
	chr1_23164664_C_T	0.01	0.05	0.03	0.06	0.32	0.10	0.03	0.00	0.08	0.10	0.00	0.15	0.58	pol (2nd)	0.33
	chr5_19564424_C_T	0.01	0.05	0.05	0.00	0.26	0.02	0.00	0.00	0.08	0.15	0.00	0.05	0.58	pol (2nd)	0.33
	chr1_118539067_A_G	0.11	0.00	0.11	0.03	0.37	0.12	0.00	0.00	0.21	0.15	0.00	0.10	0.58	pol (2nd)	0.33
	chr10_7190022_G_A	0.04	0.00	0.03	0.00	0.26	0.07	0.00	0.00	0.11	0.05	0.05	0.05	0.58	pol (2nd)	0.33
	chr1_118502383_T_C	0.11	0.00	0.11	0.03	0.37	0.12	0.00	0.00	0.21	0.10	0.00	0.10	0.57	pol (2nd)	0.33
	chr1_132950294_A_C	**0.81**	**0.83**	0.53	0.75	0.29	0.54	**0.73**	**0.89**	**0.71**	**0.70**	**1.00**	**0.75**	0.57	exponential	0.33
	chr5_39594296_T_C	0.01	0.03	0.13	0.09	0.37	0.13	0.13	0.03	0.13	0.35	0.05	0.00	0.57	pol (2nd)	0.33
	chr14_2632555_C_T	0.00	0.00	0.03	0.03	0.24	0.11	0.00	0.03	0.08	0.00	0.00	0.00	0.57	pol (2nd)	0.32
	chr14_6200065_A_T	0.06	0.13	0.11	0.13	0.47	0.18	0.13	0.08	0.11	0.15	0.10	0.20	0.57	pol (2nd)	0.32
	chr14_6200159_T_C	0.06	0.13	0.11	0.13	0.47	0.13	0.13	0.08	0.11	0.15	0.10	0.20	0.57	pol (2nd)	0.32
	chr7_30967748_A_G	0.00	0.08	0.03	0.00	0.29	0.20	0.05	0.00	0.08	0.05	0.00	0.00	0.57	pol (2nd)	0.32
	chr1_23140159_T_C	0.06	0.03	0.08	0.03	0.29	0.04	0.00	0.00	0.11	0.00	0.00	0.05	0.55	pol (2nd)	0.31
	chr6_28550216_G_A	0.04	0.10	0.16	0.13	0.34	0.18	0.03	0.08	0.16	0.00	0.00	0.05	0.55	pol (2nd)	0.31
	chr9_17960492_C_T	0.44	0.63	0.68	0.66	0.18	0.51	**0.70**	0.64	0.42	0.60	**0.75**	0.55	0.55	pol (2nd)	0.31
	chr3_38560061_T_C	0.03	0.05	0.13	0.06	0.32	0.12	0.08	0.03	0.05	0.20	0.00	0.00	0.55	pol (2nd)	0.31
	chr14_6200496_G_A	0.06	0.13	0.18	0.16	0.50	0.15	0.15	0.17	0.11	0.20	0.10	0.20	0.55	pol (2nd)	0.30
	chr1_23065714_A_G	0.06	0.03	0.13	0.09	0.34	0.04	0.00	0.06	0.08	0.05	0.05	0.15	0.55	pol (2nd)	0.30
	chr2_126471892_A_G	0.00	0.00	0.03	0.06	0.29	0.07	0.05	0.06	0.05	0.00	0.20	0.10	0.55	pol (2nd)	0.30
Outlier SNPs LandUse	Mean value (LandUse)	32.85	11.6	12.18	38.3	25	32.24	37.52	1.33	38.19	29.19	38.13	34.22			
	chr17_7056075_A_C	0.10	0.28	0.05	0.03	0.26	0.07	0.08	0.31	0.03	0.00	0.05	0.00	0.73	lineal	0.31
	chr17_7239853_G_A	0.51	0.33	0.21	0.44	0.39	0.38	0.38	0.25	0.68	0.65	0.50	0.60	0.70	lineal	0.30
	chr2_31725655_T_C	0.17	0.43	0.39	0.06	0.16	0.21	0.05	0.39	0.16	0.55	0.05	0.15	0.54	pol (2nd)	0.30
	chr2_31712297_A_G	0.19	0.45	0.39	0.06	0.24	0.28	0.03	0.31	0.13	0.30	0.05	0.05	0.53	pol (2nd)	0.28
	chr2_79029493_G_T	0.18	0.30	0.08	0.13	0.05	0.16	0.10	0.64	0.11	0.30	0.25	0.25	0.53	logarithmic	0.28

Table 5S.11. Enrichment analysis for genes overlapping selected RDA outliers (RDA correlation ≥ 0.3 -0.28).

