PHENOMIC and GENOMIC LANDSCAPE of ETHIOPIAN VILLAGE CHICKENS

By

Takele Taye Desta

BSc (Animal Sciences), MSc (Animal Breeding and Genetics)

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

May 2015

Ecology and Evolution Group

School of Life Sciences

Declaration

I, Takele Taye Desta, hereby declare that, this thesis and the work presented in it is my own and has been generated by me from my own original research, except the commercial chicken populations, and village chickens sampled outside of Ethiopia, pheasant, waterfowl and the junglefowl species SNP data obtained from Prof. David Burt and Dr David Wragg, which has been duly acknowledged. Except the use of references from other individuals' work, which I have duly acknowledged, this thesis in full or in part has not been presented elsewhere for another degree.

Dedication

This thesis is dedicated to the memory of my late father Taye Desta.

Acknowledgements

First of all, I would like to thank THE HIGHERMOST GOD for providing me with courage and stamina. Foremost, I would like to express my deepest thanks to my supervisor Prof. Olivier Hanotte for his unreserved support, encouragement and advice throughout the course of this study. Thanks Dr David Wragg for mentoring me on wet lab activities and for your help in data analysis and for allowing me to use your SNP data from village chickens sampled outside of Ethiopia, pheasant, waterfowl and the junglefowls. Dr Joram Mwashigadi Mwacharo thanks for your insightful thought and critical comment and for developing the blood sample collection protocol. I would like to acknowledge Mario Barbato for handling the effective population size analysis and Hussain Bahbahani for helping me in selective sweep analysis. I would like to thank Prof. David Burt for allowing me to use his commercial chickens' SNP data.

I would like to acknowledge, BBSRC (Biotechnology and Biological Sciences Research Council), DFID (Department for International Development) and the Scottish Government who financed this project. My appreciation also extends to Prof. Pete Kaiser who generously covered most of the genotyping cost of the Ethiopian chickens. Thanks Dr Rob Christley for your excellent coordination that brings my PhD project to this level. Dr Paul Wigley thanks a lot for your fruitful discussion; your music is an additional flavour. I would like to acknowledge The University of Nottingham for hosting my study and for awarding a scholarship.

Chicken Health for Development (CH4D) project core field team members in the UK and Ethiopia: Dr Judy Bettridge, Dr Stacey Lynch, Kassech Melese, Marisol Collins, Dr Zelalem G. Terfa, Eshetu Zerihun, Michael Temesgen, Eskinder Aklilu, Berhanu Aberra, Alemayehu Amare, Jibril Yusuf, Berhanu Bersissa and Teshome Takele thank you all for your help during field sampling. Livestock production officers in Horro and Jarso district Office of Agriculture and agricultural development agents working in each village facilitate the on spot field sampling. Dr Tadelle Dessie and Rahel Misganaw are duly acknowledged for managing logistic stuffs and for handling administrative works associated with the field sampling.

All visited households in Horro and Jarso were kindly volunteered to sample their flocks. My family in Ethiopia God bless you all for your incredible moral support and for your dedicated prayer.

Acronyms

AMOVA	Analysis of molecular variance
ANOVA	Analysis of variance
AR	Allelic richness
CA	Correspondence analysis
DAPC	Discriminant analysis of principal components
DAVID	Database for Annotation Visualization and Integrated Discovery
DENSITY	Number of SNPs forming 1Kb ROH segment
DNA	Deoxyribose Nucleic Acid
FAO	Food and Agriculture Organization of the United Nations
GGA	Gallus gallus
GWAS	Genome-wide association study
HWE	Hardy Weinberg equilibrium
IBD	Identity by descent
IBS	Identity by state
iHH	Integrated haplotype homozygosity
iHS	Integrated haplotype score
Kb	Kilo base pair
KB_{TOTAL}	Physical size of ROH segments in Kb
KEGG	Kyoto Encyclopedia of Genes and Genomes
km	kilometre
LD	Linkage disequilibrium
MAF	Minor allele frequency
masl	Meter above sea level
Mb	Mega base pair
NCBI	National Centre for Biotechnology Information
Ne	Effective population size
ng	Nano gram
NJ	Neighbour joining
NSEG	Number of ROH segments
NSIM	Number of overlapping ROH segments
NSNP	Number of SNPs forming ROH segment

PC	Principal component
PCA	Principal component analysis
QC	Quality control
QTL	Quantitative trait loci
ROH	Runs of homozygosity
SNP	Single nucleotide polymorphism
UPD	Uniparental disomy
µl	Micro litre

Table of Contents

List of Figures

List of Tables

Summary

This study involves two village chicken populations sampled from Horro and Jarso regions of Western and Eastern Ethiopia respectively. This study maps the phenomic and genomic landscape of the two chicken populations using morphological markers and a high density (600K) SNP array. Although the two chicken populations tend to display nondescript morphological characteristics, they show a subtle variation except for rare morph variants that have been in most instances scored on Jarso chickens. Morphological analysis uncovers a vast array of intrapopulation variation. Genetic diversity and population structure analyses assign the two chicken populations to two distinct genepools representing their population of origin. A high intrapopulation genetic diversity is uncovered, which shows a broad genetic base (high genetic diversity) of the two chicken populations. We hypothesized that a clearly evident genetic divergence observed between the two chicken populations may be attributed to difference in demographic history, origin (routes of introduction to Africa), breeding history of the two chicken populations and demographic structure of subsistence farmers. Absence of gene flow owing to their distant geographic location and ecological variation may have also contributed to this divergence. A population structure analysis performed on a random subset of the two Ethiopian chicken populations along with village chickens sampled from other African countries, Asia and Latin America, commercial populations and the junglefowl species reveals a unique genetic structure of Ethiopian chickens, which implicates the need for further study on the genetic landscape of the latter. To infer the extent of inbreeding we performed a run of homozygosity analysis (ROH). Our analysis indicates that ROH is more intense in Jarso than Horro chickens and in macrochromosomes than microchromosomes. The extensive ROH mapped in some chickens implicates the need to restructure the existing traditional breeding practice of subsistence farmers. Our analysis confirms the commonness of ROH in genic regions. For the first time, we detect twenty three putative uniparental disomy in twenty two Ethiopian village chickens. Signature of selection analysis detects divergently selected genomic regions in the two chicken populations indicating a considerable divergent selection imposed on the two populations. Genes involving in melanogenesis pathway are among those subjected to a divergent selection. However, some overlapping regions were also mapped in the two chicken

populations implicating the ubiquitous impact of natural selection on genes regulating vital biological processes. A genome-wide association study performed on pigmentation (earlobe, plumage and shank) traits and variants of crest, comb and a lightly feathered shank maps a number of putative loci that may underlie variations in these traits. Our GWAS analysis on pigmentation traits produced a long list of loci than that have been known to involve in the genetic control of pigmentation in the chicken, with most of these have been mapped in the mouse. We also refined further the causative variants underlying a lightly feathered shank mutation. Our GWAS analysis map a number of putative novel loci that may underlie the genetic control of the traits analysed and this has laid a foundation for subsequent work that would involve targeted sequencing and a candidate gene approach. This study is the first of its kind in Africa that uses a large number of samples and a high density SNP array to unlock phenomic and genomic landscape of the true type village chickens.

Chapter 1

General Background

The junglefowls

Chickens are among some of domesticates survived by an extant progenitor. The chicken is believed to be domesticated ~ ten thousand years ago from the junglefowl species native to the Asian continent (see Xiang et al., 2014 for the latest update). The junglefowls are subdivided into four species based on their morphological characteristics and home range (Delacour, 1977). The four junglefowl species are the red junglefowl (*Gallus gallus*) native to southern Asia (the main home range is southeast Asia); the grey junglefowl (*Gallus sonneratii*) native to southern India; the green junglefowl (*Gallus varius*) native to Indonesia and the Ceylon junglefowl (*Gallus lafayetii*) native to Sri Lanka. The red junglefowl is well known for its wide range of geographical distribution (Crawford, 1990), that extends from western India to South-East Asia, the Philippines and as far as the Pacific islands (Peterson and Brisbin 1999). Photographs displaying hen and cock of the four junglefowl species are presented in Figure1.1 – 1.4. The home range of the junglefowl species includes the south and the south-eastern Asia – the putative centre(s) for domestication of chicken (Figure 1.5). Although the junglefowl is subdivided into four species, there is no variation in their karyotype (Okamoto et al., 1988). However, the green junglefowl is thought to be distantly related to the remaining three species. Unlike the other three species, the comb of the green junglefowl is non-serrated. The green junglefowl has a single median wattle and it has sixteen tail feathers instead of fourteen. The green junglefowl is also characterized by absence of eclipse plumage (Crawford, 1990).

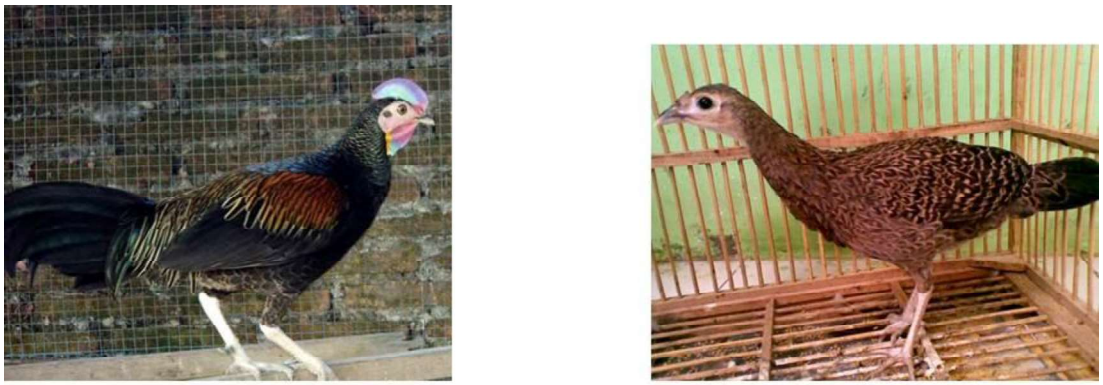
Figure 1.1 The red junglefowl cock and hens in Kaziranga National Park, Assam, India.

(i) (ii)

Figure 1.2 The grey junglefowl: (i) cock from Bandipur National Park and (ii) hen from Thattekad Bird Sanctuary (Source:http://en.wikipedia.org/wiki/Grey_junglefowl).

(i) (ii)

Figure 1.3 The Ceylon junglefowl: (i) cock from Sinharaja Forest Reserve, Sri Lanka and (ii) hen from Sinharaja Forest Reserve, Sri Lanka (Source: http://en.wikipedia.org/wiki/Sri_Lankan_junglefowl)



(i) (ii)

Figure 1.4 The green junglefowl: (i) cock and (ii) hen from Sinharaja Forest Reserve, Sri Lanka (Source: http://www.cemanifarms.com/p/green-jungle-fowl.html).

The red junglefowl is further subdivided into five subspecies (*Gallus gallus bankiva* (Java), *G. g. gallus* (Indochina), *G.g. jabouillei* (Vietnam), *G. g. murghi* (India), and *G. g. spadiceus* (Burma) based on home range, size of comb, earlobe, hackle feather and wattle and colour of earlobe and hackle feather (Crawford, 1990; Romanov et al., 2009).

Figure 1.5 The geographical distribution of the junglefowl in India, Tibet, Mongolia, China and Philippines (http://www.rareprintsgallery.com/store/product/beebe029).

Origin of the domestic fowl

There is active debate on the origin of domestic chickens between monophyletists and polyphyletists, though a third moderate group suggests the red junglefowl as the major contributor and the remaining three species (especially the grey junglefowl) as minor contributor(s) to the domestic chicken genepool (see Romanov et al., 2009 for a review). Evidence from experimental crossings and morphological studies (Darwin, 1868; Steiner, 1945; Danforth, 1958; Morejohn, 1968) and genetic studies (Fumihito et al., 1994 & 1996; Eriksson et al., 2008) often report conflicting findings. However, there is some evidence that supports the polyphyletic origin of the domestic fowl, though the contribution of the red junglefowl to the domestic chickens' genepool is thought to be proportionately high. For example, loci that underlie silver plumage (Stevens, 1991) and yellow skin (Eriksson et al., 2008) in domestic chickens are peculiar to the grey junglefowl and the locus that confers

extended black plumage is peculiar to the green junglefowl (Stevens, 1991). The spotted comb colour observed in the Kenyan village chickens (Kingori et al., 2010) may genetically associate with a multi-coloured comb variant of the green junglefowl. However, Romanov et al. (2009) noted a scarcity of reliable evidence from the polyphyletic origin school of thought, who argue to the extent of extinction of true progenitor of the domestic fowl.

Multiple maternal origins of the domestic fowl has been supported by a large number of studies (e.g. Nishibori et al., 2005; Liu et al., 2006; Oka et al., 2007; Kanginakudru et al., 2008; Storey et al., 2012; Miao et al., 2012; Xiang et al., 2014). Archaeological evidence also corroborates multiple centres of domestication (see Blench & Macdonald 2000 and the references therein). Dispersion of domestic chickens from their putative centres of domestication most likely associates with migration and trade routes of a prehistoric man (Blench & Macdonald, 2000; Mwacharo et al., 2013). Presumably, chickens were introduced to Africa continent through two main routes – the north east Africa and the Red Sea coast, which supports two main maternal lineages found from mitochondrial DNA analysis of African native chickens (*e.g.* Muchadeyi et al., 2008; Mwacharo et al., 2011). Domestic chickens may have also been introduced to Ethiopia via these two entry points.

The chicken karyotype

The chicken karyotype consists of 39 pairs of chromosomes (Yamashina, 1944) as it is schematically displayed in Figure 1.6. Unlike mammals, the chicken karyotype shows a wide variation in physical size of the autosomes (Hillier et al., 2004; Burt, 2005). Based on their physical size, chickens' autosomes have been classified into eight pairs of cytogenetically distinct macrochromosomes and thirty pairs of cytogenetically indistinguishable microchromosomes (Emara and Kim, 2003). However, Hillier et al. (2004) based on their size they classified chicken autosomes into three broad classes: five macrochromosomes (GGA1–5), five intermediate chromosomes (GGA6–10) and twenty eight microchromosomes (GGA11–38). The chicken macrochromosomes are ~ equal to an average-sized mammalian chromosome (~140Mb), while the size of the microchromosomes ranges from 2 to

15Mb (Hillier et al., 2004). The microchromosomes make 18% of the chicken genome (Axelsson et al., 2005); however, they harbour 31% of the genes (Hillier et al., 2004). Unlike mammals, in the chicken the sex chromosomes are designated as Z and W. In the chicken, the hen is heterogametic (ZW), while the cock is homogametic (ZZ). Analogous to the X chromosome of the mammals, the Z chromosome is bigger than its W counterpart, and the Z chromosome contains a large number of genes than W (Hillier et al., 2004).

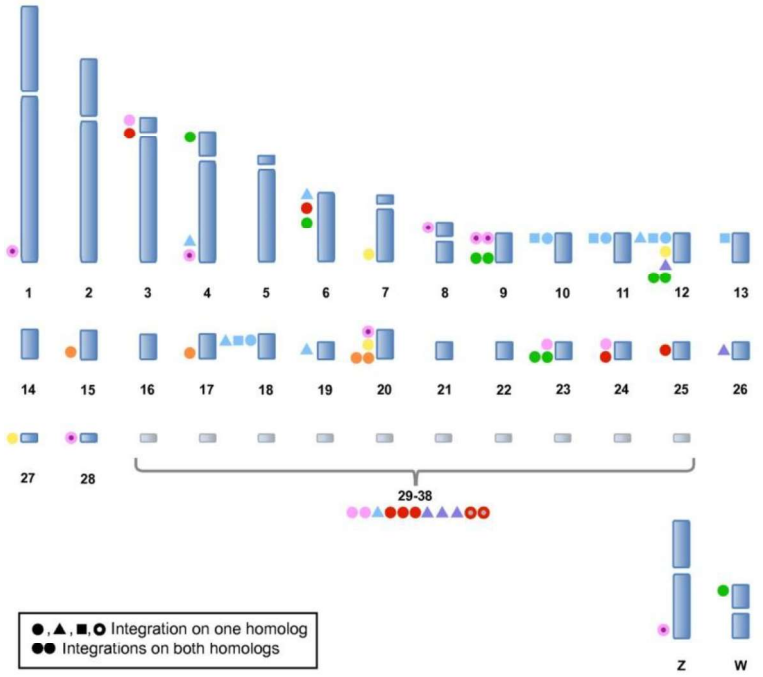


Figure 1.6 Schematic sketch of the chicken karyotype (Source: Robinson et al. Herpesviridae 2010 1:5 doi: 10.1186/2042-4280-1-5).

Genomic organization of the chicken karyotype

The chicken is the first livestock species to have had its genome of ~ 1.05 X 10^9 base pair sequenced. The chicken genome size is ~ one-third of the mammalian genome, which is attributable to substantial reduction in interspersed repeats, pseudo-genes and segmental duplications (Hillier et al., 2004). For example, interspersed repeats found ~ in 9% of the chicken genome only (Hillier et al., 2004). The microchromosomes have a higher GC content and higher density of CpG islands, genes and repeats than the macrochromosomes (Hillier et al., 2004). Moreover, the microchromosomes exhibit a high recombination rate, because each chromosome arm must have at least one crossover for meiotic segregation of chromosomes to occur (Hillier et al., 2004; Burt, 2005; Ellegren, 2005). The compact nature of microchromosomes is further evidenced by a positive correlation between chromosome size with both intron and junk DNA (Ellegren, 2005). A crossover rate of 2.8cM/Mb for the macrochromosomes and 6.4cM/Mb for the microchromosomes has been estimated for the chicken genome, which is greater than 1 to 2cM/Mb crossover observed in most of the human chromosomes (Burt, 2005). Due to a high recombination rate especially for traits encoded by microchromosomal genes, a high density genetic marker is required to map loci that underlie variation in trait of interest. Despite this, a high rate of recombination is important to reduce mapping region and to localize linkage disequilibrium; which then increases the resolution power of a fine mapping (Ellegren, 2005). This genomic landscape has made chicken an ideal species in genetic linkage studies (Burt, 2005).

The microchromosomes have 18% higher intronic sequence divergence and 26% higher rate of synonymous substitution in the coding sequences than the macrochromosomes. The microchromosomes are therefore more liable to mutation (Burt, 2005). Due to a high gene density, the microchromosomes have a higher frequency of genic SNPs per unit of a genomic segment. However, frequencies of SNPs and indels are independent of the chromosome size except chromosome 16 that harbours the hyper-variable gene family – the major histocompatibility complex (Hillier et al., 2004).

Village chickens

The village chickens have been evolved mainly under the impact of natural selection. Human driven selection is less intense and mating is commonly uncontrolled. Scavenging forms the main feed resource base and management is suboptimal. Village chickens display a spectacular morphological diversity (plumage colour, comb shape, shank colour, and earlobe colour etc., Figure 1.7). Family flock size is small and chicken rearing is subsistence oriented farming activity. Extensive gene flow occurs through local trading-networks and different forms of gift. Besides egg and meat production, village chickens are kept to fulfil a number of cultural, ritual and social roles. This study involves two village chicken landraces, Horro and Jarso sampled from western and eastern Ethiopia respectively. The map of the two study sites was generated using RgoogleMaps (Loecher 2011) and ggmap (Kahle & Wickham 2013) package for R and is presented in Figure 1.8. A study conducted on genetic diversity of five Ethiopian village chicken populations including Horro and Jarso using ten microsatellite markers reveals a considerable level of genetic divergence between Horro and Jarso chickens (Dessie, 2003). Moreover, a hyper variable region of the mitochondrial DNA analysis for a subset of our chicken populations (53 from Horro and 60 from Jarso) assigns D-haplogroup to Horro and A-haplogroup to Jarso, with a few chickens from Jarso (4/60) showing a D-haplogroup (Personal communication with Joram Mwacharo, 2013). Both studies invariably confirmed the genetic disconnectedness of Horro and Jarso chickens.

(i) (ii)

Figure 1.7 Village chicken (i) cock and (ii) (hen).

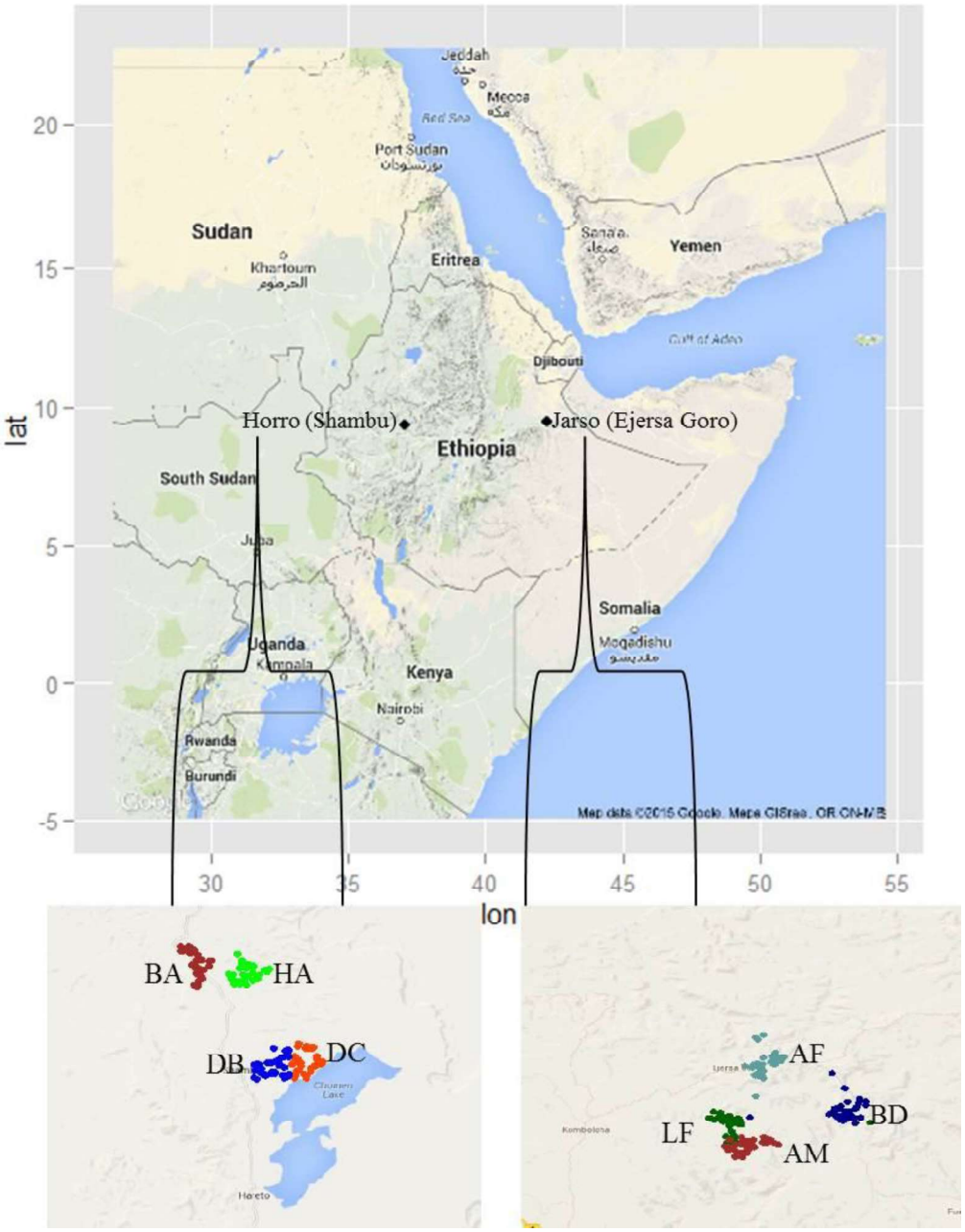


Figure 1.8 Map of the two study sites.

The genetic tool

Single nucleotide polymorphism (SNP) refers to a single base change between nucleotide sequences of homologous chromosomes. Single nucleotide polymorphism is one of the common genetic markers that have been increasingly used in genomic studies. In this study, we used a recently developed and commercially available high density Affymetrix® SNP chip (600K) for chickens (Kranis et al., 2013).

Organization of the thesis

The main aim of this study is to investigate phenomic and genomic landscape of two Ethiopian village chicken populations – Horro and Jarso. This study represents the genetic component of a wider program "Reducing the impact of diseases on village poultry in Ethiopia"– a project stemmed from Combating Infectious Diseases of Livestock for International Development (CIDLID) programme of the United Kingdom. The two chicken populations are selected based on the result of a previous study (Dessie, 2003) and their distant geographical location and contrasting production environments. This thesis makes nine chapters. The first chapter provides a brief general introduction as an entry point to the main body of the thesis. The second and the third chapters are dedicated to studies involving morphological markers. The fourth chapter deals in detail with genetic structure of the two Ethiopian chicken populations (Horro and Jarso), while chapter five compares a random subset of the two Ethiopian chicken populations with village chickens sampled from other African countries, Asia and Latin America, and commercial chickens and the four junglefowl species. Chapter six is dealing with runs of homozygosity analysis and chapter seven is dedicated to detection of selective sweep regions. Chapter eight is dealing with a genome-wide association study of threshold and Mendelian traits. Chapter nine presents a general discussion and concluding remarks. Chapter 4–7 each is accompanied by supplementary information provided in the Appendices section.

References

AXELSSON, E., WEBSTER, M. T., SMITH, N. G., BURT, D. W. & ELLEGREN, H. Comparison of the chicken and turkey genomes reveals a higher rate of nucleotide divergence on microchromosomes than macrochromosomes. Genome Research, 2005, 15(1), 120-125.

BLENCH, R. M. & MACDONALD, K. C. Chickens. In The Cambridge world history of food, Volume I, FK Kenneth & C.O. Kriemhild editions. Cambridge, Cambridge University Press, 2000, 496-499.

BURT, D. W. Chicken genome: Current status and future opportunities. Genome Research, 2005, 15, 692-1698.

CRAWFORD, R. D. Poultry Breeding and Genetics. Elsevier B.V., 1990.

DANFORTH, C. H. *Gallus sonnerati* and the domestic fowl. Journal of Heredity, 1958, 49, 167-169.

DARWIN, C. The variation of animals and plants under domestication. London: John Murray, 1868.

DELACOUR, J. The Pheasants of the World, 2nd Edition. Spur Publications, Hindhead, Surrey, 1977.

DESSIE, T. Phenotypic and genetic characterization of local chicken ecotypes in Ethiopia. A PhD thesis, Humboldt University of Berlin, Germany, 2003.

ELLEGREN, H. The avian genome uncovered. Trends in Ecology and Evolution, 2005, 20(4), 180-186.

EMARA, M. G., & KIM, H. Genetic markers and their application in poultry breeding. Poultry Science, 2003, 82(6), 952-957.

ERIKSSON, J., LARSON, G., GUNNARSSON, U., BED'HOM, B., TIXIER-BOICHARD, M., STRÖMSTEDT, L., ... & ANDERSSON, L. Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. PLoS Genetics, 2008, 4(2), e1000010.

FUMIHITO, A., MIYAKE, T., SUMI, S., TAKADA, M., OHNO, S. & KONDO, N. One subspecies of the red junglefowl (*Gallus gallus gallus*) suffices as the matriarchic ancestor of all domestic breeds. Proceedings of the National Academy of Sciences, 1994, 91(26), 12505-12509.

FUMIHITO, A., MIYAKE, T., TAKADA, M., SHINGU, R., ENDO, T., GOJOBORI, T., KONDO, N. & OHNO, S. Monophyletic origin and unique dispersal patterns of domestic fowls. Proceedings of the National Academy of Sciences, 1996, 93(13), 6792-6795.

HILLIER, L. W., MILLER, W., BIRNEY, E., WARREN, W., HARDISON, R. C., PONTING, C. P., ... & DODGSON, J. B. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature, 2004, 432(7018), 695-716.

KAHLE, D. & WICKHAM, H. ggmap: Spatial visualization with ggplot2. R Journal, 2013, 5(1).

KANGINAKUDRU, S., METTA, M., JAKATI, R. D., & NAGARAJU, J. Genetic evidence from Indian red jungle fowl corroborates multiple domestication of modern day chicken. BMC Evolutionary Biology, 2008, 8(1), 174.

KINGORI, A. M., WACHIRA, A. M. & TUITOEK, J. K. Indigenous chicken production in Kenya: a review. International Journal of Poultry Science, 2010, 9(4), 309-316.

KRANIS, A., GHEYAS, A. A., BOSCHIERO, C., TURNER, F., YU, L., SMITH, S., ... & BURT, D. W. Development of a high density 600K SNP genotyping array for chicken. BMC Genomics, 2013, 14(1), 59.

LIU, Y. P., WU, G. S., YAO, Y. G., MIAO, Y. W., LUIKART, G., BAIG, M., ... & ZHANG, Y. P. Multiple maternal origins of chickens: out of the Asian jungles. Molecular Phylogenetics and Evolution, 2006, 38(1), 12-19.

LOECHER, M. RgoogleMaps: Overlays on Google map tiles in R. R package version, 2011, 1(9.15).

MIAO, Y. W., PENG, M. S., WU, G. S., OUYANG, Y. N., YANG, Z. Y., YU, N., ... & ZHANG, Y. P. Chicken domestication: an updated perspective based on mitochondrial genomes. Heredity, 2012, 110(3), 277-282.

MOREJOHN, G. V. Breakdown of isolation mechanisms in two species of captive junglefowl. Evolution, 1968, 22, 576-582.

MUCHADEYI, F. C., EDING, H., SIMIANER, H., WOLLNY, C. B. A., GROENEVELD, E., & WEIGEND, S. Mitochondrial DNA D-loop sequences suggest a Southeast Asian and Indian origin of Zimbabwean village chickens. Animal Genetics, 2008, 39(6), 615-622.

MWACHARO, J. M., BJØRNSTAD, G., HAN, J. L., & HANOTTE, O. The history of African village chickens: an archaeological and molecular perspective. African Archaeological Review, 2013, 30(1), 97-114.

MWACHARO, J. M., BJØRNSTAD, G., MOBEGI, V., NOMURA, K., HANADA, H., AMANO, T., ... & HANOTTE, O. Mitochondrial DNA reveals multiple introductions of domestic chicken in East Africa. Molecular Phylogenetics and Evolution, 2011, 58(2), 374-382.

NISHIBORI, M., SHIMOGIRI, T., HAYASHI, T., & YASUE, H. Molecular evidence for hybridization of species in the genus Gallus except for *Gallus varius*. Animal Genetics, 2005, 36(5), 367-375.

OKA, T., INO, Y., NOMURA, K., KAWASHIMA, S., KUWAYAMA, T., HANADA, H., ... & AKISHINONOMIYA, F. Analysis of mtDNA sequences shows Japanese native chickens have multiple origins. Animal Genetics, 2007, 38(3), 287-293.

OKAMOTO, S., MAEDA, Y. & HASHIGUCHI, T. Analysis of the karyotypes of four species of jungle fowls. Japanese Journal of Zootechnical Science, 1988, 59 (2), 146-151.

PETERSON, A. T., & BRISBIN, I. L. Genetic endangerment of wild Red Junglefowl *Gallus gallus*? Bird Conservation International, 1999, 8(04), 387-394.

ROBINSON, C. M., HUNT, H. D., CHENG, H. H., & DELANY, M. E. Chromosomal integration of an avian oncogenic herpesvirus reveals telomeric preferences and evidence for lymphoma clonality. Herpesviridae, 2010, 1(1), 5.

ROMANOV, M. N., SAZANOV, A. A., MOISEYEVA, I. & SMIRNOV. A. F. Poultry. In Genome Mapping and Genomics in Domestic Animals (Vol. 3), Cockett, N. E., & Kole, C. (Eds). Springer, 2009, 75-141.

STEINER, H. *Ueber letal Fehentwicklung der Zeiten Zackommenschafts-Generation bei tieischen Arbastarden Arcb*. Julius Klaus-Stift. Verebungforscb. Socialanthropol Rossenbg, 1945, 20, 236-251.

STEVENS, L. Genetics and Evolution of the Domestic Fowl. Cambridge University Press, 1991.

STOREY, A. A., ATHENS, J. S., BRYANT, D., CARSON, M., EMERY, K., HIGHAM, C.,... & MATISOO-SMITH, E. Investigating the global dispersal of chickens in prehistory using ancient mitochondrial DNA signatures. Plos One, 2012, 7(7), e39171.

XIANG, H., GAO, J., YU, B., ZHOU, H., CAI, D., ZHANG, Y., ... & ZHAO, X. Early Holocene chicken domestication in northern China. Proceedings of the National Academy of Sciences, 2014, 111(49), 17564-17569.

YAMASHINA, M. Y. Karyotype studies in birds. Cytologia, 1944, 13(3-4), 270-296.

Chapter 2

Signature of artificial selection and ecological landscape on morphological structures of Ethiopian village chickens

DESTA, T. T., DESSIE, T., BETTRIDGE, J., LYNCH, S. E., MELESE, K., COLLINS, M., CHRISTLEY, R. M., WIGLEY, P., KAISER, P., TERFA, Z., MWACHARO, J. & HANOTTE, O.

Animal Genetic Resources, 2013, 52, 17–29.

Abstract

Village chickens have been maintained for millennia by smallholder farmers of the least developed world. Our study intends to dissect the impact of artificial selection and ecological landscape on morphological attributes of Ethiopian Horro and Jarso chickens. Morphological score of village chickens (n = 798) and a concise interview conducted to elicit preferences of farmers (n = 399) on morph variants have been used in this analysis. Statistically significant differences in morphological scores were commonly found for rare variants of the morphs scored. The rare variants were more frequent in Jarso chickens, with some unique to Jarso chickens. This morphological variation may be explained by the impact of locally driven evolutionary forces and differences in breeding history of the two chicken populations. A high intrapopulation morphological diversity was observed in the two chicken populations that have been largely evolved under uncontrolled mating. Single comb is less preferred by most of the respondents (93.8%); which was then occurring at low frequency (26.7%). Farmers show a high preference for yellow shank (42.3%), which was then frequently observed (61.1%). The reported reasons for preference of morphs were visual appeal, market demand and cultural and religious significance. A non-significant difference observed in morph preferences between the two regions is attributed to the multifunctional role of village chickens.

Keywords: village chickens, morphological structure, Mendelian inheritance, farmers' preference, ecological variation

Introduction

The village chickens have been maintained for millennia under traditional management practices of the smallholder farmers, which have made them to adapt to a wide range of ecological landscapes. Village chickens are characterized by nondescript and hyper-variable morphological characteristics (Orheruata et al., 2006; Halima et al., 2007; Dana et al., 2010a; Egahi et al., 2010; Melesse and Negesse, 2011). Village chickens also show a vast array of morphological variations in plumage color and pattern, comb shape, earlobe color, shank color etc. (Orheruata et al., 2006).

Studies conducted since the earliest twentieth century have confirmed a Mendelian mode of inheritance pattern for visual traits of the chicken (*e.g.* Bateson, 1909; Punnett, 1923). For example, it has been substantiated that single comb (the wild variant) is recessive to all comb shapes; except the comb-less variant – Breda (http://www.edelras.nl/chickengenetics/mutations2.html). The causative genetic variants underlying variation of some of these morphological traits were mapped to their respective genomic regions (*e.g.* Dorshorst et al., 2010; Wragg et al., 2012). Segregation of morphological traits in village chickens has been shaped by uncontrolled breeding and this provides a unique and a powerful resource to map the impact of natural selection (Wragg et al., 2012). Moreover, the nondescript morphological structure of village chickens can be used to study the impact of natural selection on genetic structure of the domestic fowl.

The origin and extent of this diversity, however, remains understudied and the impact of natural and/or artificial selection on morphological diversity of village chickens has not been sufficiently documented. It is strongly believed that the multipurpose role of village chickens rearing has played a significant role in shaping this phenotypic diversity. A high phenotypic diversity is inevitable for village chickens to survive in resource-constrained production systems. A high phenotypic diversity of village chickens has been partly induced by uncontrolled breeding, which represents the main difference between village and commercial and fancy chickens that have been selected for decades to produce phenotypically homogenous populations.

Genetic improvement of village chickens for production traits is exceptionally rare (Dessie et al., 2011). However, there is a mild selection practiced on visual traits by smallholder farmers who keep chickens not only for their direct use values (egg and meat production) but also to address their visual appeal and cultural and religious needs (Dessie and Ogle, 2001; Dana *et al.*, 2010a; Melesse and Negesse, 2011). In spite of a considerable variation in trait preference, individuals and/or communities may show some common preference. When preference prevails, selective breeding is practiced to maintain and increase the proportion of the preferred phenotype(s) (Bartels, 2003). Smallholder farmers usually have broad breeding objectives to fulfil their multi-functional needs (Dana et al., 2010b; Moges et al., 2010) by keeping flocks showing diverse phenotypes. Our study is intended to elicit farmers' preference to visual traits and to assess variation in morphological scores between the two chicken populations. Our study confirms the commonness of rare morph variants in Jarso chicken and absence of a significant difference in trait preference between the two communities.

Materials and methods

The study sites

The two study sites include Horro (37°01′E to 37°12′E longitudes and 9°55′N to 9°77′N latitudes, recorded for the study villages only) from Western Ethiopia and Jarso (42°10′E to 42°16′E longitudes and 9°25′N to 9°41′N latitudes) from Eastern Ethiopia. The two sites are ~ 870km apart and are known for their considerable variation in farming system and ecological landscape. Horro is characterized by a sub-humid agro-ecological zone and food self-sufficiency. According to Horro district Office of Agriculture unpublished data, the average annual rainfall is 1685mm (range: 1300 – 1800mm) and the average annual temperature is 19°C (range: 14 – 24°C). Majority of the population are Christians (~ 98%) and the agricultural (rural) population represents ~ 88% of the total population. Horro consists of twenty one rural villages. Horro has a cereal production dominated crop livestock mixed farming system. Livestock and livestock products are among the main sources of incomes. The main crops cultivated include maize, teff (*Eragrostis tef*), noug (*Guizotia abyssinica*), wheat, barley, faba bean and peas. The livestock

species kept by farmers in descending order of counts are cattle, sheep, chicken, goat, horse and donkey. The major soil types are redzinas, haplic and luvic phaeozems.

Unpublished data from Jarso district Office of Agriculture indicates that Jarso district represents a semi-arid agro-ecological zone and is food deficient. The average annual rainfall is 700mm (range: 600 – 900mm) and the average temperature is 21°C (range: 14 – 24°C). Unlike Horro, almost all the inhabitants are Muslim (~ 99%). Jarso has eighteen rural villages. Jarso is characterized by a highly rugged terrain and a degraded landscape. Most of the farmers earn their livings from khat (*Catha edulis*) and Irish and sweet potato production, and petty trading. Livestock species reared by farmers in descending order of counts are chicken, cattle, sheep, goat, donkey, camel and a few feral horses. The main soil types include lithosols, vertic luvisols, eutric regosols, vertic cambisols and eutric fluvisols.

The study populations

According to the respective district Office of Agriculture, the total number of chickens found in Horro and Jarso are 38776 and 62829, respectively. The two chicken populations subsist on scavenging and mating is uncontrolled and in most instances random.

Sampling methods

The two sites were selected based on their wide variation in socio-economic and agro-ecological setup. A reconnaissance tour was made in advance in both sites to identify sampling units (villages). A pilot study was then conducted to have hands on exercise on sampling and data collection procedures. Villages that are located near to town centres were largely excluded from sampling to reduce the impact of urban affiliated farming systems on a typical rural village-based chicken production system. Only farmers that keep indigenous chickens were visited. From each site two marketsheds were selected and each marketshed was represented by two villages. Based on this sampling strategy, we sample eight villages *i.e.* four from each study site. Equal numbers of households were visited in each village, marketshed and study

site. The data were collected in two sessions in 2011 (from April to June and from October to November).

The study households were selected from a master list of households belonged to the village and chickens were selected from a family flock in both cases using random table numbers. A random selection of chicken was performed when more than two adult chickens met the selection criteria. However, when the flock size is small, two adult chickens were directly picked up. In the absence of cock, two hens were sampled. Twenty five households were sampled in each village and sampling session and two adult chickens greater than ~ 6 month old were sampled from each household. The chickens sampled were snapshot and scored for unique morphological variants using a pre-coded format. Following this sampling strategy, we sample 798 chickens (400 in Horro and 398 in Jarso) from 399 households (200 from Horro and 199 from Jarso).

Data sources

Multiple photographs of each chicken were taken from lateral, front and dorsal side besides details of the head and the leg region. Morphs were scored on spot based on direct observation and using photographs. Due to a rudimentary and ambiguous nature of hens' comb, we analysed this trait for cocks only. Moreover, the chicken owners were briefly one-to-one interviewed for their preferences for comb shape and shank colour variants and they were also asked to describe their reasons for preference. Earlobe has no local name and farmers do not consider this trait while selecting breeding stocks. Variation in earlobe colour in the two sites is therefore largely under the impact of demographic history of the two chicken populations and natural selection.

Data management and analysis

The response variables were analysed using the non-parametric Chi-square test of the base R (R Core Team, 2012). Logistic regression of SPSS (2007) was used to analyse binomial traits (Equation 1).

$$L_i = \ln\left[\frac{p_i}{1-p_i}\right] \qquad (1)$$

Where p_i is the probability of presence of a phenotype, $1-p_i$ the probability for absence of a phenotype and L_i the natural log of odds ratio. In this model, the study site is used as a classifying variable.

Cramer's V (Equation 2) as implemented in SPSS was used to test the relation between three pigmentation traits (earlobe, shank and beak colour) and association between the study site and four morphological traits (crest, earlobe colour, shank colour and beak colour).

$$V = \sqrt{\frac{\chi^2}{n(k-1)}} \qquad (2)$$

Where V is Cramer's V, n the number of observations and k the number of traits analysed.

Correspondence analysis (CA) was used to assess the discriminating power of earlobe colour, shank colour, beak colour and crest variants using CA Package (Nenadic and Greenacre, 2007) of the R.

Results

Counts of the chickens sampled

We sampled 160 cocks and 240 hens from Horro (cock to hen ratio of 1: 1.5) and 121 cocks and 277 hens (1: 2.29) from Jarso. Our plan was to sample an equal number of cocks and hens; however this failed due to frequent absence of cocks in a family flock. Particularly, one to one ratio was considerably violated by frequent absence of cocks in flocks sampled from Jarso ($\chi^2_1 = 5.22$; $P = 0.022$). Finding fewer cocks than hens is common in village chickens, because in most instances cocks are slaughtered for religious and/ or cultural ceremonies, to welcome guests and for

family's consumption. Moreover, cocks are sold out more frequently to cover incidental expenses.

Comb-shape variants

We found seven variants of comb shapes (single, buttercup, duplex, pea, rose, strawberry and walnut) in the two sites (Table 2.1 and Figure 2.1 (i) to (v)). The proportion of the seven comb types show a significant difference ($\chi^2_6 = 802.07$; $P < 0.001$). However, we left a few ambiguous comb shapes unclassified ($n = 3$). In Horro, cocks' comb shapes were limited to rose and single, whereas, though rare, other comb types were scored on Jarso cocks. The derived comb variants were more frequent than the single comb as of the wild-type, *i.e.*, 73.3% ($n = 206$) versus 26.7% ($n = 75$) ($\chi^2_1 = 61.07$; $P < 0.001$). A logistic regression analysis shows that the chance of getting a single-combed cock in Horro is less likely by 66%.