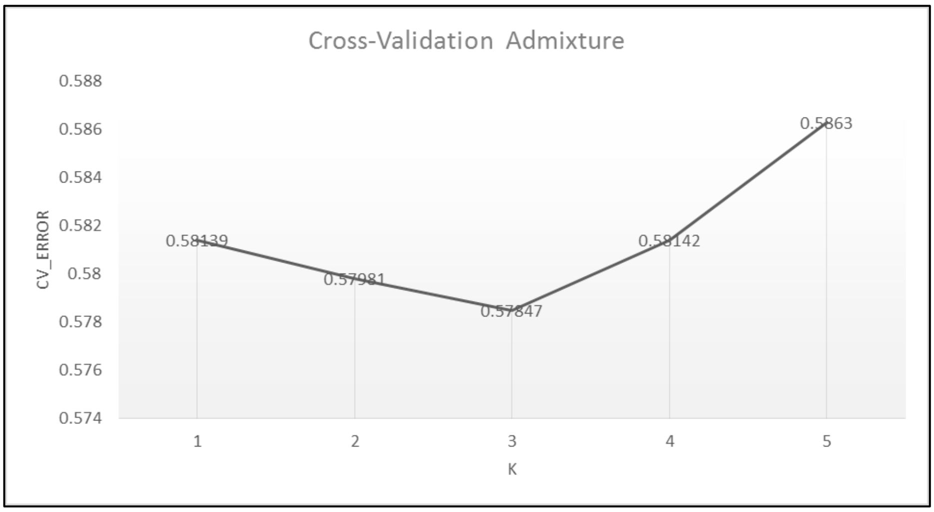
	#Term	**ID**	**Input number**	**Background number**	**P-Value**	**Corrected P-Value**	**Input**
minTemp bio6	Insulin signalling pathway	gga04910	2	121	0.007905505	0.044152659	*SREBF1|MTOR*
	Herpes simplex virus 1 infection	gga05168	2	150	0.011860285	0.044152659	*MAVS|MTOR*
	Pantothenate and CoA biosynthesis	gga00770	1	15	0.017303699	0.04905688	*PANK2*
	Glycosaminoglycan biosynthesis - heparan sulphate / heparin	gga00534	1	24	0.02691013	0.058719679	*HS6ST2*
	Folate biosynthesis	gga00790	1	28	0.03115061	0.060754737	*SPR*
	Autophagy - other	gga04136	1	28	0.03115061	0.060754737	*MTOR*
	Proteasome	gga03050	1	35	0.038528733	0.065809948	*PSMD3*
	Steroid hormone biosynthesis	gga00140	1	37	0.040626825	0.066900461	*CYP19A1*
	Cytosolic DNA-sensing pathway	gga04623	1	44	0.047935536	0.07405771	*MAVS*
	Metabolic pathways	gga01100	4	1290	0.049727782	0.074749014	*SPR|PANK2|AGK|CYP19A1*
precSeasonality bio15	Wnt signalling pathway	gga04310	2	141	0.020706731	0.081166572	*CTBP1|LRP6*
	Pantothenate and CoA biosynthesis	gga00770	1	15	0.02452847	0.082076033	*PANK2*
	One carbon pool by folate	gga00670	1	17	0.027553059	0.085611291	*MTHFS*
	Citrate cycle (TCA cycle)	gga00020	1	28	0.044025899	0.098909842	*DLST*
precWQ bio16	Retinol metabolism	gga00830	2	39	0.002601209	0.081382196	*LRAT|RDH10*
	Adrenergic signalling in cardiomyocytes	gga04261	2	128	0.023985025	0.083437269	*CACNA2D3|PPP2R5C*
	Pantothenate and CoA biosynthesis	gga00770	1	15	0.029096651	0.084168328	*PANK2*
	One carbon pool by folate	gga00670	1	17	0.03267504	0.084723484	*MTHFS*
	Glycosaminoglycan biosynthesis - heparan sulphate / heparin	gga00534	1	24	0.045098193	0.09514264	*HS6ST2*
precDQ bio17	Focal adhesion	gga04510	3	189	0.003785831	0.05314209	*PTEN|VAV2|CRKL*
	Adipocytokine signalling pathway	gga04920	2	64	0.005224917	0.05314209	*MTOR|RXRA*
	ErbB signalling pathway	gga04012	2	78	0.007587442	0.05314209	*MTOR|CRKL*
	Autophagy - animal	gga04140	2	117	0.016197992	0.05314209	*PTEN|MTOR*
	Insulin signalling pathway	gga04910	2	121	0.017239395	0.05314209	*MTOR|CRKL*
	mTOR signalling pathway	gga04150	2	140	0.022560194	0.055918007	*PTEN|MTOR*
	Cellular senescence	gga04218	2	143	0.02345512	0.057330648	*PTEN|MTOR*
	Arginine biosynthesis	gga00220	1	17	0.029019397	0.061769859	*ASS1*
	Regulation of actin cytoskeleton	gga04810	2	190	0.039272502	0.07283138	*VAV2|CRKL*
	Autophagy - other	gga04136	1	28	0.046347778	0.077885176	*MTOR*
LandUse	Synthesis and degradation of ketone bodies	gga00072	1	10	0.019190239	0.097854102	*HMGCL*
	Steroid biosynthesis	gga00100	1	17	0.031214529	0.099734228	*CEL*
	Butanoate metabolism	gga00650	1	24	0.043094875	0.11069468	*HMGCL*
	Folate biosynthesis	gga00790	1	28	0.049819664	0.113502191	*SPR*