Table 2.1 Frequency count differences observed in cocks' comb variants within and between sites.

Comb types	**Horro**	**Jarso**	**Total**	χ^2
	N (%)	N (%)	N (%)	
Rose	130 (81.3)	49 (40.5)	179 (63.7)	17.96***
Single (the wild type)	28 (17.5)	47 (38.8)	75 (26.7)	11.76**
Pea	0 (0.0)	9 (7.4)	9 (3.2)	11.90**
Walnut	0 (0.0)	6 (5.0)	6 (2.1)	7.93*
Strawberry	0 (0.0)	5 (4.1)	5 (1.8)	6.61*
Butter cup	0 (0.0)	3 (2.5)	3 (1.1)	3.97*
Duplex	0 (0.0)	1 (0.8)	1 (0.4)	1.32^{ns}
Unclassified	2 (1.3)	1 (0.8)	3 (1.1)	0.12^{ns}
χ^2	171.65***	193.91***	795.93***	

^{ns} not significant; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

(i) (ii) (iii) (iv) (v)

Figure 2.1 Comb types: (i) single, (ii) pea, (iii) rose, (iv) strawberry and (v) walnut

Preference for comb shape

Farmers adopt a dichotomous classification and preference pattern for comb-shape variants. Farmers' preference is between single and variants of the derived comb shapes. Farmers don't care as such for details of the morphological differences observed among the derived comb shape variants. We found that 98% of the Horro and 89.5% of the Jarso farmers show preference to the derived variants of comb shape with absence of statistically significant difference ($\chi^2_1 = 0.39$; $P = 0.535$). Preference of comb shapes is dictated by the reasons presented in Table 2.2.

Table 2.2 Reasons reported for comb type preferences.

Stated reasons	**Horro**	**Jarso**	**Overall**	χ^2
	N (%)	N (%)	N (%)	
Aesthetic value	82 (32.4)	81 (35.2)	163 (33.8)	0.28^{ns}
Market demand	140 (55.3)	122 (53.0)	262 (54.2)	0.12^{ns}
Religious and cultural value	15 (5.9)	4 (1.7)	19 (3.9)	5.38*
Growth rate	16 (6.3)	23 (10.0)	39 (8.1)	2.02^{ns}
χ^2	170.79***	152.43***	321.10***	

^{ns} not significant; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Ear tuft and variants of other Mendelian traits

A rudimentary type of ear tuft was observed in all the chickens sampled despite a slight variation in its size and appearance. However, multiple spur, polydactyl, heterodactyl, syndactyly and bantam chickens were not found. A single naked neck chicken (unfortunately not picked up by the random sampling) was found in Horro.

Rareness of naked neck chickens may associate with high elevation of the study sites (higher than ~ 2000m above sea level for all the villages sampled). Farmers also show less preference to the naked neck phenotype. Their main reasons were lack of visual appeal (58.0%); strangeness (41.0%) and low market demand (12.4%). Farmers also perceive that the naked neck expose chickens to cold and rain (3.8%) and the naked neck can be easily caught by predators (1.9%) than the feathered neck which is slippery.

Tufted crest

A tufted crest (Figure 2.2) was commonly observed in hens (n = 106, 83.5%) than cocks (n = 21, 16.5%) ($\chi^2_1 = 55.82$, $P < 0.001$). This is the typical characteristic feature of sex-influenced traits. A logistic regression analysis also shows that the likelihood of being crestless in cocks is more likely by 77.3% than hens. Moreover, the chance of finding crested chickens in Horro is ~ 10.5 times more likely than Jarso. The proportion of crested head scored in the entire population was 16.1% (127/791), which shows the commonness of a plain head ($\chi^2_1 = 364.56$, $P < 0.001$). In Horro, 111 crested head chickens (27.9%, 111/398) were sampled, whereas these were 16 (4.1%, 16/393) in Jarso, which exhibits a highly significant difference ($\chi^2_1 = 69.87$, $P < 0.001$).

(i) (ii)

Figure 2.2 Crest: (i) crested hen, (ii) crestless hen

Earlobe colour

The village chickens sampled display a high diversity in earlobe pigmentation (Figure 2.3 (i) – (iv)). For example, we found red and pink earlobes that are intermingled or peppered with a different proportion of white or yellow colour. A wide variation in earlobe pigmentation has made the classification of earlobe colours

into distinct phenotypic classes a challenging task. We grouped the chickens sampled into broad phenotypic classes by pooling closely related variants of earlobe colour to a common slot (Table 2.3).

Table 2.3 Earlobe colour variants of the two chicken populations.

Earlobe colour	**Horro**	**Jarso**	**Total**	χ^2
	N (%)	N (%)	N (%)	
Dark brown	1 (0.3)	0 (0.0)	1 (0.1)	0.97^{ns}
Pink	78 (19.9)	41 (10.8)	119 (15.5)	10.29**
Pink and white	63 (16.1)	75 (19.8)	138 (17.9)	1.49^{ns}
Pink and yellow	0 (0.0)	6 (1.6)	6 (0.8)	ND
Red	159 (40.7)	120 (31.7)	279 (36.3)	4.21*
Red and white	81 (20.7)	123 (32.5)	204 (26.5)	10.13**
White	6 (1.5)	11 (2.9)	17 (2.2)	1.65^{ns}
Yellow	3 (0.8)	2 (0.5)	5 (0.7)	0.17^{ns}
χ^2	359.86***	307.11***	822.81***	

^{ns} not significant; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; ND - not done

(i) (ii) (iii) (iv)

Figure 2.3 Earlobe colour: (i) red, (ii) white, (iii) yellow & purple, (iv) white spotted red

Beak colour

Usually, the two horny mandibles of the chicken beak show variation in colour. The lower mandible is usually brighter than the upper, which then creates a mosaic appearance. Dimorphism in beak colour may occur due to continuous exposure of the upper beak to sunlight besides a differential act of body region specific biochemical processes. Some of the beak colour phenotypes scored are presented in Table 4 and are displayed in Figure 2.4 (i) – (iii).

Table 2.4 Beak colour variants observed in the chicken populations studied.

Beak colour	Horro	Jarso	Total	χ^2
	N (%)	N (%)	N (%)	
Yellow	153 (39.1)	145 (38.4)	298 (38.8)	0.03^{ns}
White	2 (0.5)	5 (1.3)	7 (0.9)	1.39^{ns}
Brownish yellow	176 (45.0)	103 (27.2)	279 (36.3)	16.72***
Brown	37 (9.5)	71 (18.8)	108 (14.0)	11.89**
Black	23 (5.9)	54 (14.3)	77 (10.0)	13.55**
χ^2	328.78***	146.02***	429.23***	

^{ns} not significant; ** $P < 0.01$; *** $P < 0.001$

(i) (ii) (iii)

Figure 2.4 Beak colour: (i) brownish yellow, (ii) yellow and (iii) black

Shank colour

The commonest shank colours were yellow and white, whereas the rare variants include slate blue, green, black and brown (Table 2.5 and Figure 2.5 (i) – (v)). Dark shanks (black, slate blue and willow green) were relatively more frequent in Jarso than Horro chickens ($\chi^2_1 = 47.90$, $P < 0.001$).

Table 2.5 Variants of shank colour observed in Horro and Jarso chickens.

Shank colour	Horro	Jarso	Total	χ^2
	N (%)	N (%)	N (%)	
Yellow	279 (71.5)	190 (50.4)	469 (61.1)	14.01 **
White	93 (23.8)	91 (24.1)	184 (24.0)	0.01^{ns}
Black	7 (1.8)	23 (6.1)	30 (3.9)	9.09*
Slate blue	4 (1.0)	37 (9.8)	41 (5.3)	27.70***
Green	6 (1.5)	23 (6.1)	29 (3.9)	10.55**
Mottled	1 (0.3)	13 (3.4)	14 (1.8)	10.70**
χ^2	942.18***	370.64***	1246.83***	

^{ns} not significant; * $P < 0.05$; ** $P < 0.01$;*** $P < 0.001$

(i) (ii) (iii) (iv) (v)

Figure 2.5 Shank colour: (a) yellow, (b) green, (c) slate blue, (d) black and (e) white

Farmers' preference for shank colour

Farmers were asked to rank shank colour of their preference including white, yellow or black (slate blue); however, few farmers mentioned a red shank as an additional variant. According to these farmers red shank represents a deep yellow shank that has vertical light pink bands. Farmers show a high preference to yellow shank (Table 2.6).

Table 2.6 Preferences of farmers to shank colour variants.

Shank colour	Horro	Jarso	Total	χ^2
	N (%)	N (%)	N (%)	
Yellow	80 (39.4)	88 (45.4)	168 (42.3)	0.83^{ns}
White	57 (28.1)	44 (22.7)	101 (25.4)	1.14^{ns}
Black/slate blue	31 (15.3)	36 (18.6)	67 (16.9)	0.63^{ns}
Red	35 (17.2)	26 (13.7)	61 (15.4)	0.95^{ns}
χ^2	30.20***	46.25***	72.87***	

^{ns} not significant; *** $P < 0.001$

Lightly feathered shank and spur

A lightly feathered shank (ptilopody, Figure 2.6) was rarely observed and only 2.1% (17/798) of the chickens sampled have had this phenotype ($\chi^2_1 = 731.44$, $P < 0.001$). Ptilopody was scored in seven chickens from Horro (1.8%, 7/ 400, $\chi^2_1 = 372.49$, $P < 0.001$) and 10 from Jarso (2.5%, 10/398, $\chi^2_1 = 359.01$, $P < 0.001$). However, the proportions of ptilopody in the two sites is not significantly different ($\chi^2_1 = 0.54$, $P = 0.461$). A well grown spur was observed in cocks whereas it was a rudimentary type in most of the hens (data not shown). We noticed that the length of spur is mostly associated with the age of the bird, the older the bird, the longer the spur.

Figure 2.6 Lightly feathered shanks (ptilopody).

Correspondence analysis of morph traits

The summary function of CA shows that the three (the original four variables (morphological traits) minus one) principal inertias (eigenvalues) account for 56.5, 25.0 and 18.5% of the total variation with corresponding inertia values of 0.029, 0.013 and 0.009, respectively. Most of the chickens were tightly clustered (Figure 2.7). Variations in earlobe, shank and beak colour and crest explain 39.4, 20.2, 23.3 and 17.1% of the total inertia, respectively with corresponding inertia values of 0.020, 0.010, 0.012 and 0.009, respectively. A correspondence analysis shows that despite its binary nature, crest contributes for a considerable proportion of the total variation. Only a few outlier birds show a higher inertia whereas for most of the chickens sampled, individual differences for traits analysed are weak (Figure 2.8). All the four traits show coordinates of different signs (direction of the arrows), which indicates a weak association among them.

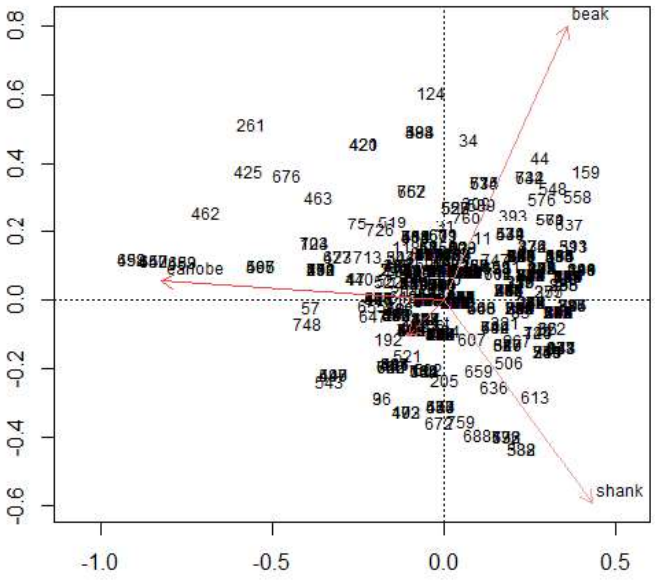


Figure 2.7 Correspondence analysis plot for variants of earlobe, shank and beak colours and crest in the two chicken populations. Serial numbers 1 to 391 and 392 to 769 represent Horro and Jarso chickens respectively.

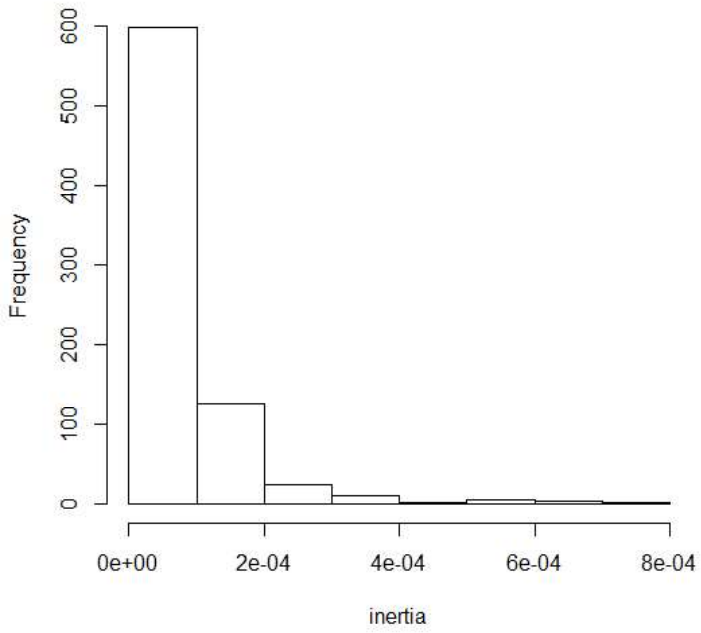


Figure 2.8 Frequency distribution of eigenvalues (inertias) explained by individual chickens.

Relationship between qualitative traits

A Cramer's V test reveals that variants of three morphological traits (earlobe, beak and shank colour) show a weak but statistically significant correlation, *i.e.* earlobe versus shank colour (0.230, $P < 0.001$); earlobe versus beak colour (0.122, $P = 0.003$) and beak colour versus shank colour (0.266, $P < 0.001$). Cramer's V analysis also shows a statistically significant correlation between the study site and variants of crest (0.324, $P < 0.001$), earlobe colour (0.230, $P < 0.001$), shank colour (0.331, $P < 0.001$) and beak colour (0.320, $P < 0.001$).

(i) (ii)

Figure 2.9 The nondescript Ethiopian village chickens and their natural habitat: (i) Horro chickens and (ii) Jarso chickens.

Discussion

The term morph refers to the outward appearance of an individual and it can be used to group members of a population into closely related sub-types. Morphological variation as it has been displayed in Figure 2.9 (i) & (ii) has been a subject of thorough studies due to its importance in evolutionary biology, socio-cultural life of subsistence farmers and in adaptive radiation. For example, interviewees show a high preference to some variants of a morph, which indicates the importance of morphological diversity in socio-cultural life of subsistence farmers.

Morphological traits may show different proportions among populations. For example, the rareness of single comb (26.7%) found in this study is inconsistent with the report of Melesse and Negesse (2011) for southern Ethiopian chickens (55.0%), Apuno et al. (2011) for Shelleng and Song chickens (96.45%), Daikwo et al. (2011) for Dekina chickens (51%) and Orheruata et al. (2006) for Edo State chickens (92%) from Nigeria and Bhuiyan et al. (2005) for Desi chickens of the Bangladesh (97%). The commonness of single comb in the latter populations may be associated with its special importance in evaporative cooling under warm climates and due to variations in morphological preference of farmers. For example, the southern Ethiopia region is dominated by Protestant Christians (55.5%, FDREPCC, 2008) and this sect is strictly against sacrificial offerings and to all ritual practices. The high demand observed for the derived comb variants elsewhere in Ethiopia (Dessie and Ogle, 2001; Dana et al., 2010a; Moges et al., 2010) is therefore of limited significance in the Southern

Ethiopia. Preference of morph is a common practice among smallholder farmers of Ethiopia, because plumage colour and pattern, sex, comb shape and age are used as main criteria to scarify chicken for ritual ceremonies (Dessie and Ogle, 2001). The best combination of the traits preferred a chicken has, the highest is its market demand and price. Most likely inclusion of morphological traits as selection criteria may have been driven by this socio-cultural significance (Muchadeyi et al., 2009; Dana et al., 2010b; Moges et al., 2010). Muchadeyi et al. (2009) also noted that in some instances chickens have been culled based on their morphological appearance. Systematic culling of single-combed chickens via sale and a slaughter that is intended to consumption is common in Ethiopia. In some parts of Ethiopia this has been practiced for centuries to reduce the frequency of single-comb phenotype. Owing to its homozygous recessive pattern of expression, considerable reduction in allelic variants conferring single comb phenotype could be achieved through systematic culling.

However, in line with our findings, Dana et al. (2010a) reported a low average proportion of single comb in five indigenous Ethiopian chicken populations (13%) and Dong Xuan et al. (2006) in Dong Tao chickens in Vietnam (10%). Even in our study populations, a single comb is more common in Jarso (38.8%) than Horro chickens (17.5%). The commonness of single comb variant in Jarso may be associated with demographic structure of the community. Jarso area is almost entirely inhabited by Muslims who hardly sacrifice chickens for ritual purposes, which otherwise under the context of Ethiopian tradition requires sacrificing chickens displaying the derived comb shape variants (Moges et al., 2010; personal observation).

Dong Xuan et al. (2006) in Vietnamese Dong Tao chickens (90%), Dana et al. (2010a) in Horro chickens and Halima et al. (2007) in north western Ethiopia chicken reported pea comb as a common variant. Unlike, Dana et al. (2010a), we have not scored pea comb in Horro chickens. However, congruent to our findings, rareness of pea comb was reported in Beninese chickens (Youssao et al., 2010). Double (v-shaped) comb was reported in 13.4% of north western Ethiopian chickens (Halima et al., 2007); however, it is rarely observed in the two chicken populations. The most frequent rose comb variant we found partly agrees with Melesse and

Negesse (2011) who report rose comb as the second most common variant. However, all these findings need to be treated with care because inconsistencies may arise among individuals due to variation in defining the derived variants of comb shapes.

Through a continuous selection as it has been observed for comb shapes, farmers unintentionally affect production traits influenced by the pleiotropic effect of loci underlying variation of Mendelian traits. For instance, the homozygous rose is known to reduce fertility (Crawford and Smyth, 1964) and pea comb is noted to reduce tissue mass (see Wright et al., 2009 for a review). A phenotypic correlation found between morphological traits shows the impact of farmers' selection on associated traits. Farmers' preference can also be assessed from ecological adaptation perspectives, for example, in elevated highland areas experiencing temperate like climate, chickens with derived comb variants with a reduced surface area may adapt better by reducing heat loss (see Wright et al., 2009 for the pea comb).

Crest is rare in Nigerian chicken (17.03%, Egahi et al., 2010), Ugandan chickens (12%, Ssewannyana et al., 2008) and Dana et al. (2010a) in five Ethiopian chicken populations (range: 1 – 75%) including Horro (34%), which all support of our findings. However, Halima et al. (2007) reported higher proportion of crest (48.8%) for north western Ethiopia chickens. The frequency of crest shows variation among geographical regions, for example, it is less frequent in south and south-western Ethiopia than north western and western Ethiopia (Dana et al., 2010a). The proportion of crest also shows variation between Horro and Jarso chickens. Wang et al. (2012) reported a voluminous crest in cocks than hens, which disagrees with our findings. We found low proportion of cocks having a rudimentary type crested head, which may partly linked with large comb size in cocks.

In line with our findings; earlobe, comb and wattle colours are commonly red in fancy chickens (Wragg et al., 2012). Unlike a plain earlobe colour reported for southern Ethiopian chickens (Melesse and Negesse, 2011), we found a considerable proportion of earlobe showing a combination of two colours (spotted earlobes). Some Kenyan indigenous chicken populations also have a multi-coloured earlobe

(Kingori et al., 2010). The proportion of white earlobe we found (2.2%) is lower than Beninese Forest type (60.8%) and Savannah-type chickens (45.1%) (Youssao et al., 2010), Nigerian chickens (73.21%, Egahi et al., 2010) and Ugandan chickens (48%, Ssewannyana et al., 2008). White earlobe is most frequent in chickens of the Mediterranean region (FAO, 2010). However, Orheruata et al. (2006) reported red earlobe colour as the commonest variant (60%) in Nigeria Edo State chickens, though, reasonably higher proportion of white earlobe (39%) was observed in this population and 47% in Ugandan chickens (Ssewannyana et al., 2008). Vij et al. (2006) reported brown earlobe as the commonest variant in Punjab brown chickens, however, we found only a single chicken (0.1%) showing this phenotype. Earlobe was clearly visible in all the chickens sampled; however, it is less visible in most (64%) of the Ugandan chickens (Ssewannyana et al., 2008). Despite controversies on the inheritance pattern of earlobe colour (Warren, 1928), polygenic and sex-linked patterns have been suggested (Warren, 1928; Wragg et al., 2012), which indeed contributes to a high variation observed.

Unlike our chicken populations that have been dominated by brownish yellow beaks, Kenyan indigenous chickens usually have black and dark grey beaks (Kingori et al., 2010). Beak colour in Punjab brown chickens of India is yellow; however, its upper part turns to black as the chicken gets older (Vij et al., 2006). A comparable type of mosaic beak colour was found in our populations.

Yellow shank is commonly observed in village chickens: Halima et al. (2007, 64.4%), Dana et al. (2010a, 60%) and Melesse and Negesse (2011, 52.5%) from Ethiopia; Daikwo et al. (2011) from Nigeria (40.5%), Ssewannyana et al. (2008) from Ugandan (42%), Bhuiyan et al. (2005) in Bangladeshi Desi chickens (32%) and Orheruata et al. (2006) in Nigerian chickens (30%). However, Youssao et al. (2010) reported low proportion of yellow shank (5%) in Beninese chickens. White shank is the second common variant in the two chicken populations (24.0%), which is comparable with Ugandan chickens (20%, Ssewannyana et al., 2008) and Bangladesh Desi chickens (29%, Bhuiyan et al., 2005). However, white shank is more common in Nigerian chickens (41%, Orheruata et al., 2006) and in Beninese chickens (40.1%, Youssao et al., 2010). Black shank is found in 39% of the Bangladeshi Desi chickens (Bhuiyan et al., 2005); 42.2 or 29% in Nigerian chickens

(Orheruata et al., 2006; Egahi et al. 2010) and 21% in Ugandan chickens (Ssewannyana et al., 2008). Slate blue shank is the most frequent in Beninese chickens (43.3%, Youssao et al., 2010). Dark shank colour is a typical characteristic feature of the red junglefowl (for reviews, see Brisbin and Peterson, 2007 and the references therein). However, the proportion of black and slate blue shank is low in our chicken populations. Green shank was rarely observed (3.9%) in Horro and Jarso chickens and a comparable proportion (0.5%) was reported for Ugandan chickens (Ssewannyana et al., 2008). However, Halima et al. (2007) reported higher proportion of green-shanked chickens (12%) from north western Ethiopia. Rareness of green shank may associate with its unfavourable correlation with viability (see McGibbon, 1979 and the references therein). Apuno et al., (2011) have reported highest proportion of pink shank (38.8%) in Nigerian chickens, which can be considered as an outlier variant. The low proportion of a lightly feathered shank found in our study (2.1) is invariably confirmed by Melesse and Negesse (2011) (2%), Halima et al. (2007) (2.5%) and Ssewannyana et al. (2008) (4%).

In line with our findings (data not shown), almost all combs and wattles of the indigenous chickens in Kenya are red, except a few spotted variants having white and black colours (Kingori et al., 2010). However, we have not observed any spotted variants of comb and wattle. It is biologically important to study comb and wattle colour because these are usually indicators of chickens' health and egg laying status (Hume, 2011). Moreover, Navara et al., (2012) reported a positive correlation between brightness of comb and sperm viability.

It is thought that, unlike quantitative traits, the environment has less impact on traits that show Mendelian mode of inheritance. However, we found that even qualitative traits that haven't been under the influence of intentional artificial selection are significantly differing between the two chicken populations (e.g. some variants of the earlobe and beak colour). This may indicate the impact of ecological variation and the breeding history of village chickens. Moreover, four of the morphological traits (crest, earlobe colour, shank colour and beak colour) show statistically significant correlation with the study sites, which implicates the impact of ecological landscape on variation in morphological traits. Although preference of farmers for comb shape and shank colour was not significantly different between the two sites,

differences in culture and religion may have contributed especially for difference in single comb proportions observed between the two chicken populations. The intrapopulation low frequency of single comb and high frequency of yellow shank likely indicates the impact of artificial selection. However, a mild selection of farmers is intended at keeping chickens that display diverse phenotypes to address their different socio-cultural needs. This implies that population specific morph variant is less evident in the village chickens.

Except a tufted crest, all other rare morph variants in most instances were scored in Jarso chickens, which may associate with difference in breeding history of the two chicken populations. There is a historical and archaeological evidence for earlier settlement of people in northern Ethiopia (e.g. D'Andrea et al., 2011) and perhaps human settlement in eastern Ethiopia where Jarso district is located had been initiated closer to this prehistoric time (Betemariam, 2011). The degraded landscape of the Jarso also indicates an ancient practice of agriculture. On the other hand, elderly farmers in Horro recalled the very recent (~ 40 years back) cover of a dense forest, which has now been however cleared for agricultural activities and human settlement. We also noticed that Horro still has patches of dense natural forest and Horro district land use pattern profile also show an extant cover of ~ 42.6% natural forest. Jarso is relatively closer to an ancient trade route between Ethiopia and the Middle East involving the Red Sea coast – one of the putative routes for introduction of domestic chickens to Africa (Blench and Macdonald, 2000). This may have made chickens to arrive earlier in Jarso and chicken rearing may had started earlier in Jarso.

Conclusion

Most of the rare morph variants found in our study were observed in Jarso chickens (except crest), which shows a high phenotypic diversity of Jarso chickens. This indicates that Jarso chickens may have descended from earlier founding population. Phenotypic diversity is favourably associated with the length of breeding history and inversely related to geographical distance from centre of domestication (Jarso is relatively closer to the putative centre of chicken domestication). Owing to its ancient inhabitation, stochastic factors and evolutionary forces may have had more

time to shape the germline mutation in the Jarso chickens. Moreover, due to the impact of a long history of mild artificial selection, Jarso chickens most likely have had better chance to accumulate some of the rare variants that favoured by human-driven selection. Difference in their ancestral genetic background (genetic structure of the founder population) and variation in ecological landscape may have also contributed to difference observed. A relatively lower phenotypic diversity observed for most of the morphological traits in Horro chickens on the other side may indicate the consequence of a very recent introduction of chickens following the short history of human settlement. The relative abundance of crested chickens in Horro, however, may associate with a high frequency of crested alleles in the founding population.

References

APUNO, A. A., MBAP, S. T. & IBRAHIM, T. Characterization of local chickens (*Gallus gallus domesticus*) in Shelleng and Song Local Government Areas of Adamawa State, Nigeria. Agriculture and Biology Journal of North America, 2011, 2(1), 6-14.

BARTELS, T. Variations in the morphology, distribution, and arrangement of feathers in domesticated birds. Journal of Experimental Zoology, 2003, 298B, 91-108.

BATESON, W. Mendel's principles of heredity. Cambridge, Cambridge University Press, 1909.

BETEMARIAM, E. A. Forest diversity in fragmented landscapes of northern Ethiopia and implications for conservation. PhD Thesis, University of Bonn, Germany, 2011.

BHUIYAN, A. K. F. H., BHUIYAN, M. S. A. & DEB, G. K. Indigenous chicken genetic resources in Bangladesh: current status and future outlook. Animal Genetic Resources Information, 2005, 36, 73-84.

BLENCH, R. M. & MACDONALD, K. C. Chickens. In F.K. Kenneth & C.O. Kriemhild, eds, The Cambridge world history of food, Volume I, pp. 496–499. Cambridge, Cambridge University Press, 2000.

BRISBIN, JR., I. L. & PETERSON, A. T. Playing chicken with red jungle fowl: identifying phenotypic markers of genetic purity in *Gallus gallus*. Animal Conservation, 2007, 10, 429-435.

CRAWFORD, R. D. & SMYTH, J. R. Studies of the relationship between fertility and the gene for rose comb in the domestic fowl. 1. The relationship between comb genotype and fertility. Poultry Science, 1964, 43, 1009-1017.

D'ANDREA, A. C., RICHARDS, M. P., PAVLISH, L. A., WOOD, S., MANZO, A. & WOLDE-KIROS, H. S. Stable isotopic analysis of human and animal diets from two pre-Aksumite/proto-Aksumite archaeological sites in northern Ethiopia. Journal of Archaeological Science, 2011, 38, 367-374.

DAIKWO, I. S., OKPE, A. A. & OCHEJA, J. O. Phenotypic characterization of local chickens in Dekina. International Journal of Poultry Science, 2011, 10(6), 444-447.

DANA, N., DESSIE, T., VAN DER WAAIJ, L. H. & VAN ARENDONK, J. A. M. Morphological features of indigenous chicken populations of Ethiopia. Animal Genetic Resources, 2010a, 46, 11-23.

DANA, N., VAN DER WAAIJ, L. H., DESSIE, T. & VAN ARENDONK, J. A. M. Production objectives and trait preferences of village poultry producers of Ethiopia: implications for designing breeding schemes utilizing indigenous chicken genetic resources. Tropical Animal Health and Production, 2010b, 42, 1519-1529.

DESSIE, T. & OGLE, B. Village poultry production systems in the central highlands of Ethiopia. Tropical Animal Health and Production, 2001, 33(6), 521-537.

DESSIE, T., TAYE, T., DANA, N., AYALEW, W. & HANOTTE, O. Current state of knowledge on phenotypic characteristics of indigenous chickens in the tropics. World's Poultry Science Journal, 2011, 67, 507-516.

DONG XUAN, D. T., SZALAY, I., SU, V. V., TIEU, H. V. & DANG VANG, N. Animal genetic resources and traditional farming in Vietnam. Animal Genetic Resources Information, 2006, 38, 1-17.

DORSHORST, B., OKIMOTO, R. & ASHWELL, C. Genomic regions associated with dermal hyperpigmentation, polydactyly and other morphological traits in the Silkie chicken. Journal of Heredity, 2010, 101 (3), 339-350.

EGAHI, J. O., DIM, N. I., MOMOH, O. M. & GWAZA, D. S. Variations in qualitative traits in the Nigerian local chicken. International Journal of Poultry Science, 2010, 9(10), 978-979.

FAO. Chicken genetic resources used in smallholder production systems and opportunities for their development, by P. Sørensen. FAO Smallholder Poultry Production Paper No. 5. Rome, 2010 (available at http://www.fao.org/docrep/013/al675e/al675e00.pdf).

FDREPCC (Federal Democratic Republic of Ethiopia Population Census Commission). Summary and statistical report of the 2007 population and housing census: Population size by age and sex. December 2008, Addis Ababa, Ethiopia (available at http:// www.csa.gov.et/pdf/Cen2007_firstdraft.pdf) (accessed 1 May 2012).

HALIMA, H., NESER, F. W. C., VAN MARLE-KOSTER, E. & DE KOCK, A. Phenotypic variation of native chicken populations in northwest Ethiopia. Tropical Animal Health and Production, 2007, 39, 507-513.

HUME, T. Backyard chicken husbandry: Part 1. Companion Animal, 2011, 16(6), 43-46.

KINGORI, A. M., WACHIRA, A. M. & TUITOEK, J. K. Indigenous chicken production in Kenya: a review. International Journal of Poultry Science, 2010, 9(4), 309-316.

MCGIBBON, W. H. Green shanks and adult mortality in chickens. Journal of Heredity, 1979, 70(1), 44-46.

MELESSE, A. & NEGESSE, T. Phenotypic and morphological characterization of indigenous chicken populations in southern region of Ethiopia. Animal Genetic Resources, 2011, 49, 19-31.

MOGES, F., MELLESSE, A. & DESSIE, T. Assessment of village chicken production system and evaluation of the productive and reproductive performance of local chicken ecotype in Bure district, north-west Ethiopia. African Journal of Agricultural Research, 2010, 5(13), 1739-1748.

MUCHADEYI, F. C., WOLLNY, C. B. A., EDING, H., WEIGEND, S. & SIMIANER, H. Choice of breeding stock, preference of production traits and culling criteria of village chickens among Zimbabwe agro-ecological zones. Tropical Animal Health and Production, 2009, 41, 403-412.

NAVARA, K. J., ANDERSON, E. M. & EDWARDS, M. L. Comb size and color relate to sperm quality: a test of the phenotype-linked fertility hypothesis. Behavioral Ecology, 2012, 23(5), 1036-1041.

NENADIC, O. & GREENACRE, M. Correspondence analysis in R, with two-and three-dimensional graphics: The ca Package. Journal of Statistical Software, 2007. 20(3), http:/www.jstatsoft.org/v20/i03/

ORHERUATA, A. M., ADEGITE, A. V. & OKPEKU, M. Morphological and egg characteristics of indigenous chicken in Edo State, Nigeria. Nigerian Agriculture Journal, 2006, 37, 114-123.

PUNNETT, R. C. Heredity in Poultry. London, Macmillan and Co, 1923.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012, ISBN 3-900051-07-0, http://www.R-project.org/.

SPSS. SPSS 16.0 for Windows User's Guide Release. SPSS Inc., Chicago, 2007.

SSEWANNYANA, E., SSALI, A., KASADHA, T., DHIKUSOOKA, M., KASOMA, P., KALEMA, J., KWATOTYO, B. A. & AZIKU, L. On-farm characterization of indigenous chickens in Uganda. Journal of Animal and Plant Sciences, 2008, 1(2), 33-37.

VIJ, P. K., TANTIA, M. S. & VIJH, R. K. Characterization of Punjab brown chicken. Animal Genetic Resources Information, 2006, 39, 65-76.

WANG, Y., GAO, Y., IMSLAND, F., GU, X., FENG, C., LIU, R., SONG, C., TIXIER-BOICHARD, M., GOURICHON, D., LI, Q., CHEN, K., LI, H., ANDERSSON, L., HU, X. & LI, N. The crest phenotype in chicken is associated with ectopic expression of HOXC8 in cranial skin. Plos One, 2012, 7(4), e34012.

WARREN, D. C. Inheritance of earlobe color in poultry. Genetics, 1928, 13, 470-487.

WRAGG, D., MWACHARO, J. M., ALCALDE, J. A., HOCKING, P. M. & HANOTTE, O. Analysis of genome-wide structure, diversity and fine mapping of Mendelian traits in traditional and village chickens. Heredity, 2012, 109, 6-18.

WRIGHT, D., BOIJE, H., MEADOWS, J. R. S., BED'HOM, B., GOURICHON, D., VIEAUD, A., TIXIER-BOICHARD, M., RUBIN, C-J., IMSLAND, F., HALLBOOK, F. & ANDERSSON, L. Copy number variation in intron 1 of SOX5 causes the pea comb phenotype in chickens. PLoS Genetics, 2009, 5(6), e1000512.

YOUSSAO, I. A. K., TOBADA, P. C., KOUTINHOUIN, B. G., DAHOUDA, M., IDRISSOU, N. D., BONOU, G. A., TOUGAN, U. P., AHOUNOU, S., YAPI-GNAORÉ, V., KAYANG, B., ROGNON, X. & TIXIER-BOICHARD, M. Phenotypic characterisation and molecular polymorphism of indigenous poultry populations of the species *Gallus gallus* of savannah and forest ecotypes of Benin. African Journal of Biotechnology, 2010, 9(3), 369-381.

Chapter 3

Morphological diversity of plumage colour and pattern in Ethiopian village chickens

Abstract

Village chickens are well known for their high diversity in plumage colour and pattern. This diversity represents the legacy of polygenic inheritance pattern and the impact of evolutionary forces and a mild artificial selection. Our analysis is based on hens (n = 485) and cocks (n = 281) sampled from Horro and Jarso regions of Western and Eastern Ethiopia, respectively. Variants of plumage colour were scored using photograph. A concise one-to-one interview of farmers (n = 399) was conducted to elicit farmers' preference to plumage colour and pattern. Regardless of the sex of the chicken, red was the most preferred plumage colour whereas black is the least. Farmers' preferences are dictated by visual appeal, market demand, cultural and religious values. In line with preference of farmers, red plumage was most frequently observed in cocks and hens (red plumage in hens according to farmers' description represents a brownish red). Self-coloured (black or white) varieties were infrequent. This may be to some extent attributed to selection pressure, for example, exposure of white chickens to visually hunting predators and poor preference of farmers to black chickens and a sex-linked nature of the self-black plumage colour. High diversity in plumage colour and pattern observed in village chickens is a good resource to study the genetic control of vertebrates' pigmentation.

Key words: plumage colour, plumage pattern, village chickens, dichromatism, farmers preferences

Introduction

Based on their management and breeding history, village chickens are known with different synonyms. The synonymous among others are native or local or indigenous or scavenging or free roaming or free ranging or smallholders' or backyard or bush or ranging or runner or rural chickens or family or rural or village or farmyard poultry or indigenous fowl. These synonyms invariably indicate the traditional

management systems under which village chickens have been maintained for time immemorial. Village chickens presumably constitute a substantial proportion of the global chicken diversity. According to FAO (2000) African indigenous chickens account for 8% of the extant chicken breeds despite their smaller share to the global chicken population (5%). The actual number of breeds most likely excels this as characterization works that have been conducted in the least developed nations to the level of defining breeds are limited. Moreover, the landrace nature of village chickens complicates the definition of breed. The wide range of variation in plumage colour and pattern observed in village chickens is most likely linked with a high phenotypic diversity.

Plumage colour is a complex polygenic trait (Moore and Smyth, 1972) and underlies the combined effect of dominance and epistatic interaction (Smyth, 1990). Due to its complex inheritance pattern, plumage pigmentation shows a high variation among chicken populations (Hellström et al., 2010). However, there is a possibility for a particular colour pattern to evolve due to stochastic forces like genetic drift (Protas and Patel, 2008). For example, genetic drift may reduce local diversity (increase divergence among populations) in plumage colour due to the impact of vicariance particularly evident as the population size is reduced (Johnson & Burnham, 2012). Besides a high diversity in plumage colour, village chickens are also characterized by a vast array of plumage patterns. This forms the basis of intriguing plumage colours we see today in the domestic fowl. Interestingly, these patterns are inherited in a Mendelian fashion (Pearl and Boring, 1914), which has made them good resources to study the genetic basis of pigmentation.

The common colours that form the basis of plumage colour and pattern in chickens are red, white and black. The basic plumage colours display different patterns when they combine among themselves and with non-basic and infrequent variants to a varying extent at different parts of the chicken body to form the entire set of a plumage. Regional variation in plumage colour may arise from the impact of localization effect of plumage pattern in different parts of the chicken body (Pearl and Boring, 1914). Plumage colours and patterns can vary with respect to the number and type of loci underlying their genetic control. This likely has made

chickens to display a spectacular array of plumage colour compared to other livestock species (Chang et al., 2007).

Pigmentation compounds in chicken are of two main types – melanin and carotenoids. The two compounds interact to each other and/or other cell types to produce structural sheens of the plumage. However, it is the melanin that mostly determines plumage colour and pattern in chickens (Smyth, 1990). Intensity and distribution of melanin is associated with age and sex of the chicken. There are primary and secondary plumage pattern genes. The primary pattern genes are those which determine the pterylar and multipterylar distribution of a plumage colour (Kimball, 1952). The secondary pattern genes determine the distribution of eumelanin within individual feather and most of the secondary patterns are controlled by autosomal inheritance (Smyth, 1990). Carotenoids which are also responsible for yellowish colour in chickens are obtained from feed. In this regard, environmentally induced variation may have impact on plumage colour (Paxton, 2009).

Morphological traits like plumage colour and pattern are important to smallholder farmers to address their cultural and religious needs and to feed their visual appeal. Plumage colour and pattern to some extent have been evolved under the impact of artificial selection. Natural selection has also been involved in shaping plumage colour and pattern of village chickens. Chickens having vivid plumage, for example, are commonly exposed to predators. We have gained insight from farmers' conversations that easily noticed chickens like self-white are commonly killed by predators and as a result self-white chickens are less preferred by some farmers. A similar condition was noted in many avian species that have a wild-type plumage (Hellström et al., 2010). Plumage colour therefore could have adaptive values like predator avoidance, social signalling and communication, thermoregulation, reproductive fitness and abrasion reduction (for details see Riegner 2008 and the references therein; Sheppy, 2011). This study identifies a subtle variation in plumage colour and pattern between Horro and Jarso chickens and a high intrapopulation variation in plumage colour and pattern.

Materials and methods

The study sites and the chicken populations

Two village chicken populations Horro and Jarso are the focus of this study. Horro is located in sub-humid agro-ecological zone of western Ethiopia. Jarso is located in semi-arid part of eastern Ethiopia. Desta et al. (2013) provides a detailed description on production system of the two study sites. The two chicken landraces are native to Ethiopia and have been maintained and bred for millennia by smallholder farmers. The two study sites are selected based on their wide difference in production system and due to their distant geographical location.

Data collection and analysis

The sampling procedure followed was described in detail in Desta et al. (2013). Preference of farmers to plumage colour was recorded for hens and cocks separately. The non-parametric chi-square test of the base R (R Core Team, 2013) was used to analyse the within and between population difference in plumage colour and pattern and owner's preference to plumage colour in hen and cock populations. A summary statistics was performed for frequency counts.

Results

Plumage preference

A vast array of plumage colour was observed in the two chicken populations. Farmers were interviewed to choose their favourite type of plumage colours for hen and cock populations' separately. We did this because plumage colour displays sexual dichromatism. However, farmers classify plumage colour to wide phenotypic classes each containing a number of variants. For a considerable proportion of the plumage colour significantly different preferences were recorded between the two study sites both in hen and cock populations (Table 3.1 & 3.2). However, a high variation in preference of farmers was observed at study site level. Nonetheless, 12.4% of Horro and 15.7% of Jarso farmers' did not show preference to plumage

colour. Those farmers who show no preference welcome any plumage colour and their assertion is this is a natural gift and we should have to appreciate this.

Table 3.1 Farmers' preferences to variants of hens' plumage colour.

Preferred colour	Horro	Jarso	Total	χ^2_1
	N (%)	N (%)	N (%)	
Red (brownish red)	187 (38.72)	189 (39.79)	376 (39.25)	0.01^{ns}
White	110 (22.77)	136 (28.63)	246 (25.68)	2.75^{ns}
Black	34 (7.04)	87 (18.32)	121 (12.63)	23.21***
Wheaten	73 (15.11)	29 (6.11)	102 (10.65)	18.98***
Multi-coloured	59 (12.22)	23 (4.84)	82 (8.56)	15.8049***
Brown	7 (1.45)	9 (1.89)	16 (1.67)	0.25^{ns}
Red pyle	7 (1.45)	2 (0.42)	9 (0.94)	2.78^{ns}
Salmon breasted	6 (1.24)	0 (0.00)	6 (0.63)	ND
χ^2_7	463.8986 ***	590.1116 ***	996.3549 ***	

^{ns} not significant; ***P < 0.0001; ND - not done

Table 3.2 Farmers' preferences to variants of cocks' plumage colour.