Supplementary Figures

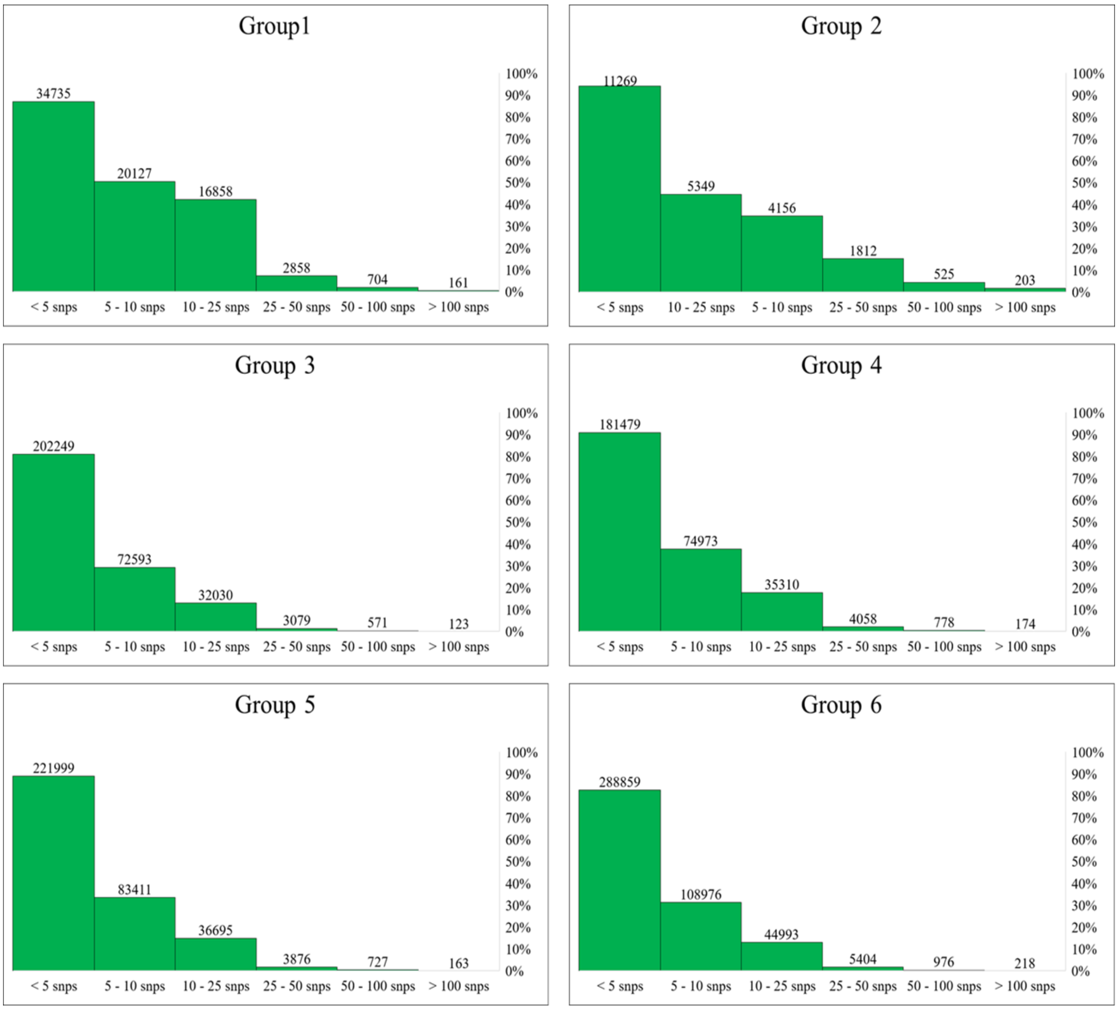
Chapter 2

Genetic diversity and population structure of indigenous Ethiopian village chicken

Supplementary Figure S2.1. Cross-validation showing K=3 as the best value explaining ancestry for the 25 Ethiopian chicken population investigated.