Preferred colour	Horro	Jarso	Total	χ^2_1
	N (%)	N (%)	N (%)	
Red	188 (40.00)	191 (42.83)	379 (41.38)	0.0237^{ns}
White	106 (22.55)	118 (26.46)	224 (24.45)	0.6429^{ns}
Black	19 (4.04)	73 (16.37)	92 (10.04)	31.6957***
Silver birchen	76 (16.17)	8 (1.79)	84 (9.17)	55.0476***
Wheaten	29 (6.17)	22 (4.93)	51 (5.57)	0.9608^{ns}
Speckled	15 (3.19)	26 (5.83)	41 (4.48)	2.9512^{ns}
Red-pyle	16 (3.40)	2 (0.45)	18 (1.97)	10.8889**
Black breasted red	19 (4.04)	2 (0.45)	21 (2.29)	13.7619**
Brown	2 (0.43)	4 (0.90)	6 (0.66)	0.6667^{ns}
χ^2_9	683.95***	703.86***	1422.69***	

ns not significant; **P < 0.01; ***P < 0.0001; ND - not done

Farmers were also interviewed to list variants of plumage colour they less prefer. Farmers commonly show less preference to some variants of plumage colour (Table 3.3). Loss of preference may be partly arise from exposure to predators in chickens displaying vivid plumage colour (personal communication) and due to low market demand, which is largely dictated by cultural and religious landscape of the community (Table 3.4). From analysis of farmers preference we found that red is the most preferred plumage colour whereas black is the least. Cultural and religious value of plumage colour considerably varies between the two sites implicating the impact of demographic structure of the two communities. Village chickens show a vast array of plumage colour as it is displayed in Figure 3.1 – 3.3, even at family flock level, which makes them a good resource in genetic mapping of pigmentation traits.

Table 3.3 Plumage colour variants that are less preferred by farmers.

Plumage colour	**Horro**	**Jarso**	**Total**	χ^2_1
	N (%)	N (%)	N (%)	
Black	119 (63.98)	90 (54.22)	209 (59.38)	4.02*
White	42 (22.58)	39 (23.49)	81 (23.01)	0.11^{ns}
Wheaten	7 (3.76)	23 (13.86)	30 (8.52)	8.53**
Multi-coloured	17 (9.14)	7 (4.22)	24 (6.82)	4.17*
Coppery black	1 (0.54)	7 (4.22)	8 (2.27)	4.50*
χ^2_4	251.20***	142.67***	383.54***	

^{ns} not significant; *P < 0.05; **P < 0.01

Table 3.4 Stated reasons for plumage colour preference.

Reasons	**Horro**	**Jarso**	**Total**	χ^2_1
	N (%)	N (%)	N (%)	
Aesthetic value	108 (42.52)	116 (58.29)	224 (49.45)	0.29^{ns}
Market demand	131 (51.57)	78 (39.20)	209 (46.14)	13.44**
Religious and cultural values	15 (5.91)	5 (2.51)	20 (4.42)	5.00*
χ^2_2	89.1102***	95.9497***	171.2185***	

^{ns} not significant; *P < 0.05; **P < 0.01

Plumage colour

We scored variants of plumage colour in hen and cock populations separately. Cocks show high diversity in plumage colour than hens (Table 3.5 versus Table 3.6). A high variation in plumage colour was observed at intrapopulation level. The two chicken populations show a significant difference for some plumage colours of hens.

Table 3.5 Frequency count of plumage colour variants in hens.

Plumage colour	**Horro**	**Jarso**	**Total**	χ^2_1
	N (%)	N (%)	N (%)	
Brown	139 (62.05)	133 (57.33)	272 (59.65)	0.1866^{ns}
Wheaten	35 (15.63)	48 (20.69)	83 (18.20)	0.7049^{ns}
Black	9 (4.02)	34 (14.66)	43 (9.43)	6.0605**
Red	26 (11.61)	7 (3.02)	33 (7.24)	5.0436**
White	15 (6.70)	9 (3.88)	24 (5.26)	0.7516^{ns}
Lavender	0 (0.00)	1 (0.43)	1 (0.22)	ND
χ^2_5	256.5357***	229.5844***	472.3297***	

^{ns} not significant; *P < 0.05; **P < 0.01; ***P < 0.0001

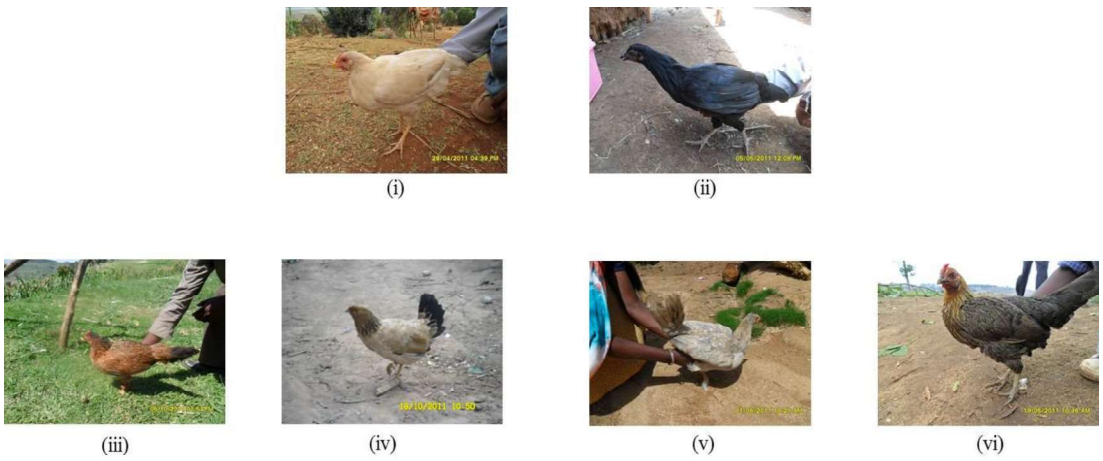


Figure 3.1 Some of the hens sampled: (i) white, (ii) black, (iii) brown, (iv) wheaten brown, (v) lavender and (vi) dull brown

Table 3.6 Frequency count of plumage colour variants in cocks.

Plumage colour	Horro	Jarso	Total	χ^2_1
	N (%)	N (%)	N (%)	
Red	74 (45.96)	64 (46.72)	138 (46.31)	0.0062^{ns}
Brown	32 (19.88)	17 (12.41)	49 (16.44)	1.7281^{ns}
Silver birchen	13 (8.07)	20 (14.60)	33 (11.07)	1.8809^{ns}
Silver birchen	12 (7.45)	20 (14.60)	32 (10.74)	2.3185^{ns}
Red pyle	20 (12.42)	12 (8.76)	32 (10.74)	0.6325^{ns}
White	9 (5.59)	4 (2.92)	13 (4.36)	0.8377^{ns}
Barred	1(0.62)	0 (0.00)	1 (0.34)	ND
χ^2_6	156.1739***	136.1022***	283.4228***	

^{ns} not significant; *P < 0.05; ***P < 0.0001; ND – not done

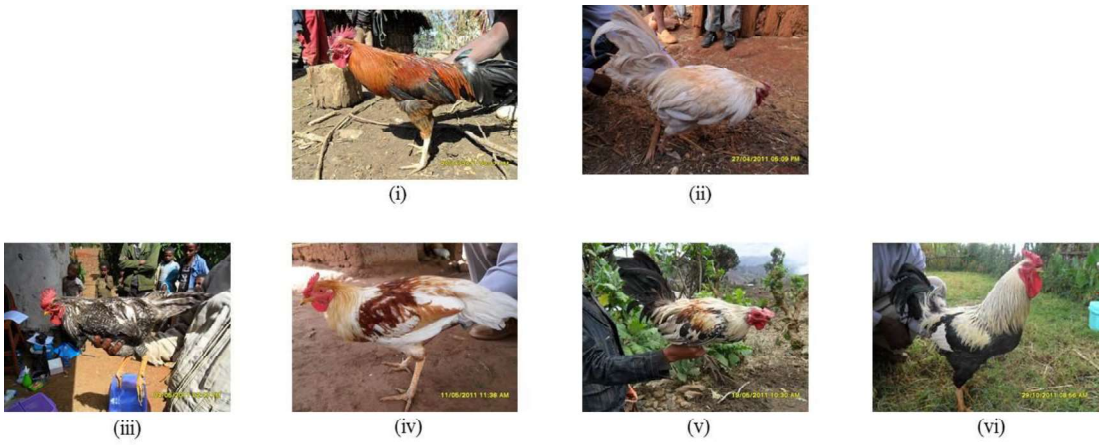


Figure 3.2 Some of the cocks sampled: (i) red, (ii) white, (iii) barred, (iv) red Pyle, (v) golden splashed silver birchen and (vi) silver birchen

(i) (ii)

Figure 3.3 Family flocks in Horro (i) and Jarso (ii).

Discussion

In most instances, plumage colour is used as a morphological marker to describe chicken breeds as in white leghorn; however, this is not the case in village chickens that display a vast array of plumage colours even at family flock level (Figure 3.3 (i) and (ii)). Farmers use plumage colour and pattern as a biological marker to identify individual chickens within their flock. Animal identifiers like tags are entirely absent under subsistence farmers management system, which then make farmers to identify individual chickens based on their plumage colour and pattern (Abdelqader et al., 2007). This shows that traits showing a Mendelian pattern of segregation could serve as animal identifier. It has been also reported that chicken breeds are named based on their morphological attributes (Muchadeyi et al., 2009). Despite this, in most instances, chicken populations are named after their home range and communities maintaining them as in the case of Horro and Jarso chickens.

Like other avian species, plumage colour and pattern in chickens is a complex trait involving the genetic control of many loci (Moore and Smyth, 1972). Plumage colour and pattern shows a wide variation among breeds of chickens (Hellström et al., 2010). However, few mutations have been so far identified (Chang et al., 2007). Structural coloration is mainly divided into iridescent (i.e. the hue varies at different angles of observation) or non-iridescent. Studying the role of melanin in pigment synthesis will help to understand the evolution of basic components of colouration over time to result in a spectacular diversity of plumage colour in the avian species (Shawkey and Hill, 2008). To better understand the genetic basis of plumage

evolution, it is more appropriate to study variants of plumage colour in natural populations (Protas and Patel, 2008).

There are primary and secondary pattern genes that determine the type of plumage. Accordingly, E and Co loci and their alleles have been classified as primary pattern genes while others such as Sg, Pg, and Bg as secondary pattern genes (Kimball, 1952; Moore and Smyth, 1972). The primary pattern genes control distribution of black and red pigments. The primary pattern genes interact with secondary pattern genes that control individual feather's pigmentation and determine the final set of plumage colour (Moore and Smyth, 1972). An extended black mutation E1-12 is the most dominant one among many alleles found in E locus and is primarily responsible for self-black plumage in chickens. Self-black plumage is produced due to increased deposition of eumelanin (Moore and Smyth, 1971). However, barring gene could be a partial inhibitor to black (Warren, 1928). Self-white plumage as in White Leghorns is produced by dominant action of white gene (I) that differs from the E locus. The two forms of white plumage (the dominant I and the recessive c) epistatically interact with other colours to produce different patterns. A recessive white plumage (c) is associated with a truncated transcript of tyrosinase gene (Chang et al., 2007). The c locus more efficiently inhibits pheomelanin than eumelanin; whereas, I locus is relatively weak inhibitor of pheomelanin (Campo, 1997).

Domestication has resulted in a wide range of colours in livestock species. Most of these colour variants may associate with demographic structure of communities maintaining livestock species (Sheppy, 2011). Minor modifications may occur in plumage colour due to environmental impact, however in most instances this is not strong enough to induce variation; therefore morphological diversity of plumage colours is largely the impact of bird's genotype (Paxton, 2009). Plumage colour of the galliforms is highly varied and commonly shows sexual dimorphism. In many genera of the galliforms it is the male which exhibit vividly coloured plumage, with most of the females showing cryptic plumages. The highly conspicuous colour in males is associated with attraction of mates, while the sombre colours seen in hens are used to avoid predators (Sheppy, 2011). However, there is a trade-off between being conspicuous to attract a mate and exposure to predators in cocks.

Due to the nondescript characteristics of village chickens, proportion of different plumage colour across populations may not significantly differ. A large proportion of red plumage in cocks and brown plumage in hens observed in the two chicken populations may reflect the impact of their progenitor – the red junglefowl alleles for plumage colour. A wide-range preference of farmers ascends the diversity of plumage colours (Cabarles et al., 2012). Plumage colour is used as a culling criterion in chickens by some farmers as it has been observed in our study for black plumage and as it has been reported for Nigerian Turkeys (Yakubu et al., 2013). For example, consumers in Ethiopia usually prefer brown chickens and pay higher price, while black plumage is considered as a sign of bad fortune. White birds are considered as carriers of bad spirit inflicting disease, and the communities believe that bad spirits that target a family can be diverted to someone else or to somewhere through white chickens (Aklilu, 2007). Moreover, white plumage is considered as symbol of peace and hence it is the most preferred variant by Beninese farmers, whereas in some places dark colours like the self-black are considered as signs of misfortune. Following this, black chickens are used for magic and red ones are presented as gift for relative's spirit (Faustin et al., 2010). Plumage colour is used by farmers as selection criterion for replacement/breeding cock and determines the market price (Moges et al., 2010).

Plumage preference has a long lasting history in human life (Paxton, 2009). Plumage colour preference may be to some extent influenced by demographic structure (e.g. religion, culture) of the community and this likely leads to re-ranking of plumage colour preferences. For example, white chickens are mostly preferred in north western Ethiopia whereas red is commonly preferred by communities living in western Ethiopian and by some ethnic groups of the southern Ethiopia (Dana et al., 2010). Plumage colour therefore has a multi-dimensional socio-cultural significance.

Conclusion

Our analysis shows that village chickens display a high intrapopulation variation in plumage colour. The difference observed in the proportion of plumage colour variants between the two chicken populations is low implicating the landrace nature of the two chicken populations. Our analysis shows that besides local adaptation,

demographic structure of the community may have shaped the diversity of plumage colour in the two chicken populations.

References

ABDELQADER, A., WOLLNY, C. B. A. & GAULY, M. Characterization of local chicken production systems and their potential under different levels of management practice in Jordan. Tropical Animal Health and Production, 2007, 39(3), 155-164.

AKLILU, H. A. Village poultry in Ethiopia: Socio-technical analysis and learning with farmers. PhD thesis, Wageningen University, the Netherlands, 2007.

CABARLES, J. C., LAMBIO, A. L., VEGA, S. A., CAPITAN, S. S. & MENDIORO, M. S. Distinct morphological features of traditional chickens (*Gallus gallus domesticus* L.) in Western Visayas, Philippines. Animal Genetic Resources, 2012, 51, 73-87.

CAMPO, J. L. The hypostatic genotype of the recessive White Prat breed of chickens. Poultry Science, 1997, 76, 432–436.

CHANG, C-M., FURET, J-P., COVILLE, J-L., COQUERELLE, G., GOURICHON, D. & TIXIER-BOICHARD, M. Quantitative effects of an intronic retroviral insertion on the transcription of the tyrosinase gene in recessive white chickens. Animal Genetics, 2007, 38, 162-167.

DANA, N., VAN DER WAAIJ, L. H., DESSIE, T. & VAN ARENDONK, J. A. M. Production objectives and trait preferences of village poultry producers of Ethiopia: implications for designing breeding schemes utilizing indigenous chicken genetic resources. Tropical Animal Health and Production, 2010, 42, 1519–1529.

DESTA T. T., DESSIE T., BETTRIDGE J., LYNCH S. E., MELESE K., COLLINS M., CHRISTLEY R. M., WIGLEY P., KAISER P., TERFA Z., MWACHARO J. M. & HANOTTE O. Signature of artificial selection and ecological landscape on morphological structures of Ethiopian village chickens. Animal Genetic Resources, 2013, 52, 1-13.

FAO. World Watch List for Domestic Animal Diversity. 3rd edn. (Scherf B.D. (Ed.)), FAO, Rome, Italy, 2000.

FAUSTIN, V., ADÉGBIDI, A. A., GARNETT, S. T., KOUDANDÉ, D. O., AGBO, V. & ZANDER, K. K. Peace, health or fortune? Preferences for chicken traits in rural Benin. Ecological Economics, 69 (9), 2010, 1848-1857.

HELLSTRÖM, A. R., SUNDSTRÖM, E., GUNNARSSON, U., BED'HOM, B., TIXIER-BOICHARD, M., HONAKER, C. F., ... & ANDERSSON, L. Sex-linked barring in chickens is controlled by the CDKN2A/B tumour suppressor locus. Pigment Cell & Melanoma Research, 2010, 23(4), 521-530.

JOHNSON, J. A., & BURNHAM, K. K. Population differentiation and adaptive selection on plumage color distributions in gyrfalcons. RT WatsonTJ CadeM FullerG HuntE Potapov. Gyrfalcons and Ptarmigan in a Changing World. Boise, ID: The Peregrine Fund, 2012, 1-19.

KIMBALL, E. Wild type plumage pattern in the fowl. Journal of Heredity, 1952, 43(3), 129-132.

MOGES, F., TEGEGNE, A. & DESSIE, T. Indigenous chicken production and marketing systems in Ethiopia: Characteristics and opportunities for market-oriented development. IPMS (Improving Productivity and Market Success) of Ethiopian Farmers Project Working Paper 24. Nairobi, Kenya, ILRI, 2010.

MOORE, J. W. & SMYTH, J. R. Melanotic: Key to a phenotypic enigma in the fowl. The Journal of Heredity, 1971, 62, 215-219.

MOORE, J. W. & SMYTH, J. R., Inheritance of the silver-laced Wyandotte plumage pattern. The Journal of Heredity, 1972, 63, 179-184.

MUCHADEYI, F. C., WOLLNY, C. B. A., EDING, H., WEIGEND, S. & SIMIANER, H. Choice of breeding stock, preference of production traits and culling criteria of village chickens among Zimbabwe agro-ecological zones. Tropical Animal Health and Production, 2009, 41(3), 403-412.

PAXTON, E. H. The utility of plumage coloration for taxonomic and ecological studies. The Open Ornithology Journal, 2009, 2, 17-23.

PEARL, R. & BORING, A. M. Some physiological observations regarding plumage patterns. Science, 1914, 143-144.

PROTAS, M. E. & PATEL, N. H. Evolution of coloration patterns. Annual Review of Cell and Developmental Biology, 2008, 24, 425-46.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013, ISBN 3-900051-07-0. http://www.R-project.org/.

RIEGNER, M. F. Parallel evolution of plumage pattern and coloration in birds: Implications for defining avian morphospace. Condor, 2008, 110, 559-614.

SHAWKEY, M. D. & HILL, G. E. Significance of a basal melanin layer to production of non-iridescent structural plumage color: evidence from an amelanotic Steller's jay (*Cyanocitta stelleri*). The Journal of Experimental Biology, 2008, 209, 1245-1250.

SHEPPY, A. The colour of domestication and the designer chicken. Optics and Laser Technology, 2011, 43, 295-301.

SMYTH, J.R. JR. Genetics of plumage, skin and eye pigmentation in chickens. pp. 109–167 in Poultry Breeding and Genetics, edited by R. D. Crawford. Elsevier, Amsterdam, 1990.

WARREN, D. C. Sex-linked characters of poultry. Genetics, 1928, 13, 421-433.

YAKUBU, A., ABIMUKU, H. K., MUSA-AZARA, I. S., IDAHOR, K. O. & AKINSOLA, O. M. Assessment of flock structure, preference in selection and traits of economic importance for domestic turkey (*Meleagris gallopavo*) genetic resources in Nasarawa State, Nigeria. Livestock Research for Rural Development, 2013, 25(1).

Chapter 4

Genetic structure and contrasting demographic history of Ethiopian village chickens

Abstract

African village chickens have a complex history of origin and dispersion. Village chickens have been kept and mildly selected by smallholder farmers for traits of socio-cultural significance. Mating is typically random and uncontrolled. A high phenotypic diversity is often observed owing to lack of strong artificial selection. Village chickens are mainly evolved under the impact of natural selection and this has mainly shaped their genetic structure. However, at a genome-wide level, genetic diversity and population structure of village chickens remains understudied. The genetic structure of two indigenous chicken populations Horro (n = 380) and Jarso (n = 367) was analysed using a high density (600K) SNP array. Principal component and genetic relationship and admixture analyses reveal Horro and Jarso chickens as genetically distinct populations. A high intrapopulation genetic diversity and weak geographic substructuring were observed at village level in the two chicken populations. However, admixture analysis reveals a genetic substructure at marketshed level in Horro, indicating the impact of trading-network on genetic structure of village chicken. Current effective population size (Ne) is similar for the two chicken populations; however, historical Ne is much larger in Jarso. Two distinct trends in Ne were found in Jarso chickens, suggesting two independent demographic histories in this population. Difference in origin, routes of chicken introduction and demographic history and anthropogenic effects may have contributed to variation in Ne and have led to genetic divergence. Our results demonstrate that large number of samples and informative genetic markers are required to map the genetic structure of nondescript village chickens at fine-scale level. The broad genetic base of the two chicken populations suggests that this genetic diversity would serve as a substrate to improve performance of village chickens.

Key words: demographic history, genetic admixture, natural selection, ecological variation, panmictic populations

Introduction

Village chickens constitute a sizeable portion of the extant global chicken diversity (FAO, 2000). This diversity is the consequence of a cumulative effect of founder events (domestication and dispersion), genetic drift, natural and/or artificial selection (Granevitze et al., 2007). Village chickens support the livelihoods of millions of subsistence-oriented smallholder farmers across Africa, Asia and Latin America. Unlike commercial chickens, which have been intensively selected for production traits (Muir et al, 2008), and fancy/traditional breeds that have been bred to conform to pre-defined standards, village chickens serve multiple functions ranging from subsistence to socio-cultural (Desta et al., 2013).

Genetic diversity of village chickens is higher than commercial chickens (e.g., Lyimo et al., 2014). Village chicken populations are good models to study the evolutionary history (origins and dispersion, local adaptation and demographic history) of domesticates. Village chickens are typically subjected to mild or no artificial selection and occur in almost all agro-ecological zones. High flock turnover is common due to predation, mortality associated with disease outbreak, portability and their use as trade and gift item.

Village chickens subsist on scavenging which leads to frequent flocks intermix and uncontrolled mating. Although mild artificial selection is practiced by smallholder farmers in favour of preferred phenotypes (Dana et al., 2010, Melesse and Negesse, 2011; Desta et al., 2013), it is thought not too intense to create phenotypically homogenous populations. In village chickens, the impact of natural selection for local adaptation most likely surpasses the influence of artificial selection (Wragg et al., 2012). The combination of these factors results in a mosaic phenotypes and propensity to adapt to a wide range of production environments. It is demanding to define the nondescript village chicken populations as a breed only based on their phenotypic appearance. The local communities describe and name village chickens based on their home range and after ethnic groups maintaining them.

There is growing evidence that corroborates as the history of African chicken involved several episodes of introduction and dispersion (Blench & Macdonald,

2000; Muchadeyi et al., 2008; Mwacharo et al., 2011; Mwacharo et al., 2013a). Mitochondrial DNA (mtDNA) D-loop analyses reveal two main African haplogroups, which may have distinct Asian origins: the Indian subcontinent and/or the South-East/East Asia, including the Indonesian islands (Muchadeyi et al., 2008; Mwacharo et al., 2011). Based on archaeological, linguistic and historical evidences two entry points of domestic chicken into the African continent have been suggested: the North-East (Egypt) and the East African coast (Blench & Macdonald, 2000).

Several studies have examined the genetic diversity of African village chicken using microsatellite markers. Most of the studies report weak genetic divergence amongst village chicken populations (Muchadeyi et al., 2007; Osei-Amponsah et al., 2010; Dana, 2011; Lyimo et al., 2014). However, a few regional studies reveal a geographic substructuring among African village chickens. For example, across West African countries (Cameroon, Benin, Ghana and Côte d'Ivoire) Leroy et al. (2012) identified three genetic groups matching with the existing farming systems. Moreover, a study involving Kenya, Uganda, Ethiopia and Sudan (Mwacharo et al., 2013b) also provides evidence for three distinct genepools.

Here, we investigate at genome-wide level the genetic diversity and population structure of two Ethiopian village chicken populations (Horro and Jarso) using Affymetrix 600K SNP array (Kranis et al., 2013). Our findings show that Horro and Jarso chickens are genetically distinct and exhibit a high intrapopulation genetic diversity. Moreover, we reveal a fine scale genetic substructure at population level, with evidence of considerable panmixia at village level. A considerable level of genetic substructure was observed down to marketshed level in Horro, while two distinct historical trends of effective population size have been identified in Jarso chickens.

Materials and methods

Study populations and sampling strategy

This study was conducted in western (Horro) and eastern (Jarso) Ethiopia. The two sites are ~ 870 km apart. The two sites were deliberately chosen to minimize the likelihood of interpopulation gene flow and to study the impact of contrasting agro-ecological and socio-cultural landscape on the genetic structure of the two chicken populations (see Desta et al. (2013) for further detail). The two populations have been managed for centuries by smallholder farmers, with farming practice of Jarso thought to be of an ancient origin (Desta et al., 2013). In each site, two marketsheds (villages sharing a common trading-network) were selected and from each marketshed two villages were sampled, giving a total of eight villages. The villages sampled are (i) Didibe Chistana (DC) and Doyo Beriso (DB) (marketshed one) and Harro Aga (HA) and Bonne Abunna (BA) (marketshed two) from Horro; and (ii) Afgug (AF) and Bedhasa (BD) (marketshed one) and Latin Fedho (LF) and Aman (AM) (marketshed two) from Jarso. Two adult chickens (birds more than 6 months old) with no known recent history of pedigree relationship were sampled following householder's consent.

Blood samples was collected and spotted on the FTA Cards® (Whatman Biosciences) over four sampling sessions (April to May and October to November in 2011 and 2012). DNA was extracted from 760 samples following the method 4 as suggested in Smith and Burgoyne (2004). The DNA was genotyped using the Axiom® 600K Affymetrix SNP array (Kranis et al., 2013) by the Ark Genomics facility of The Roslin Institute, University of Edinburgh (http://www.ark-genomics.org/).

Data pruning and statistical analysis

Selection of SNPs for downstream analysis involves two quality control (QC) steps performed using GenABEL package (Aulchenko et al., 2007) of the R (R Core Team, 2013). Prior to the QC, among 580961 SNPs assayed on the array, unmapped and non-autosomal SNPs were removed leaving 546120 autosomal SNPs (GGA1–

28). The entire dataset was used in the first QC performed using the check.marker function of GenABEL based on these criteria: SNP and sample calling rate ≥ 97.5%, minor allelic frequency (MAF) ≥ 5% and identical by state (IBS) threshold ≤ 90%. In IBS calculation all markers were included. When the IBS between a pair of samples exceeded the threshold of 90%, a bird with the lowest calling rate was excluded. The first QC removed 86727 SNPs showing a MAF of less than 5% and 21587 SNPs for low calling rate. Thirteen chickens (four from Horro and nine from Jarso) were also removed due to high IBS. The second QC for Hardy–Weinberg equilibrium (HWE) test was performed at population level using the HWE.show function of GenABEL with a cut-off p-value of 0.001. The HWE test excluded 36773 and 35154 SNPs from Horro and Jarso chickens respectively, which include 7459 SNPs common to both populations. The two QC steps retained 375213 SNPs and 747 chickens (380 from Horro and 367 from Jarso).

Descriptive statistics (MAF, observed (Ho) and expected (He) heterozygosity and inbreeding coefficient (F)) were calculated at village, markershed and population level using the descriptives.marker and perid.summary functions of GenABEL. A panmictic index (1 − F) was derived from F values calculated. At population level, the amount of genetic diversity was also assessed based on allelic richness and proportion of private allele's calculated using ADZE software (Szpiech et al., 2008). Descriptive values were tested for their significant difference between the two chicken populations using the Welch t-test (Welch, 1947) as implemented in R. Pairwise F_{ST} (Wright, 1951) was calculated using a custom R script. The hierarchical distribution of intra and interpopulation genetic variation was inferred from the analysis of molecular variance (AMOVA) using pegas package (Paradis, 2010) of the R.

Global and fine-scale population structure was inferred using principal components analysis (PCA) performed both at the entire dataset and population level using ade4 package (Dray & Dufour, 2007) of the R. The optim.a.score function of the R package adegenet (Jombart, 2008) was used to identify the optimal number of principal components (PCs) (Figure S4.2). The optimal number of genetic clusters was identified using the Bayesian Information Criterion (BIC) (Figure S4.3) as

implemented in the find.clusters function of adegenet. Genetic admixture analysis was performed using the dapc function of adegenet.

The IBS (Identity by State) matrix (not weighted for allelic frequency) was calculated using the IBS function of GenABEL. The genetic distance matrix (1− IBS) was then imported to MEGA5 (Tamura et al., 2011) to calculate the average inter, intrapopulation and net genetic distance. A dendrogram was constructed from a pairwise genetic distance using the neighbour-joining (NJ) algorithm in MEGA5, and visualized using FigTree v1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/). Geographic distance (km) between villages was calculated using fossil package (Vavrek, 2011) of the R from GPS (geographic positioning system) coordinates of each homestead. A Mantel test was performed to investigate the relationships between genetic and log transformed geographic distance in IBDWS v3.23 (Jensen et al., 2005).

Linkage disequilibrium (LD) was calculated using the r2fast function of GenABEL. Effective population size (Ne) was estimated using SNeP software (Barbato et al., 2015) by Mario Barbato, a PhD student at Cardiff University. For Ne analysis the dataset was phased using fastPHASE (Scheet & Stephens, 2006). The linkage disequilibrium (LD) between each pair of syntenic SNPs with a physical distance between 5Kb and 1Mb was calculated using Hill and Robertson squared correlation r^2 (Hill & Robertson, 1968). The resulting r^2 estimates were then binned in distance classes and for each bin the average r^2 and distance values were calculated. For each bin the historical effective population size (Ne) estimate was calculated using E (r^2) = (1+4cNe)-1+n-1 (Sved, 1971; Weir & Hill, 1980) where n corrects r^2 for finite sample size and equals 2*sample size. The recombination rate c is measured in Morgan and was estimated from the physical distance according to the approximation 0.4Mb ~ 1cM and applied in the formula as c = c [(1-c/2) (1-c)-2] (Sved, 1971). The time point calculated as generations ago (t) for each bin was calculated as t = 1/ (2c) (Hayes et al., 2003).

Results

Quality control and summary statistics

The average SNP and sample calling rates were greater than 99% for the entire dataset, i.e., for all the genotyped chickens (n = 760) and all SNPs assayed on the array (n = 580961). In the entire dataset 97.56% (532804/546120) of the autosomal SNPs were polymorphic. At population level 96.76% (528426/546120) and 96.92% (529284/546120) of the autosomal SNPs were polymorphic in Horro and Jarso chickens respectively. MAF of the SNPs that passed the QC exceeds 20% (Figure 4.2) in 55.6% (208626/375213) of the SNPs in Horro chickens and in 51.2% (192108/375213) of the SNPs in Jarso chickens; however these proportions are not significantly different (χ^2_1 = 0.9500, P = 0.6703). The physical distance between consecutive syntenic SNPs used in this analysis (n = 375213) shows a considerable variation among the autosomes and within each autosome and is presented in Table S4.1.

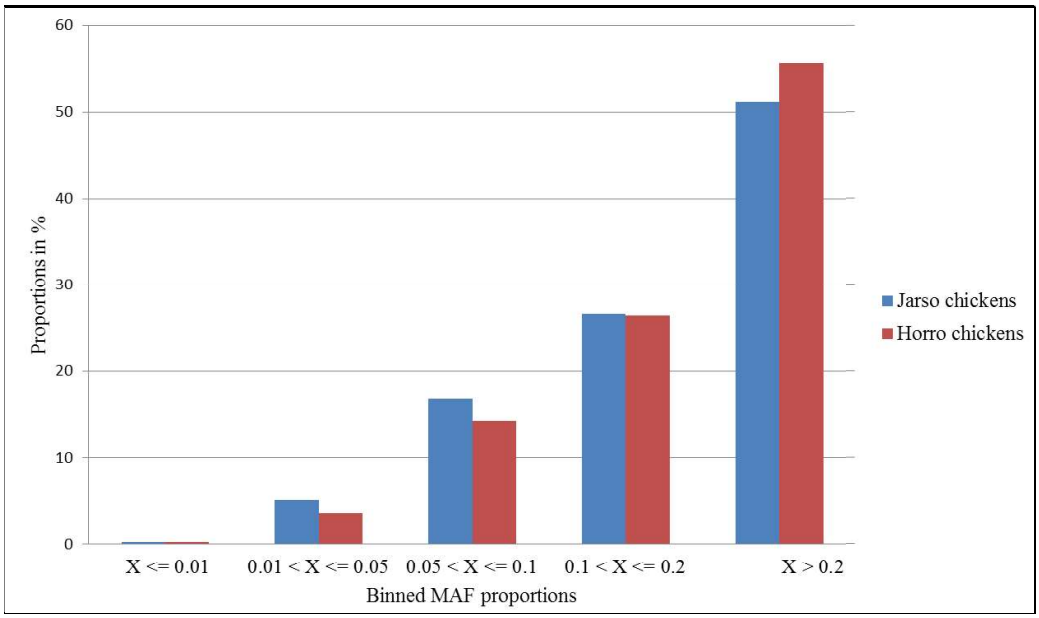


Figure 4.1 The binned proportion of MAF in Horro and Jarso chickens.

Intrapopulation genetic diversity and interpopulation genetic divergence

A summary statistics for indicators of genetic diversity at village, marketshed and population level is presented in Table 4.1. The average allelic richness (AR) is identical between the two chicken populations (AR ~ 1.984; t736.821 = 0.0478, P = 0.9619). The mean private allelic richness for Horro (0.0113) and Jarso (0.0117) do not also significantly differ ($t_{728.845}$ = 0.145, P = 0.8847). However, the mean observed heterozygosity for Horro (0.299 ± 0.04) and Jarso chickens (0.293 ± 0.04) significantly differ ($t_{735.104}$ = 2.3758, P = 0.0177). Indices of genetic diversity for cock and hen population are presented in Table S4.2. We calculate inbreeding coefficient for two datasets containing either two or single bird from each household. However, the inbreeding coefficient of the two datasets is not significantly different both in Horro and Jarso chickens (P > 0.9, Supplementary Information of the Appendix 4). Heatmap plots constructed from an IBS matrix at village level (See the Supplementary Information for details) show that two chickens sampled from the same household have not displayed a higher genetic relatedness than pair of birds sampled from different households (Figure S4.1a–h).

Table 4.1 Indicators of genetic diversity in Horro and Jarso chickens.

Population	*n*	*IBS*	*F*	H_e	H_o
Horro	380	0.723 (0.013)	0.082 (0.112)	0.332 (0.00003)	0.304 (0.038)
DC	95	0.721 (0.015)	0.085 (0.112)	0.336 (0.00004)	0.307 (0.037)
DB	94	0.723 (0.013)	0.075 (0.121)	0.333 (0.00003)	0.308 (0.040)
Horro 1	189	0.721(0.013)	0.079(0.116)	0.334	0.308(0.039)
HA	95	0.725 (0.016)	0.089 (0.125)	0.329 (0.00003)	0.300 (0.041)
BA	96	0.728 (0.016)	0.075 (0.098)	0.327 (0.00002)	0.303 (0.032)
Horro 2	191	0.726(0.014)	0.082(0.112)	0.328	0.301(0.037)
Jarso	367	0.734 (0.016)	0.085 (0.124)	0.316 (0.00003)	0.290 (0.041)
AF	89	0.741 (0.019)	0.068 (0.116)	0.311 (0.00003)	0.290 (0.036)
BD	89	0.738 (0.019)	0.099 (0.124)	0.311 (0.00005)	0.280 (0.039)
Jarso 1	178	0.739(0.017)	0.083(0.121)	0.311	0.285(0.038)
LF	96	0.728 (0.017)	0.062 (0.116)	0.329 (0.00003)	0.308 (0.038)
AM	93	0.738 (0.017)	0.109 (0.135)	0.313 (0.00003)	0.279 (0.042)
Jarso 2	189	0.731(0.016)	0.085(0.134)	0.321	0.294(0.043)
Overall	747	0.709 (0.023)	0.129 (0.112)	0.341 (0.00003)	0.297 (0.040)

Results for identity by state (*IBS*), inbreeding coefficient (*F*), expected (H_e) and observed (H_o) heterozygosities with mean and the standard deviation in parentheses. Horro1 represents DC & DB while Horro2 HA and BA; Jarso1 includes AF & BD and Jarso2 LF & AM.

Fixation indices and analysis of molecular variance

Heterozygote deficiency in the entire dataset (F_{IT}) was 0.129. The inbreeding coefficient (F_{IS}) was positive both in Horro (0.082) and Jarso chickens (0.085), indicating a mild deficiency of heterozygotes. High panmictic index was obtained at population level (91.9% and 91.5% for Horro and Jarso respectively) compared to the entire dataset (87.1%), and a little increase in panmictic index was observed at marketshed and village level. Pairwise estimate of F_{ST} (Wright, 1978) indicates a weak genetic divergence (F_{ST} = 0.042) between Horro and Jarso chickens. At the marketshed and village levels F_{ST} value tends to decrease (Table S4.4 & 5), with little divergence at marketshed and village level. AMOVA indicates that 22.92% of the total genetic variation was due to interpopulation genetic divergence, while 77.08% was explained by intrapopulation variation. Genetic variation among villages explained 1.56% and 0.122% of the total variation in Horro and Jarso chickens respectively however the difference is not statistically significant (χ^2_1 = 0.1167, P = 0.2673). AMOVA results indicate that genetic variation at population level is largely attributed to individual level differences. There is a considerable discrepancy between the calculated values of F_{ST} and AMOVA. Unlike F_{ST} that deals with a mean difference in allelic frequency, AMOVA deals with dispersion of genetic distance values around their mean. A genetic distance between pair of individuals sampled from different populations in most instances can be higher than a genetic distance between pairs of individuals belonging to the same population. This most likely makes the interpopulation level genetic distance to disperse widely around its mean compared to the intrapopulation genetic distance, which then increases the genetic variation among populations. Among populations variation may increase as the sample size increases. Although, no discussion has been made, a comparable finding was reported for a global F_{ST} (0.1042) and an AMOVA (28.6%) performed among seven geographical regions in the work of Ding & Kullo (2011) using 158 SNPs and 938 samples. F_{ST} is sensitive to variation in sample size and it increases with as the sample size reduces (Sinclair & Hobbs, 2009; Willing et al., 2012). A large number of samples included in this analysis may have reduced the F_{ST} value. The large number of markers used may also impose some impact.

Population structure

Principal component analysis assigned all the chickens to their respective population of origin. The BIC statistics also revealed the optimal number of clusters to be two. The first PCA axis (PC1) accounts for 9.18% of the total variation and it clearly separates the two chicken populations (Figure 4.2). Chickens in each population are scattered along the PC2 axis. The remaining axes including the second PC however each accounted for less than 1% of the total variation, implicating a high intra-population genetic variation. The low proportion of variance explained by the first PC axis may be attributed to the panmictic nature of the two chicken populations and the large number of samples included in the analysis. A PCA performed at population level (Figure S4.4a & b) reveals an irregular distribution pattern of individual chickens, indicating a weak intrapopulation genetic substructure both at marketshed and village level. A PCA performed using a random subset of Horro (n = 25) and Jarso (n = 25) chickens has increased the amount of variance explained by the first PC from 9.18 to 10.97% and the second PC from 0.92 to 2.71% (Figure S4.4c). Intuitively, this indicates inverse relationship between sample size and the proportion of genetic variation explained by each PC. As the sample size increases the level of genetic variation increases at decreasing rate than the number of PCs, which then reduces the proportion of variance assigned to each PC.

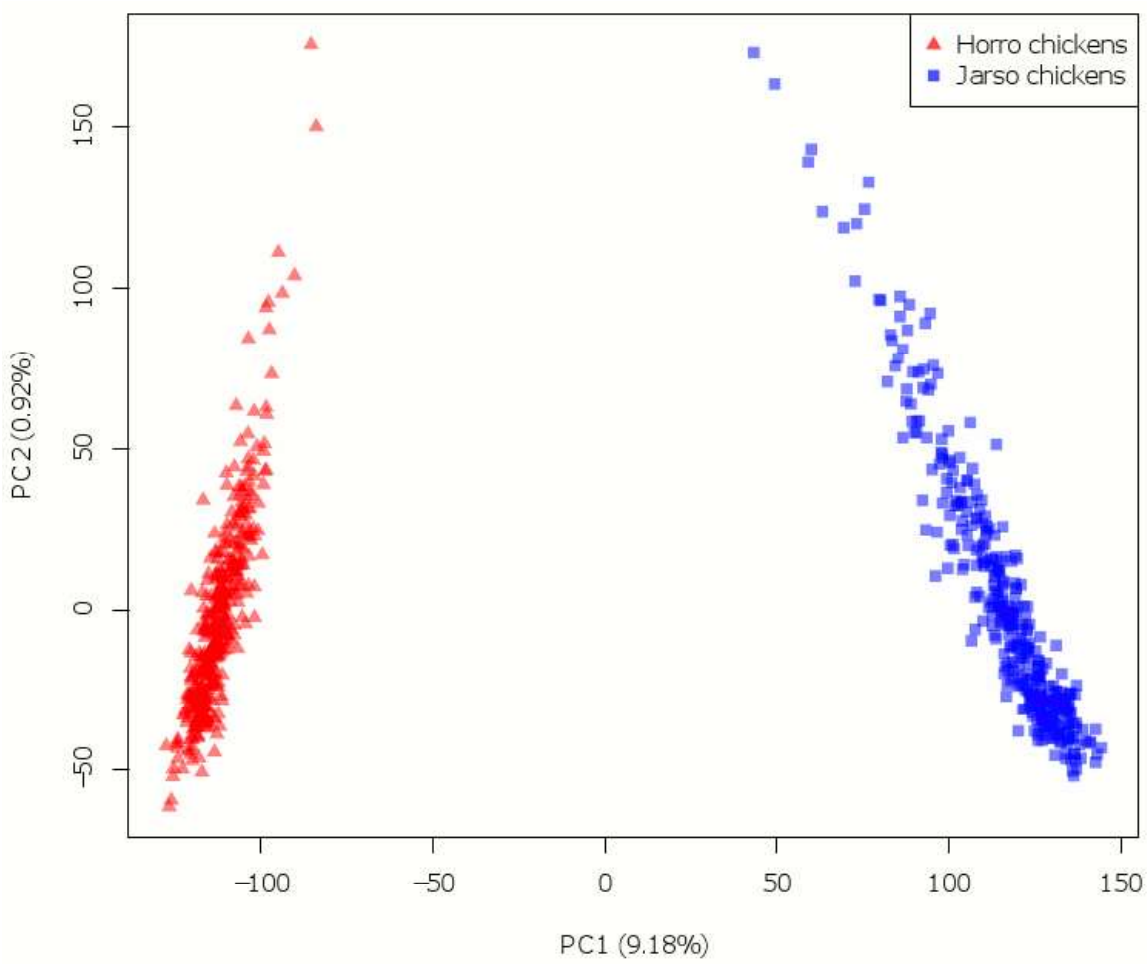


Figure 4.2 Clustering of Horro and Jarso chickens using PCA.