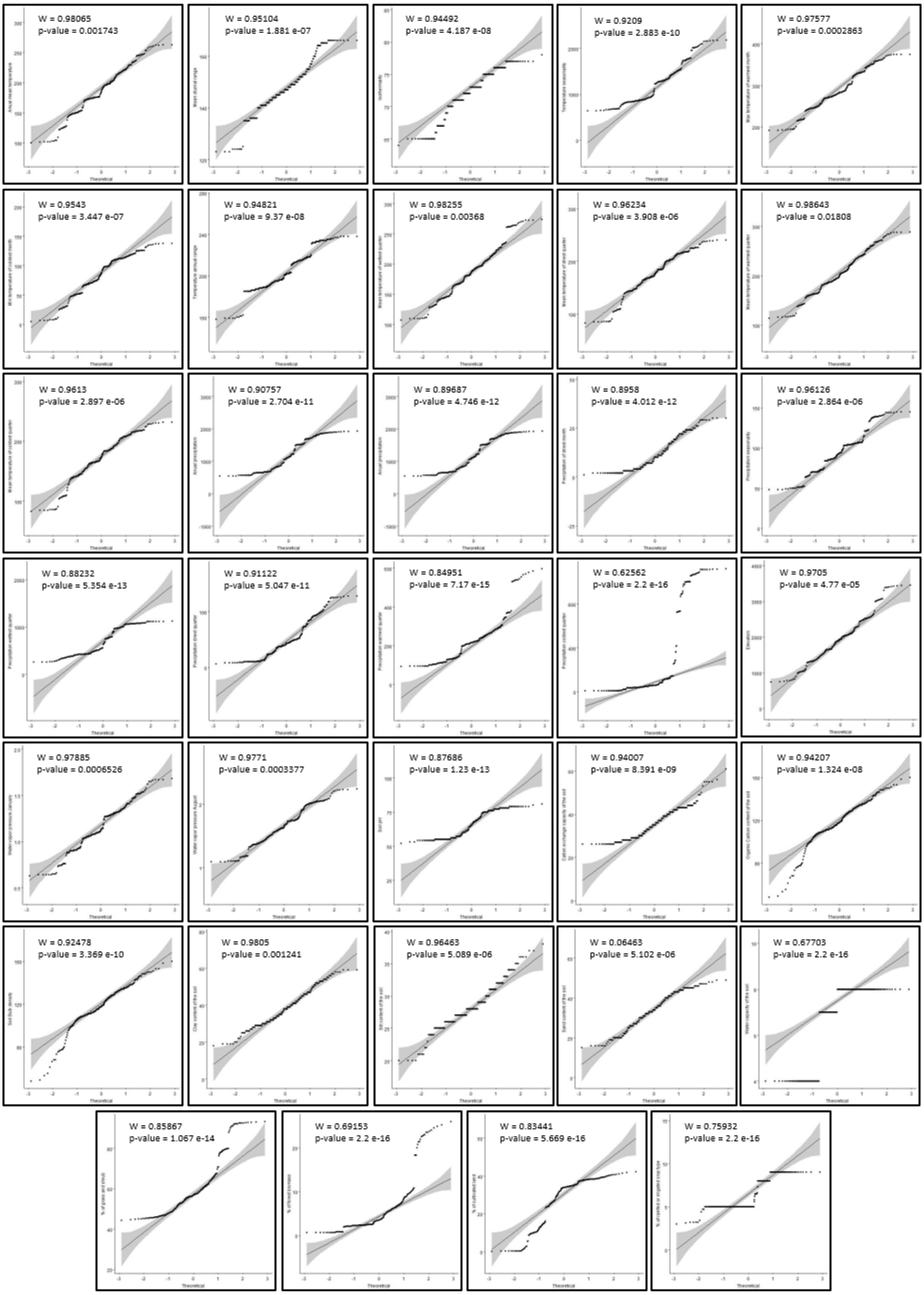
Supplementary Figure S2.2. Distribution of haplotype blocks (LD blocks) based on the number of SNPs per block in 6 population groups.



Chapter 3

Ecological niche modelling for environmental characterisation of indigenous Ethiopian village chickens

Figure S3.1. Shapiro-Wilk normality test of the agro-ecological studied variables



Chapter 4

Whole-genome sequences reveal the drivers of environmental adaptation in Ethiopian village chicken