Genetic admixture

A genetic admixture analysis performed using DAPC also reveals the optimal number of genetic clusters (K) to be two, each representing a single population. Membership coefficient matrix generated from a genetic admixture analysis does not provide any evidence of admixture between Horro and Jarso chickens. At K = 4, the DAPC plot indicates a common genetic background peculiar to each marketshed in Horro chickens; however, only Lafin Fedho village in Jarso tend to show different genetic structure. At K = 8, no further geographical substructuring was observed following the number of village sampled (Figure 4.3). A genetic admixture analysis performed at population level also shows some evidence of genetic substructure at marketshed level in Horro chickens (Figure S4.5a & b).

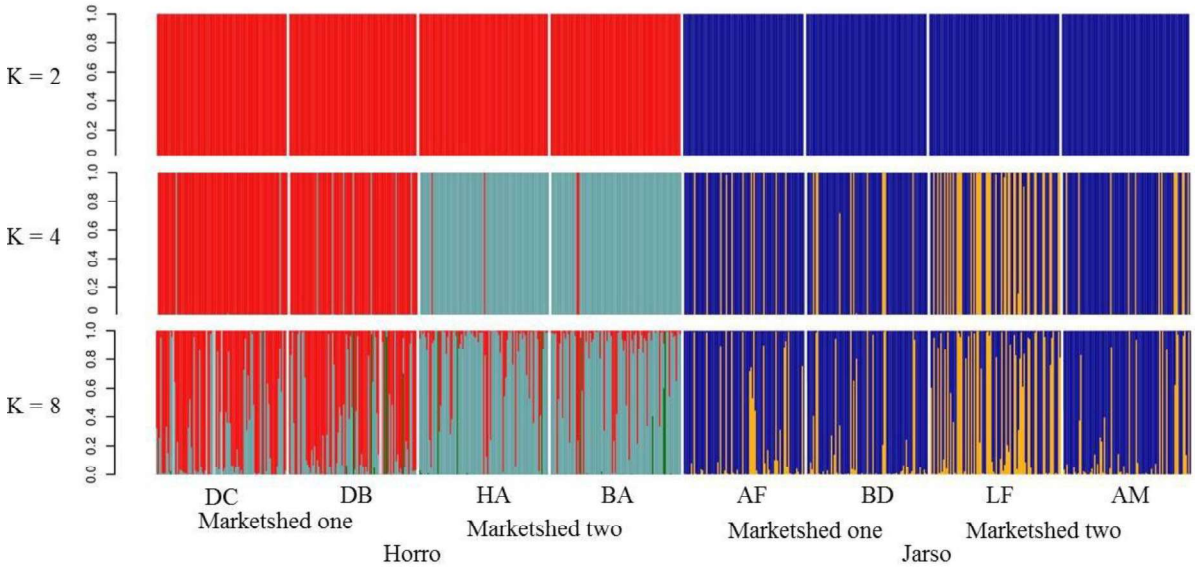


Figure 4.3 Genetic admixture map of Horro and Jarso chickens.

Genetic distance and phylogenetic relationship

The average interpopulation genetic distance was 0.310, with a corresponding average net interpopulation genetic distance of 0.038. The mean intrapopulation genetic distance for Horro and Jarso chickens was 0.277 and 0.266, respectively. Genetic distance between village and marketshed is presented in Table S4.5 and Table S4.6 respectively. A Mantle test performed between genetic (averaged at village level) and a log-transformed geographic distance returned a positive but non-significant correlation in Horro chickens (Figure S4.6a); while this is negative and significant in Jarso chickens (Figure S4.6b). Jarso is dominated by a highly rugged terrain that likely limit flock movement even among closely located households, which then leads to a fine-scale genetic divergence among closely located households. In contrast, small proportion of the Jarso has undulating terrain that facilitates flock intermix even among distantly located households, which consequently may result in high genetic diversity and an associated low genetic divergence among distantly located households. A Mantle test performed among individual chickens using ade4 package also shows a significant positive correlation ($r = 0.5$, $P = 0.0001$) between genetic and geographic distance.

In line with PCA and genetic admixture analysis, the unrooted NJ dendrogram (Figure 4.4) constructed from a pairwise genetic distance matrix among individual

chickens clustered each chicken to its respective population of origin. Dendrograms were constructed at population and marketshed level and are displayed in Figure S4.7a–f. The dendrograms constructed at population and marketshed levels show clades of some chickens, which implicates a fine-scale genetic substructure congruent to the genetic structure found from dapc analysis of Horro chickens.

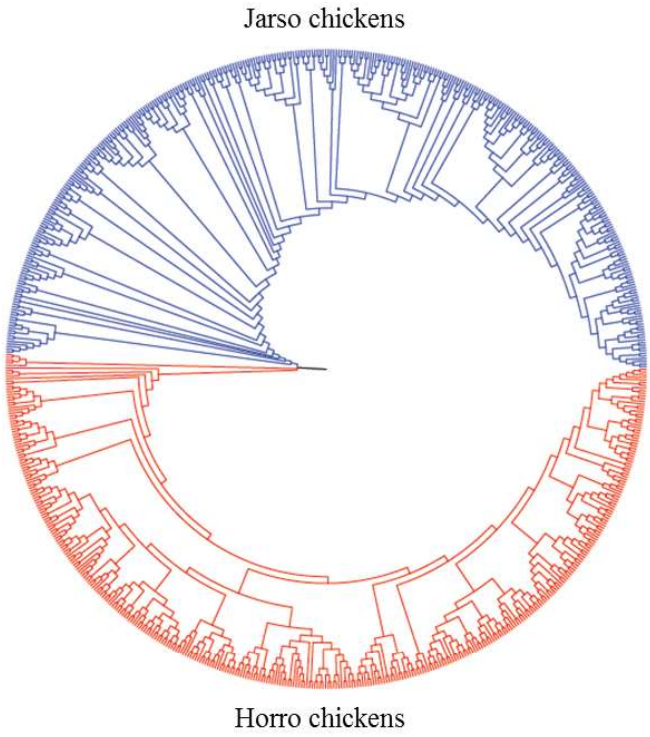


Figure 4.4 Dendrogram illustrating relationship (IBS) within and between Horro and Jarso chickens.

Linkage disequilibrium

The mean value of binned linkage disequilibrium (r^2) was higher in Jarso than Horro chicken across all the autosomes (Figure 4.5 and Figure S4.8a–c). The standard deviation of the mean r^2 value was higher than its mean for all binned distances across the autosomes. The size of LD block negatively correlates with the physical size of the autosomes. The LD value calculated as r^2 was lower in the microchromosomes (GGA11–28) than both macrochromosomes (GGA1–5, $t_{20.02}$ = 8.5662, P < 0.0001) and the intermediate-sized autosomes (GGA6–10, $t_{7.9}$ = 3.7696, P = 0.0056), which supports previous findings (e.g., Megens et al., 2009). However, no significant difference was found for average value of binned LD calculated in

macrochromosomes and intermediate-sized-chromosomes ($t_{4.403}$ = 1.4242, P = 0.2212). The commonly used r^2 values of ≥ 0.3 for genetic mapping studies (e.g. Aerts et al., 2007; Wragg et al., 2012) did not extend beyond 5Kb in the two chicken populations (data not shown), which indicates that more than 200K evenly spaced SNPs are required for genetic map studies in village chickens.

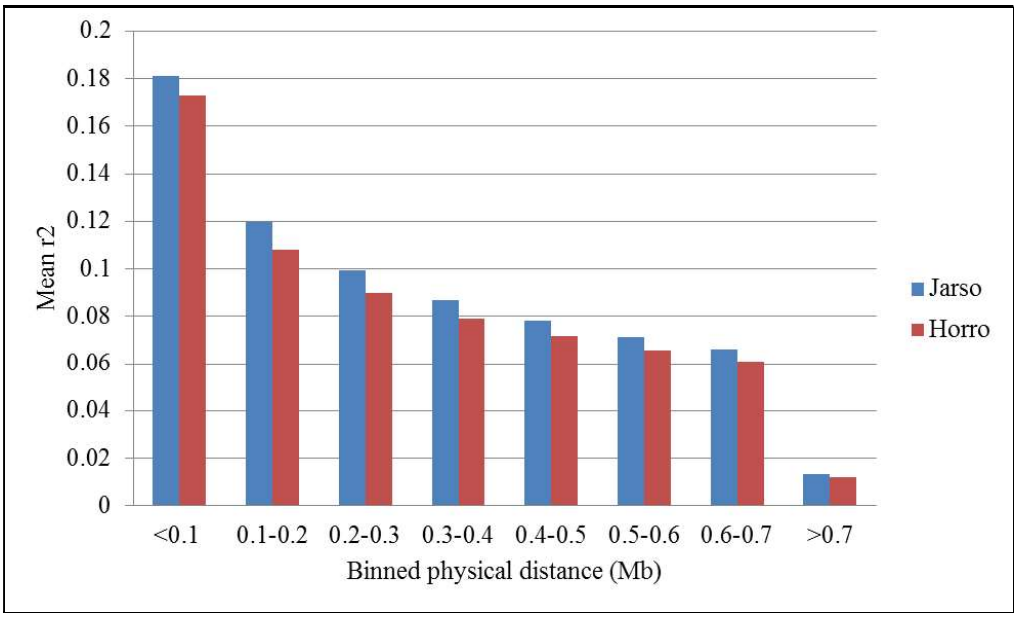


Figure 4.5 Mean r2 for all the autosomes (GGA 1–28) in Horro and Jarso chickens.

Effective population size

Past effective population size was larger in Jarso than Horro chickens (Figure 4.6 & 4.7); however, the reduction in effective population size is more evident in Jarso chickens (except Aman village). The Ne trend observed in Aman village of Jarso is comparable to Horro chickens (Figure 4.7). The large historical effective population size observed in Jarso chickens may indicate a long breeding history of this population. Effective population size fluctuates in the distant past; however, this trend has gradually declined in recent times. An improved management provided by farmers following domestication may have reduced the fluctuation in Ne. The pattern of Ne was homogenous among the four villages in Horro. However, a unique pattern was observed in Aman village of Jarso (Figure 4.7). Similarly, the marketshed from which Aman is sampled also show a different trend (Figure S4.9). Our Ne estimate

indicates that the evolutionary history of the Jarso chickens may have involved two separate demographic events.

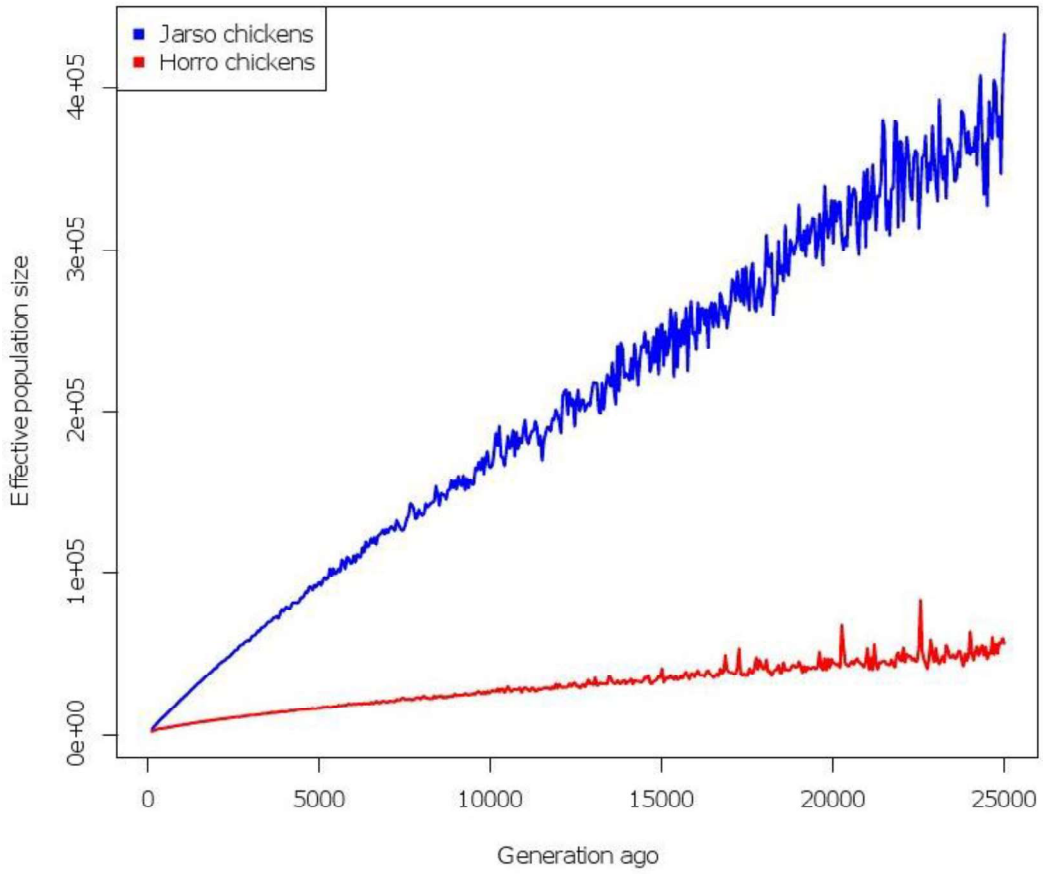


Figure 4.6 Trend of effective population size in Horro and Jarso chickens.

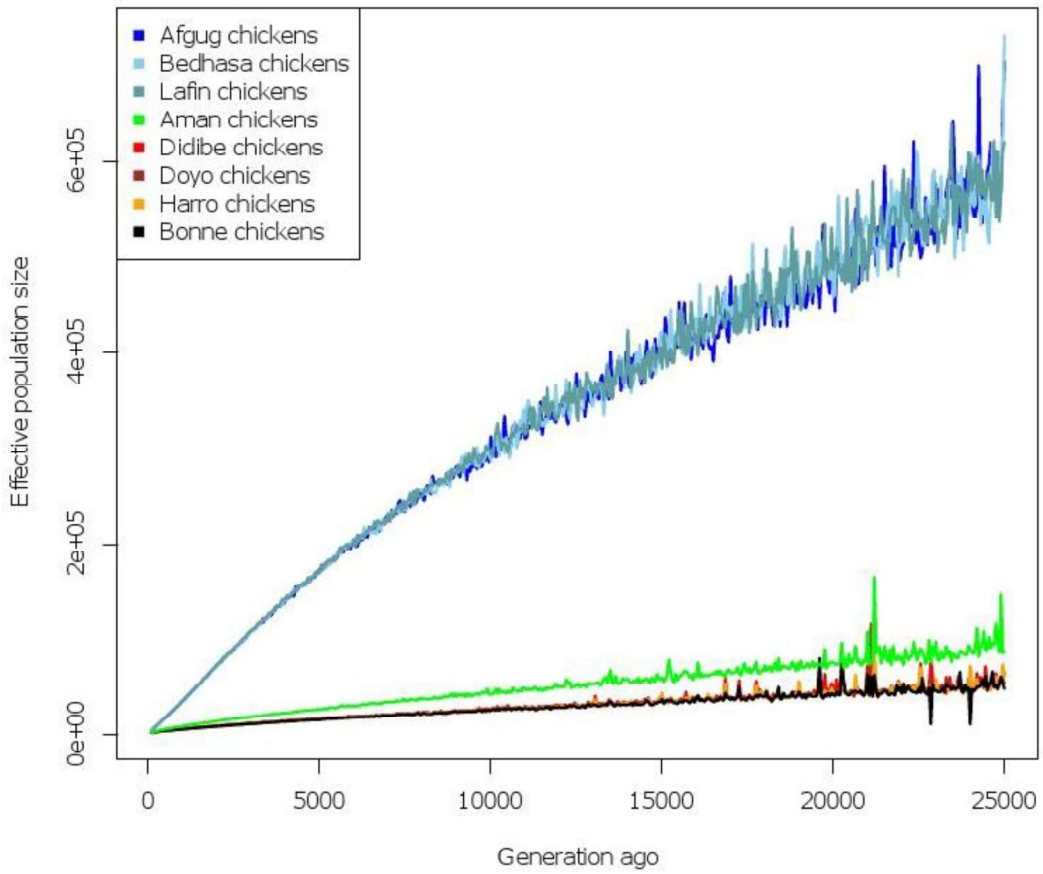


Figure 4.7 Village level effective population size estimate for Horro and Jarso chickens. In the legend the first four villages belonged to Jarso while the remaining four villages belonged to Horro.

Discussion

Unlike uniparental genetic markers, like mitochondrial DNA, nuclear markers are inherited bi-parentally and are subject to meiotic recombination, making them markers of choice for genetic diversity study. Nonetheless, even among nuclear markers there is variation in their informativeness. For instance, compared to microsatellites, SNPs are bi-allelic and are therefore less informative at locus level. However, SNPs are becoming increasingly popular in genetic diversity studies due to their abundance, amenability to high-throughput genotyping and high resolution power.

Studies on village chickens have been largely concentrated on phenotypic characteristics. Absence of population substructure among village chickens have

been reported by a number of studies conducted using microsatellites markers (e.g. Muchadeyi et al., 2007; Osei-Amponsah et al., 2010; Dana, 2011; Lyimo et al., 2014). Most of these studies were based on a limited number of microsatellite markers ($n \leq 30$) largely selected from the ISAG/FAO panel (http://www.fao.org/docrep/014/i2413e/i2413e00.pdf). Though highly polymorphic, thirty microsatellite markers may not be adequate enough to unlock the genetic structure of village chickens that is typically characterized by panmixia and high intrapopulation genetic diversity.

A high intrapopulation genetic variation is common among village chickens (Muchadeyi et al., 2007; Osei-Amponsah et al., 2010; Dana, 2011; Lyimo et al., 2014). This most likely blurs the signals of genetic divergence among populations (Wilson et al., 2008). Genetic diversity in chickens is elevated by high rate of recombination, particularly in the microchromosomes (Hillier et al., 2004; Burt, 2005; Ellegren, 2005) and in the recombination hotspot regions of a genome (Spencer et al., 2006). A high rate of recombination leads to accumulation of several recombination events within a short timespan by breaking extended linkage disequilibrium blocks. Moreover, a prolific reproductive capacity of chickens leads to rapid fixation of de novo mutations and likely enables them to purge out population genetic bottlenecks that have been introduced in the course of their demography history. Random mating, natural selection, high reproductive rate, short generation interval and high cock to hen ratio have also enabled village chickens to maintain their genetic diversity. The prolific characteristics of chickens leave little room for genetic drift to act, which may otherwise gradually eliminates intrapopulation genetic variation by introducing genetic divergence.

Demographic characteristics such as sex structure, flock size and social hierarchy can alter the genetic structure of village chickens. For example, in polygynous species like domestic chickens socially dominant birds may produce more offspring in their lifetime than their submissive conspecifics. There may also be preferential (assortative) mating in which sexual mates are recruited based on their morphological appearance. Assortative mating leads to change in allelic frequencies (Hedrick, 2011), which could then alter the genetic structure of populations. Moreover, variation in fertility rates enables some individuals to produce more

offspring than their peers, which reduces genetic variation even in panmictic populations.

Although artificial selection has been negligible, a mild level of inbreeding is common in village chickens (e.g. Muchadeyi et al., 2007; Mtileni et al., 2011; Mwacharo et al., 2013b) and other panmictic populations (Hamilton, 2009), which commonly associates with Wahlund effect (Wahlund, 1928). Mild heterozygote deficiency may also arise due to an ascertainment bias of SNP chips, given that arrays are usually developed based on genomic information of commercial populations (e.g., Kranis et al., 2013). However, the mild level of inbreeding observed might be inconsequential given the high level of genetic diversity in the two chicken populations.

The mild selection practiced by smallholder farmers for millennia has not been conducted with the intention of breed formation; it has rather focused on selection of traits of socio-cultural significance. Mild selection is expected to have had little impact on genetic diversity and structure of village chickens, as these traits, which are typically morphological (e.g. comb type), are thought to be under the control of a few genes. Selection intensity of these traits may vary among communities based on their preferences (Dana et al., 2010; Melesse and Negesse, 2011; Desta et al., 2013). For example, a high socio-cultural value attached to comb shape in Ethiopia (Dana et al., 2010) may be of little importance elsewhere. However, the cumulative effect of selection of visual traits for many generations especially in a population experiencing a minimum gene flow may have a considerable impact on genetic structure. Furthermore, there may be a pleiotropic effect involving loci subjected to artificial selection, which may negative impact population's fitness (see Desta et al. (2013) for a review), which then may reduce genetic vigour and diversity of a population.

The genetic divergence found between Horro and Jarso chickens may indicate their different origin and contrasting demographic history. The long breeding history of village chickens (Fuller et al., 2011) could result in a large number of recombination events and when this combined with extensive interbreeding it may result in a considerably admixed genetic structure. Variation in genetic structure between the

two chicken populations could be also a consequence of different routes of introduction of domestic chicken to Africa (Blench & Macdonald, 2000; Mwacharo et al., 2013a). Demographic structure of the farming communities dictates the use value of chickens and this may have led to different breeding histories, and consequently has shaped the population structure. Moreover, as it was revealed by a Mantle test (see the Supplementary Information), absence of gene flow between the two populations owing to their distant geographic location may also have led to genetic divergence.

The extent of genetic diversity may vary among populations depending on their demographic history (Cuc et al., 2010), although a phylogenetic relationship is often associated with geographical proximity (e.g., Ya-Bo et al., 2006). A moderate genetic divergence we found indicates a high level of genetic diversity within Horro and Jarso chickens. This divergence might be partly a consequence of contrasting production environments and difference in the demographic structure of the two communities (Desta et al., 2013). Several factors could impact the extent of genetic divergence among populations, including population expansions and genetic bottlenecks, time since divergence, effective population size, geographical proximity and ecological variability. For example, geographical proximity of the conspecifics, ecological similarity and socio-cultural homogeneity of the communities may facilitate a continuous gene flow, which then limits genetic divergence and leads to formation of a genetically homogeneous population.