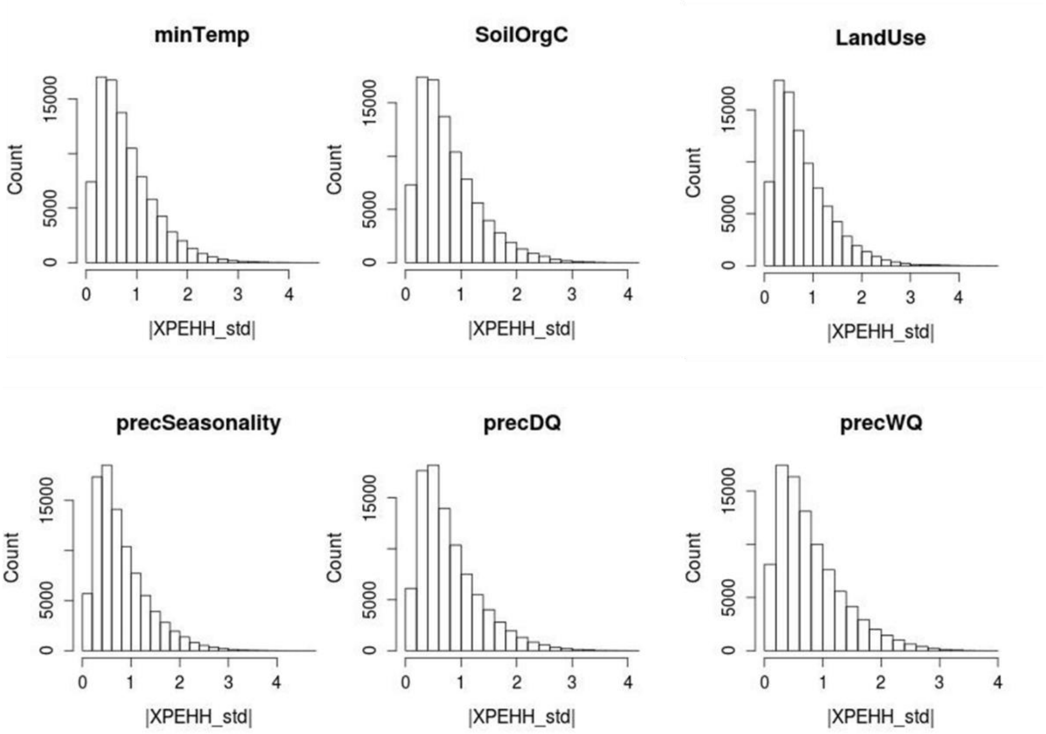


Figure S4.1. Histograms of |*XP-EHH*| values in windows in different analyses

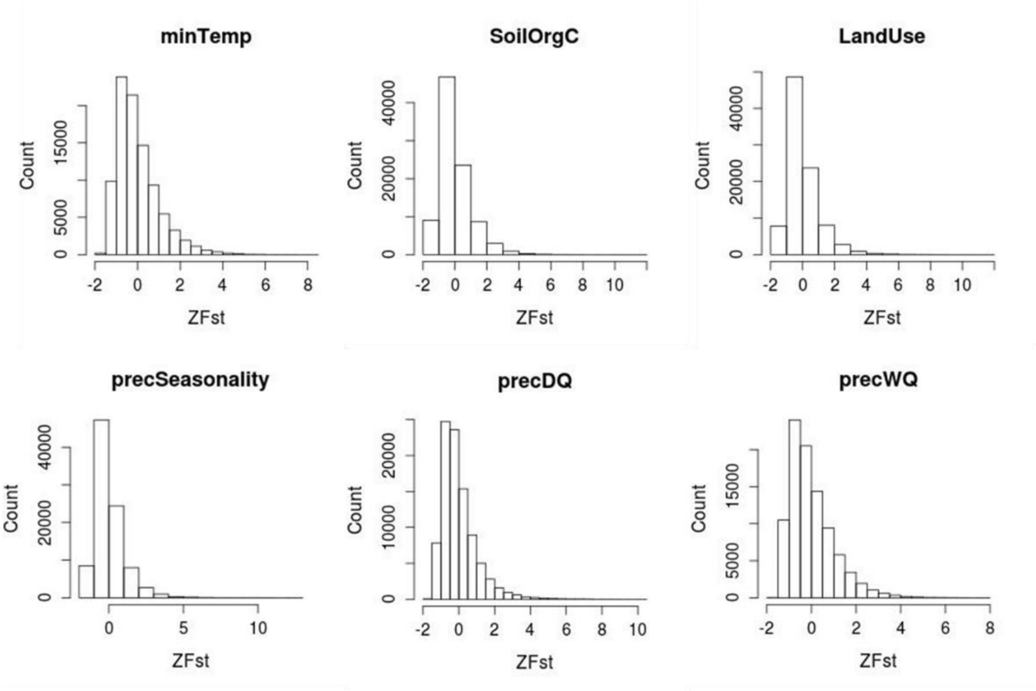


Figure S4.2. Histograms of Z*Fst* values in windows in different analyses

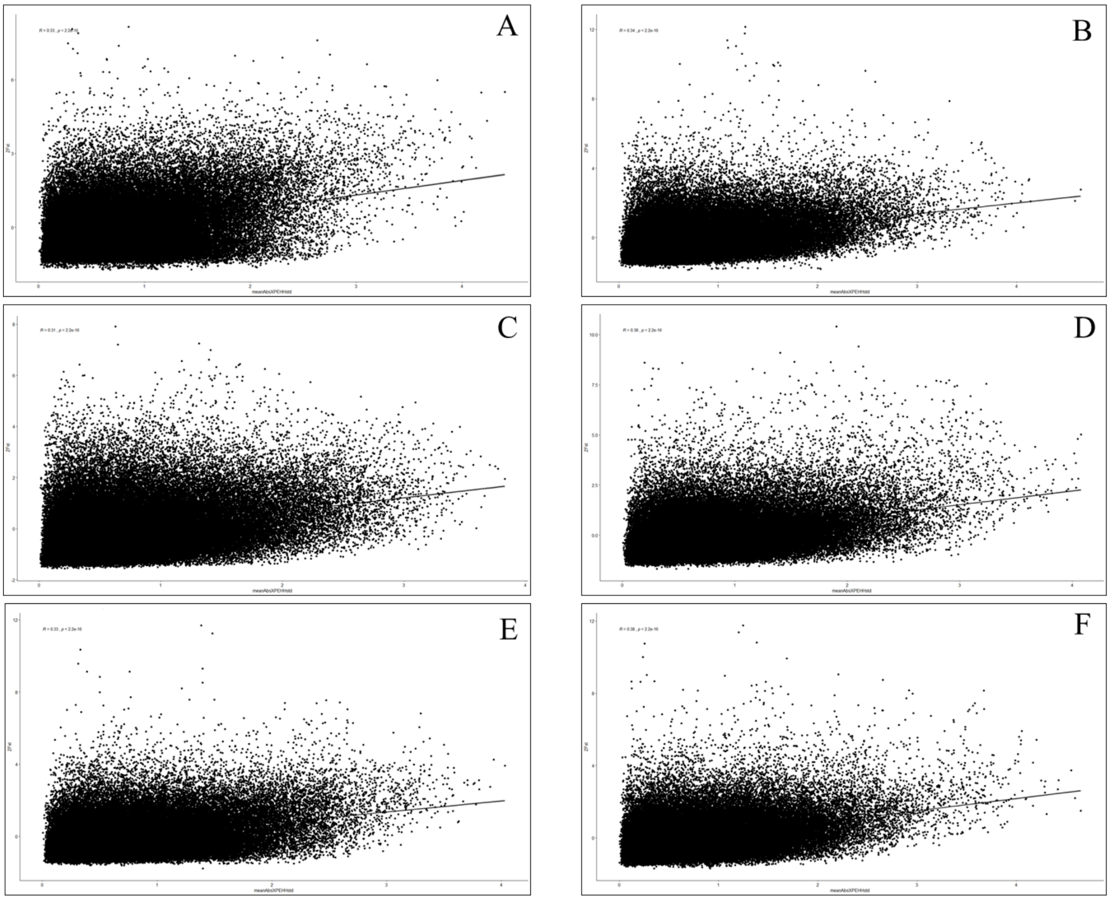


Figure S4.3. Correlation between (Z*Fst*) and (|*XP-EHH*|) signature of selection results across analyzed environmental population groups. A: minTemp, $r_s = 0.33$, $P < 2.2e\text{-}16$; B: precSeasonality, $r_s = 0.34$, $P < 2.2e\text{-}16$; C: precWQ, $r_s = 0.31$, $P < 2.2e\text{-}16$; D: precDQ, $r_s = 0.38$, $P < 2.2e\text{-}16$; E: SoilOrgC, $r_s = 0.33$, $P < 2.2e\text{-}16$, and F: LandUse, $r_s = 0.38$, $P < 2.2e\text{-}16$).

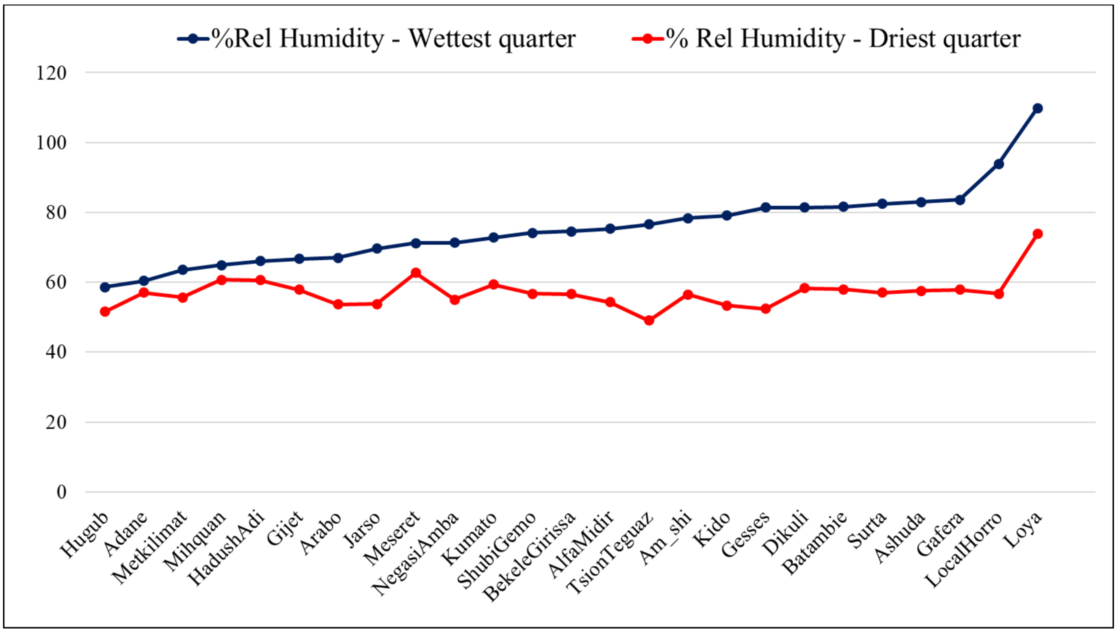


Figure S4.4. Mean percentage relative humidity in different populations in the wettest and driest quarter.

Chapter 5

Dissecting genome-environment associations in Ethiopian chicken ecotypes

Figure S5.1a Histograms of *ZHp* values in windows in different ecotypes

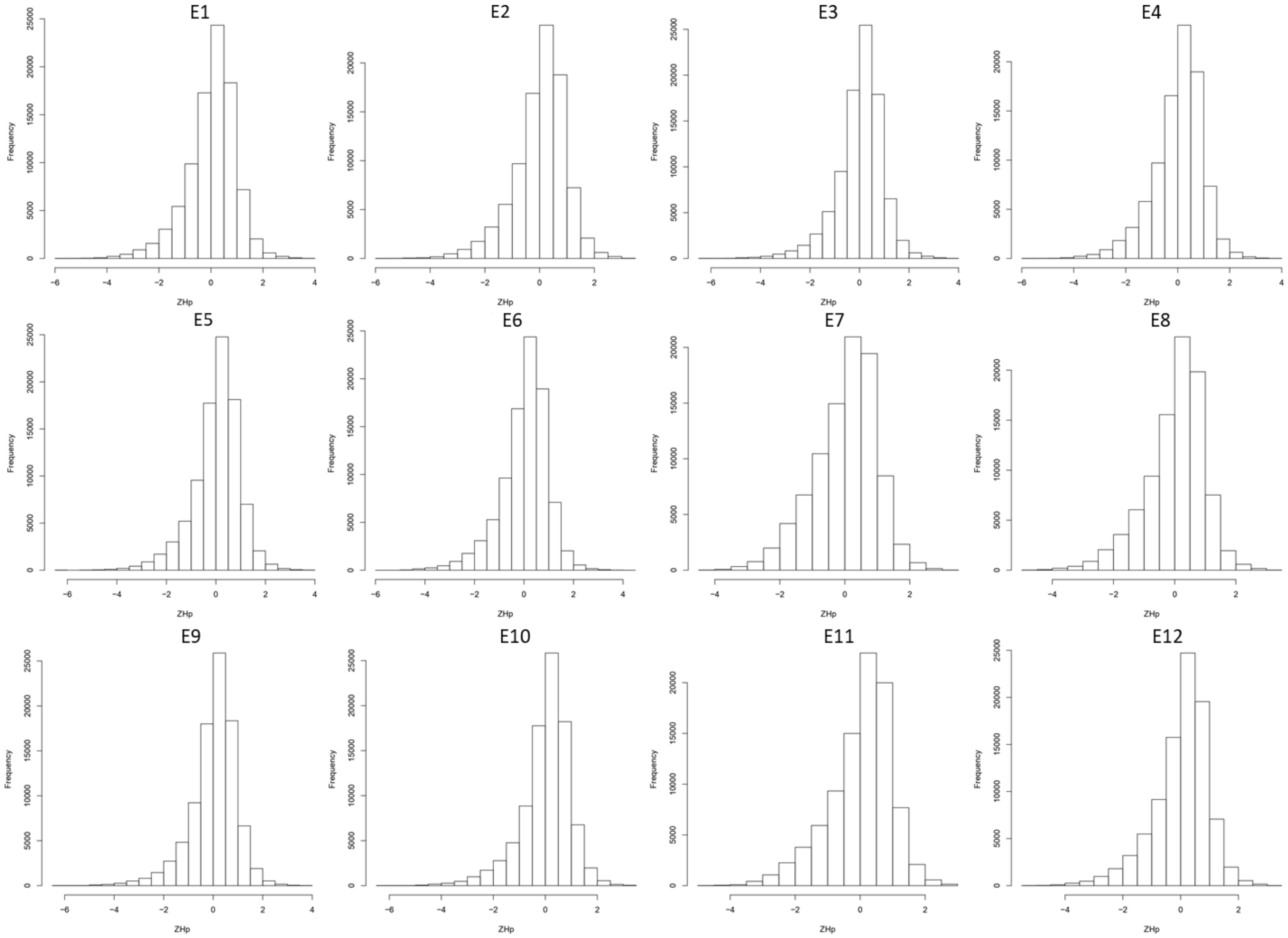


Figure S5.1b Histograms of *iHS* values in windows in different ecotypes

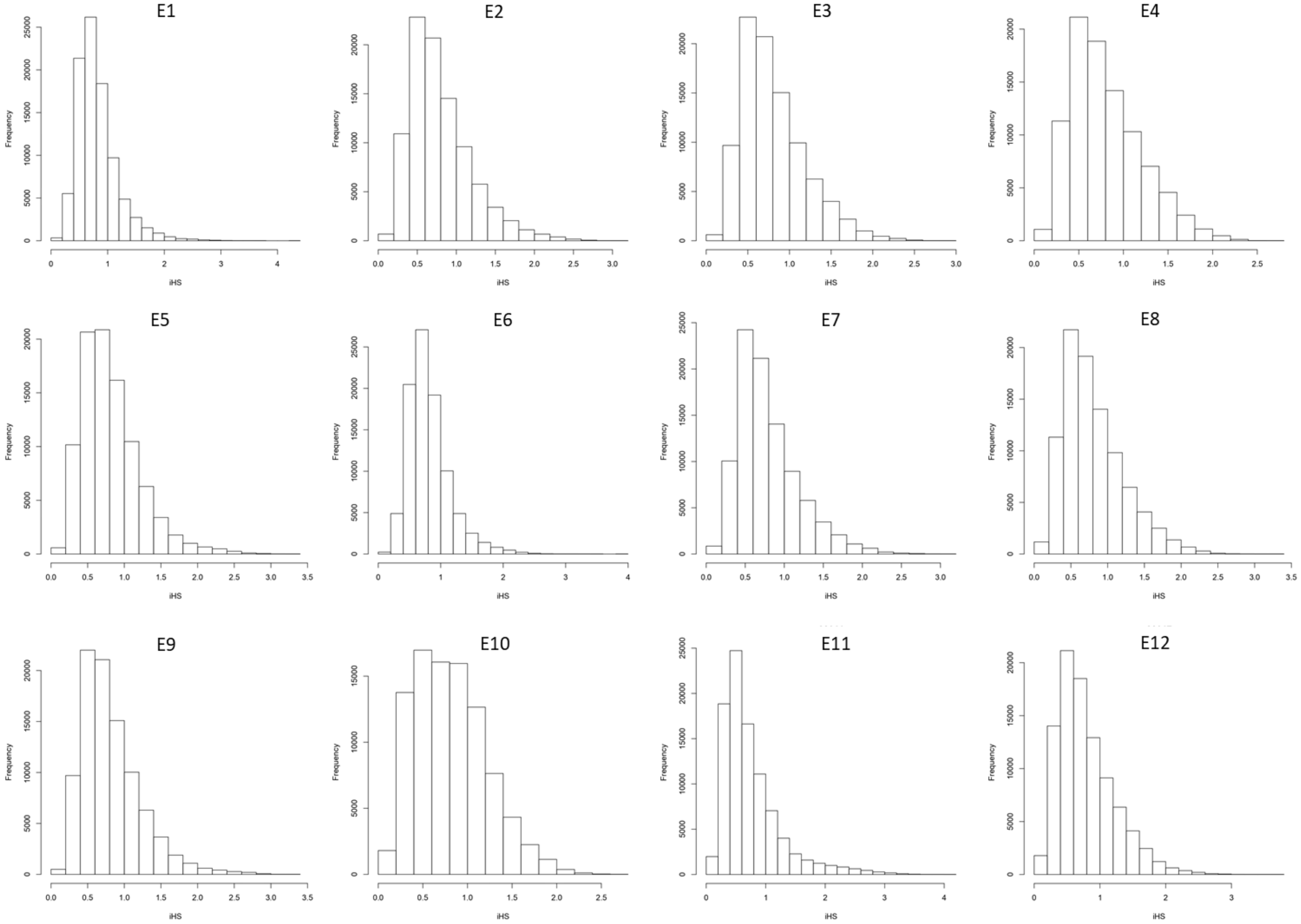


Figure S5.2 Correlation between *Hp* and mean unstandardized *iHS* values by ecotype

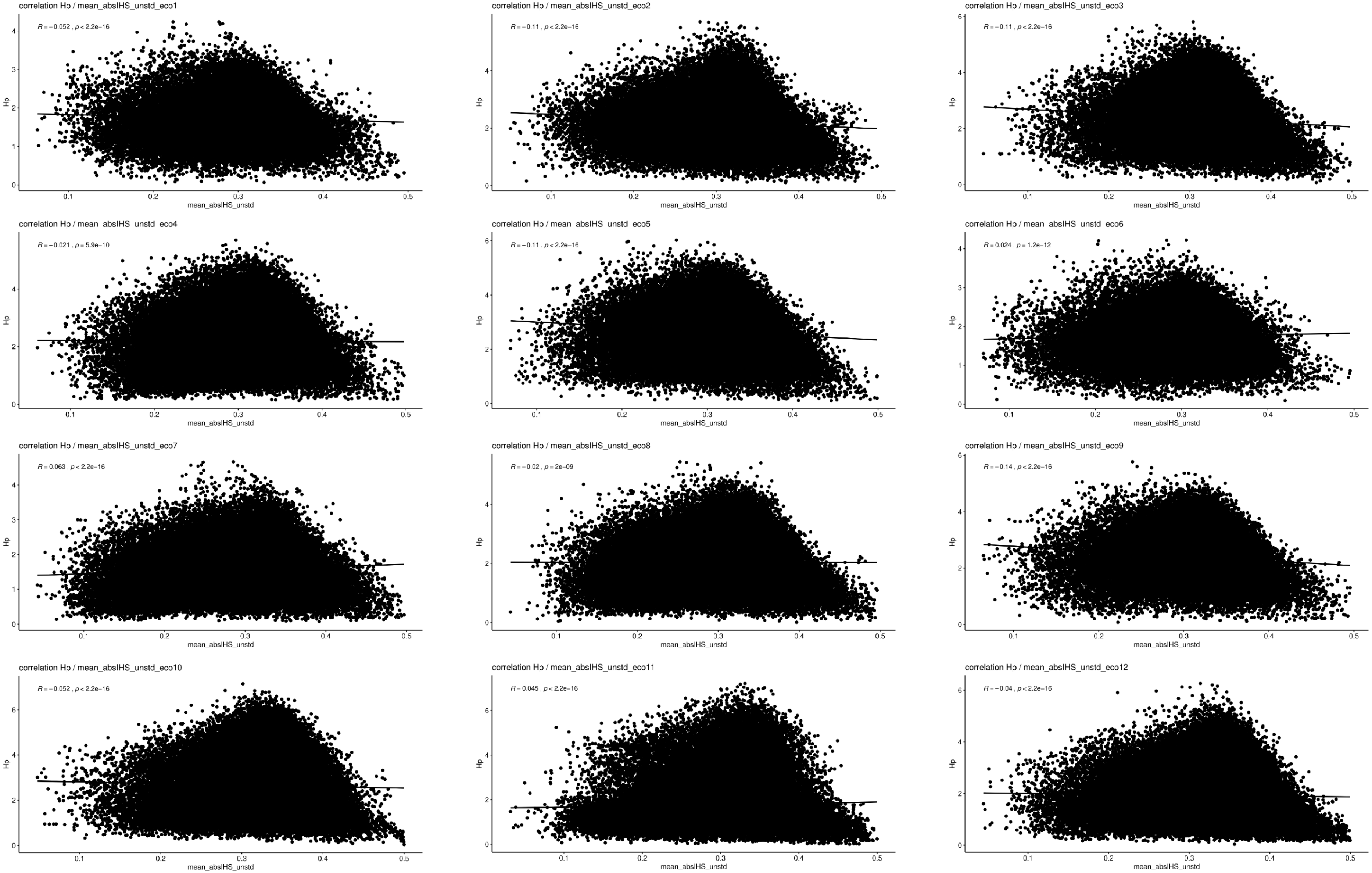


Figure S5.3 Correlation between *Z(Hp)* and mean standardized *iHS* values by ecotype