Conclusion

Our hypothesis that set a prior as individual chickens would cluster to their respective origin has been supported by PCA, genetic admixture and genetic relationship analyses, which corroborates the power of using a large number of markers. This has enabled us to unlock the genetic structure of village chickens characterized by a cryptic genetic substructure and an extensive genetic admixture. Although genetic clustering was found to associate with geographical proximity, this genetic divergence may have also occurred due to limited gene flow between the two populations and/or differences in their demographic history. Moreover, ancestral chicken populations may have arrived in the two regions through different routes

and/or at different times. Alternatively, the two chicken populations may have been derived from different origins and/or have been established and developed under different management and demographic histories.

References

AERTS, J., MEGENS, H. J., VEENENDAAL, T., OVCHARENKO, I., CROOIJMANS, R. P. M. A., GORDON, L., ... & GROENEN, M. Extent of linkage disequilibrium in chicken. Cytogenetic and Genome Research, 2007, 117(1-4), 338-345.

AULCHENKO, Y. S., RIPKE, S., ISAACS, A. & VAN DUIJN, C. M. GenABEL: an R package for genome-wide association analysis. Bioinformatics, 2007, 23(10), 1294-1296.

BARBATO, M., OROZCO-TERWENGEL, P., TAPIO, M., & BRUFORD, M. W. SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. Frontiers in Genetics, 2015, 6, 109.

BLENCH, R. M. & MACDONALD, K. C. Chickens. In The Cambridge world history of food, Volume I, FK Kenneth & C.O. Kriemhild editions. Cambridge, Cambridge University Press 2000, 496-499.

BURT, D. W. Chicken genome: Current status and future opportunities. Genome Research, 2005, 15, 692-1698.

CUC, N. T. K., SIMIANER, H., EDING, H., TIEU, H. V., CUONG, V. C., WOLLNY, C. B. A., GROENEVELD, L. F. & WEIGEND, S. Assessing genetic diversity of Vietnamese local chicken breeds using microsatellites. Animal Genetics, 2010, 41(5), 545-547.

DANA, N. Breeding programs for indigenous chicken in Ethiopia: analysis of diversity in production systems and chicken populations. PhD thesis, Wageningen University, 2011.

DANA, N., DESSIE, T., VAN DER WAAIJ, L. H. &VAN ARENDONK J. A. M. Morphological features of indigenous chicken populations of Ethiopia. Animal Genetic Resources, 2010, 46, 1-23.

DESTA, T. T., DESSIE, T., BETTRIDGE, J., LYNCH, S. E., MELESE, K., COLLINS, M., CHRISTLEY, R. M., WIGLEY, P., KAISER, P., TERFA, Z., MWACHARO, J. M. & HANOTTE, O. Signature of artificial selection and ecological landscape on morphological structures of Ethiopian village chickens. Animal Genetic Resources, 2013, 52, 1-13.

DING, K., & KULLO, I. J. Geographic differences in allele frequencies of susceptibility SNPs for cardiovascular disease. BMC Medical Genetics, 2011, 12(1), 55.

DRAY, S. & DUFOUR, A. B. The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software, 2007, 22(4), 1-20.

ELLEGREN, H. The avian genome uncovered. Trends in Ecology and Evolution, 2005, 20(4), 180-186.

FAO. World Watch List for Domestic Animal Diversity. Edited by BD Scherf, 3rd edition, FAO, Rome, Italy, 2000.

FULLER, D. Q., BOIVIN, N., HOOGERVORST, T. & ALLABY, R. Across the Indian Ocean: the prehistoric movement of plants and animals. Antiquity, 2011, 85(328), 544-58.

GRANEVITZE, Z., HILLEL, J., CHEN, G. H., CUC, N. T. K., FELDMAN, M., EDING, H. & WEIGEND, S. Genetic diversity within chicken populations from different continents and management histories. Animal Genetics, 2007, 38, 576–583.

HAMILTON, M. B. Population Genetics. Oxford, UK: Wiley-Blackwell, 2009.

HAYES, B. J., VISSCHER, P. M., MCPARTLAN, H. C. & GODDARD, M. E. Novel multilocus measure of linkage disequilibrium to estimate past effective population size. Genome Research, 2003, 13(4), 635-643.

HEDRICK, P. W. Genetics of Populations. Jones & Bartlett Publishers, 2011.

HILL, W. G. & ROBERTSON, A. Linkage disequilibrium in finite populations. Theoretical and Applied Genetics, 1968, 38(6), 226-231.

HILLIER, L. W., MILLER, W., BIRNEY, E., WARREN, W., HARDISON, R. C., PONTING, C. P., ... & DODGSON, J. B. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature, 2004, 432(7018), 695-716.

JENSEN, J. L., BOHONAK, A. J. & KELLEY, S. T. Isolation by distance, web service. BMC Genetics, 2005, 6, 13. v.3.23. http://ibdws.sdsu.edu/.

JOMBART, T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics, 2008, 24, 1403-1405.

KRANIS, A., GHEYAS, A. A., BOSCHIERO, C., TURNER, F., YU, L., SMITH, S., ... & BURT, D. W. Development of a high density 600K SNP genotyping array for chicken. BMC Genomics, 2013, 14(1), 59.

LEROY, G., KAYANG, B. B., YOUSSAO, I. A., YAPI-GNAORÉ, C. V., OSEI-AMPONSAH, R., N'GORAN, E. L., ... & ROGNON, X. Gene diversity, agroecological structure and introgression patterns among village chicken populations across North, West and Central Africa. BMC Genetics, 2012, 13(1), 34.

LYIMO, C. M., WEIGEND, A., MSOFFE, P. L., EDING, H., SIMIANER, H. & WEIGEND, S. Global diversity and genetic contributions of chicken populations from African, Asian and European regions. Animal Genetics, 2014, 45(6), 836–848.

MEGENS, H. J., CROOIJMANS, R. P., BASTIAANSEN, J. W., KERSTENS, H. H., COSTER, A., JALVING, R., ... & GROENEN, M. A. Comparison of linkage disequilibrium and haplotype diversity on macro-and microchromosomes in chicken. BMC Genetics, 2009, 10(1), 86.

MELESSE, A. & NEGESSE, T. Phenotypic and morphological characterization of indigenous chicken populations in southern region of Ethiopia. Animal Genetic Resources, 2011, 49, 19-31.

MTILENI, B. J., MUCHADEYI, F. C., MAIWASHE, A., GROENEVELD, E., GROENEVELD, L. F., DZAMA, K., & WEIGEND, S. Genetic diversity and conservation of South African indigenous chicken populations. Journal of Animal Breeding and Genetics, 2011, 128(3), 209-218.

MUCHADEYI, F. C., EDING, H., SIMIANER, H., WOLLNY, C. B. A., GROENEVELD, E., & WEIGEND, S. Mitochondrial DNA D-loop sequences suggest a Southeast Asian and Indian origin of Zimbabwean village chickens. Animal Genetics, 2008, 39(6), 615-622.

MUCHADEYI, F. C., EDING, H., WOLLNY, C. B. A., GROENEVELD, E., MAKUZA, S. M., SHAMSELDIN, R., ... & WEIGEND, S. Absence of population substructuring in Zimbabwe chicken ecotypes inferred using microsatellite analysis. Animal Genetics, 2007, 38(4), 332-339.

MUIR, W. M., WONG, G. K. S., ZHANG, Y., WANG, J., GROENEN, M. A., CROOIJMANS, R. P., ... & CHENG, H. H. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. Proceedings of the National Academy of Sciences, 2008, 105(45), 17312-17317.

MWACHARO, J. M., BJØRNSTAD, G., HAN, J. L. & HANOTTE, O. The history of African village chickens: an archaeological and molecular perspective. African Archaeological Review, 2013a, 1-18.

MWACHARO, J. M., BJØRNSTAD, G., MOBEGI, V., NOMURA, K., HANADA, H., AMANO, T., ... & HANOTTE, O. Mitochondrial DNA reveals multiple introductions of domestic chicken in East Africa. Molecular Phylogenetics and Evolution, 2011, 58(2), 374-382.

MWACHARO, J. M., NOMURA, K., HANADA, H., HAN, J. L., AMANO, T. & HANOTTE, O. Reconstructing the origin and dispersal patterns of village chickens across East Africa: insights from autosomal markers. Molecular Ecology, 2013b, 22(10), 2683-2697.

OSEI-AMPONSAH, R., KAYANG, B. B., NAAZIE, A., OSEI, Y. D., YOUSSAO, I. A., YAPI-GNAORE, V. C., ... & ROGNON, X. Genetic diversity of Forest and Savannah chicken populations of Ghana as estimated by microsatellite markers. Animal Science Journal, 2010, 81(3), 297-303.

PARADIS, E. pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics, 2010, 26(3), 419-420.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013, ISBN 3-900051-07-0. http://www.R-project.org/.

SCHEET, P. & STEPHENS, M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. The American Journal of Human Genetics, 2006, 78(4), 629-644.

SINCLAIR, E. A., & HOBBS, R. J. Sample size effects on estimates of population genetic structure: implications for ecological restoration. Restoration Ecology, 2009, 17(6), 837-844.

SMITH, L. M. & BURGOYNE, L. A. Collecting, archiving and processing DNA from wildlife samples using FTA® databasing paper. BMC Ecology, 2004, 4(1), 4.

SPENCER, C. C., DELOUKAS, P., HUNT, S., MULLIKIN, J., MYERS, S., SILVERMAN, B., ... & MCVEAN, G. The influence of recombination on human genetic diversity. PLoS Genetics, 2006, 2(9), e148.

SVED, J. A. Linkage disequilibrium of chromosome segments. Theoretical Population Biology, 1971, 141(2), 125-141.

SZPIECH, Z. A., JAKOBSSON, M. & ROSENBERG, N. A. ADZE: a rarefaction approach for counting alleles private to combinations of populations. Bioinformatics, 2008, 24(21), 2498-2504.

TAMURA, K., PETERSON, D., PETERSON, N., STECHER, G., NEI, M. & KUMAR, S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution, 2011, 28, 2731-2739.

VAVREK, M. J. fossil: Palaeoecological and palaeogeographical analysis tools. Palaeontologia Electronica, 2011, 14(1), 16.

WAHLUND, S. *Zusammensetzung von population und korrelationserscheinung vom standpunkt der vererbungslehre aus betrachtet*. Hereditas, 1928, 11, 65-106.

WEIR, B. S. & HILL, W. G. Effect of mating structure on variation in linkage disequilibrium. Genetics, 1980, 95(2), 477-488.

WELCH, B. L. The generalization of Student's problem when several different population variances are involved. Biometrika, 1947, 34(1/2), 28-35.

WILLING, E. M., DREYER, C., & VAN OOSTERHOUT, C. Estimates of genetic differentiation measured by FST do not necessarily require large sample sizes when using many SNP markers. PLoS One, 2012, 7(8), e42649.

WILSON, D. J., GABRIEL, E., LEATHERBARROW, A. J., CHEESBROUGH, J., GEE, S., BOLTON, E., FOX, A., FEARNHEAD, P., HART, C. A. & DIGGLE, P. J. Tracing the source of campylobacteriosis. PLoS Genetics, 2008, 4(9), e1000203.

WRAGG, D., MWACHARO, J. M., ALCALDE, J. A., HOCKING, P. M. & HANOTTE, O. Analysis of genome-wide structure, diversity and fine mapping of Mendelian traits in traditional and village chickens. Heredity, 2012, 109, 6-18.

WRIGHT, S. Evolution and the genetics of population, variability within and among natural populations. The University of Chicago Press, Chicago, 1978.

WRIGHT, S. The genetical structure of populations. Annals of Eugenics, 1951, 15(1), 323-354.

YA-BO, Y., JIN-YU, W., MEKKI, D. M., QING-PING, T., HUI-FANG, L., RONG, G., QING-LIAN, G., WEN-QI, Z. & KUAN-WEI, C. Evaluation of genetic diversity and genetic distance between twelve Chinese indigenous chicken breeds based on microsatellite markers. International Journal of Poultry Science, 2006, 5(6), 550-556.

Chapter 5

A genome-wide high density single nucleotide polymorphism analysis uncovers the unique genetic structure of Ethiopian village chickens

Abstract

Domestic chickens are developed under contrasting breeding histories following the formation of commercial breeds in the mid-20th century. The commercial chickens have been intensively selected as production breeds, while the nondescript village chickens have been kept to fulfil the multi-dimensional needs of subsistence farmers. We investigate the impact of breeding and demographic history on genetic structure of domestic chickens and the junglefowl species. We analyse three main gene pools: village chickens, commercial chickens (white and brown egg layers and the dual purpose Rohde Island Red) and the junglefowl species (*Gallus gallus, G. lafayetti, G. sonneratti and G. varius*) using a high density (600K) SNP array. The first principal component (10.01%) separates commercial chickens from the remaining genepool while the second principal component that accounts for 6.59% of the total genetic variation separates white egg layers and Ethiopian village chickens from brown egg layers, the junglefowls and the remaining village chickens. A phylogenetic analysis shows a higher genetic divergence between the Ethiopian Horro and Jarso chickens than between Kenyan and Nigerian chickens. Our results show that genetic structure of Ethiopian chickens is considerably varied from village chickens sampled outside of Ethiopia but have been evolved under a comparable management system. The genetic disconnectedness of Ethiopian village chickens may be associated with their demographic and breeding history. Nevertheless a further extensive study that involves representative samples from non-Ethiopian village chickens and the junglefowl species is required to corroborate our findings.

Keywords: junglefowl species, village chickens, commercial chickens, phylogenetic relationship, breeding history, demographic history

Introduction

The chicken is one of the most widely studied livestock species with respect to its genetic structure and demographic history. However, no consensus has yet been reached as to whether the chicken is of monophyletic origin (Darwin, 1868; Fumihito et al., 1994 & 1996) or if it is polyphyletic (Eriksson et al., 2008) origin and concerning its routes of dispersion from the putative centre (s) of domestication (Blench & Macdonald, 2000; Storey et al., 2012). The polyphyletic origin school of thought is supported by a study reporting introgression of the grey junglefowl yellow skin gene to domestic fowl (Eriksson et al., 2008). However, this was questioned by Flink et al. (2014), who substantiate the post domestication occurrence of this mutation. The extended black plumage gene also thought to be introgressed from the green junglefowl (Stevens, 1991), however this species is believed to be genetically distant even from other junglefowl species (Crawford, 1990). The monophyletic origin of domestic fowl from red junglefowl is supported by experimental crossing made between the junglefowls and the domestic chickens and due to morphological similarity between the red junglefowl and the domestic chickens (Darwin, 1868). Analysis of the mtDNA D-loop sequence of the domestic fowl and the junglefowl species (Fumihito et al., 1994 & 1996) also suggest a monophyletic origin.

A large scale comprehensive study that involves all representative species and subspecies of the junglefowl, the nondescript village chickens and commercial chickens and covering a wide range of geographical region is required to sufficiently address this controversy. The contribution of the red junglefowl to the domestic fowl genepool may be proportionally higher than the other junglefowl species (*e.g.* Hillel et al., 2003), which may have also contributed to this controversy. Majority of the studies performed to date have been concentrated on analysis of the hypervariable control region of the mtDNA (*a.k.a.* D-Loop). However, this short region may in itself be insufficient to unlock the complex demographic history of domestic chickens and their wild relatives (Miao et al., 2012) and it only represents the maternal lineage.

The fact that the ancestral species of modern chicken continues to survive in Asia makes the chicken of special importance in evolutionary genetics studies.

Presumably, following their domestication, chickens could have been dispersed more quickly than other livestock species owing to their portable nature and a high rate of reproduction. Domestic chickens are widely distributed across agriculturally important parts of the world. Depending on the demand of breeders and consumers, the domestic chickens have been subjected to different breeding histories ranging from the uncontrolled (that mimics panmixia) to the vertically integrated commercial breeding scheme. Uncontrolled breeding and the impact of hypervariable production environments have created a vast array of phenotypic diversity in the extant village chickens (FAO, 2000).

It might be that due to their different breeding histories, village chickens have been genetically diverged from the commercial populations (Mwacharo et al., 2007). Village chickens have been largely screened by natural selection for local adaptation, while the commercial chickens have been intensively selected for production traits. The junglefowl species on the other side, being natural populations have been experiencing little impact of direct human interference in their breeding history. The junglefowl species therefore could serve as a reference population to separate the impact of artificial selection from evolutionary events associated with domestication and natural selection. Gene flow in the junglefowl species could be constrained by physical (large water bodies, habitat fragmentation for example due to deforestation) and biological barriers (poor flight and swimming capacity and a limited ability of walking for long distance, *i.e.* narrow home range). Physical and biological barriers reduce effective population size or in the extreme cases, they may even have led to speciation of the junglefowls species. Predation and hunting (Collias and Saichuae, 1967) may also have reduced population of the junglefowl. Gene flow among commercial populations is strictly regulated and when this combines with intensive artificial selection it gradually reduces the standing genetic variation. Gene flow in village chickens is considerably high attributable to uncontrolled breeding. Consquently, the genetic diversity in village chickens is high (Hillel et al., 2003; Granevitze et al., 2007; Muchadeyi et al., 2007; Silva et al., 2009), which could reflect the impact of a mild artificial selection and broad breeding objectives of subsistence farmers that favour keeping of chickens of diverse phenotypes.

Despite a number of studies conducted involving large genepools of chickens (e.g. Hillel et al., 2003; Granevitze et al., 2007), a comparative study that includes village, commercial and the junglefowl to the best of our knowledge is infrequent. We analyse population structure and genetic relationship of village chicken, commercial layers and the junglefowl species using a high density SNP array. Our analysis uncovers the unique demographic history of the three genepools and all analyses invariably identify Ethiopian village chickens as genetic outliers.

Materials and methods

The study populations

This study involves three main genepools. Village chickens sampled from Africa (Ethiopian Horro and Jarso chickens (n = 756), Kenyan chicken (n = 4), Nigerian chicken (n = 5)), Asia (Sri Lankan chicken, n = 5) and Latin America (Chilean chicken, n = 15). The junglefowl species (subspecies) include *G. g. gallus* (n = 4), *G. g. spadiceus* (n = 5), *G. g. bankiva* (n = 1), *G. lafayetii* (n = 2), *G. sonneratii* (n = 1) and *G. varius* (n = 1). The commercial layers genepool is represented by 86 chickens: 20 Brown Egg White Rock (BEWR), 30 White Leghorn (WLH), 24 Brown Egg Female Line (BEFL) and 12 White Egg Male Line (WEML) and the dual purpose Rhode Island Red (RIR, n = 20). The SNP genotype data for commercial layers was obtained from David Burt lab, The Roslin Institute, while for village chickens sampled outside of Ethiopia and the junglefowl species it belonged to David Wragg. The RIR represents an outbred population. Henceforth, village chickens sampled outside of Ethiopia are referred as “other village chickens”.

DNA library preparation and data analysis

A whole blood was collected from the wing vein and was spotted on Whatman® FTA® cards for a subsequent DNA extraction following Smith & Burgoyne (2004) method 4. The DNA samples were suspended in Tris-EDTA buffer (pH 8.0) at a concentration of 50 ng μl^{-1} and a volume of 40 µl. The DNA was genotyped by Ark Genomics (Roslin Institute, Edinburgh) using the Axiom® 600K Affymetrix SNP array (Kranis et al., 2013).

Two quality control (QC) steps were performed using GenABEL package (Aulchenko et al., 2007) of the R (R Core Team, 2013). The criteria of the first QC are SNP and individual calling rate ≥ 90%; minor allele frequency (MAF) ≥ 5%; all markers were considered in identity by state (IBS) calculation, and the sample with a lower calling rate in the pair-wise comparisons was excluded if the IBS exceeds 0.95. From the raw data containing 528019 SNPs and 924 samples, 482460 autosomal (GGA1–28) SNPs and 905 samples were retained after the first QC. Nineteen chickens were excluded due to high IBS, 44435 and 1134 SNPs were excluded due to low MAF and low calling rate respectively. A second QC was performed at population level using the Hardy-Weinberg Equilibrium (HWE) test as implemented in GenABEL ($P \geq 10^{-3}$). The data from each population was then merged by common SNPs that passed the HWE test in all populations, leaving 135841 SNPs and 905 chickens for a downstream analysis.

In a subsequent analysis, except a single principal component analysis (PCA), a random subset of either 25 or 5 samples from each region (Horro and Jarso) were included in the analysis to reduce the bias that may arise due to variation in sample size. To infer global and fine-scale population substructure, principal component analysis (PCA) was performed for different sets of genepools using ade4 package (Dray and Dufour, 2007) of the R. The BIC statistics as implemented in the R package adegenet (Jombart, 2008) was used to identify the optimal number of clusters. For phylogenetic analysis we use a genetic distance matrix generated from IBS matrix using the IBS function of GenABEL. The genetic distance matrix was then imported to MEGA5 (Tamura et al., 2011) to construct a phylogenetic tree using the neighbour-joining (NJ) option. The constructed NJ tree was visualized using FigTree v1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/).

Results

Principal component analysis of domestic and junglefowls

Analysis of twenty five chickens from each Horro and Jarso along with the remaining populations clearly separates the commercial layers from the remaining genepool by the first principal component (PC) explaining 10.01% of the total

genetic variation (Figure 5.1). Ethiopian chickens were separated from other village chickens that have been managed and developed under a comparable management system by the second PC. Based on this analysis, other village chickens were genetically closely related to the junglefowls than to Ethiopian chickens. PC2 explained 6.59% of the total genetic variation and it separates white egg layers and Ethiopian chickens from other village chickens, brown egg layers and the junglefowls. Both PC1 and PC2 to a considerable extent reveal the genetic closeness of village chickens to the junglefowl species than commercial chickens. Different lines of the commercial chickens made a tight cluster at intrapopulation level but there is a considerable divergence among lines, which reflects their divergent breeding and selection history. From BIC statistic of adegenet the optimum number of genetic clusters was found to be ten (see Figure S5.2 in the Appendices), which is comaparable with the PCA result (Figure 5.1).

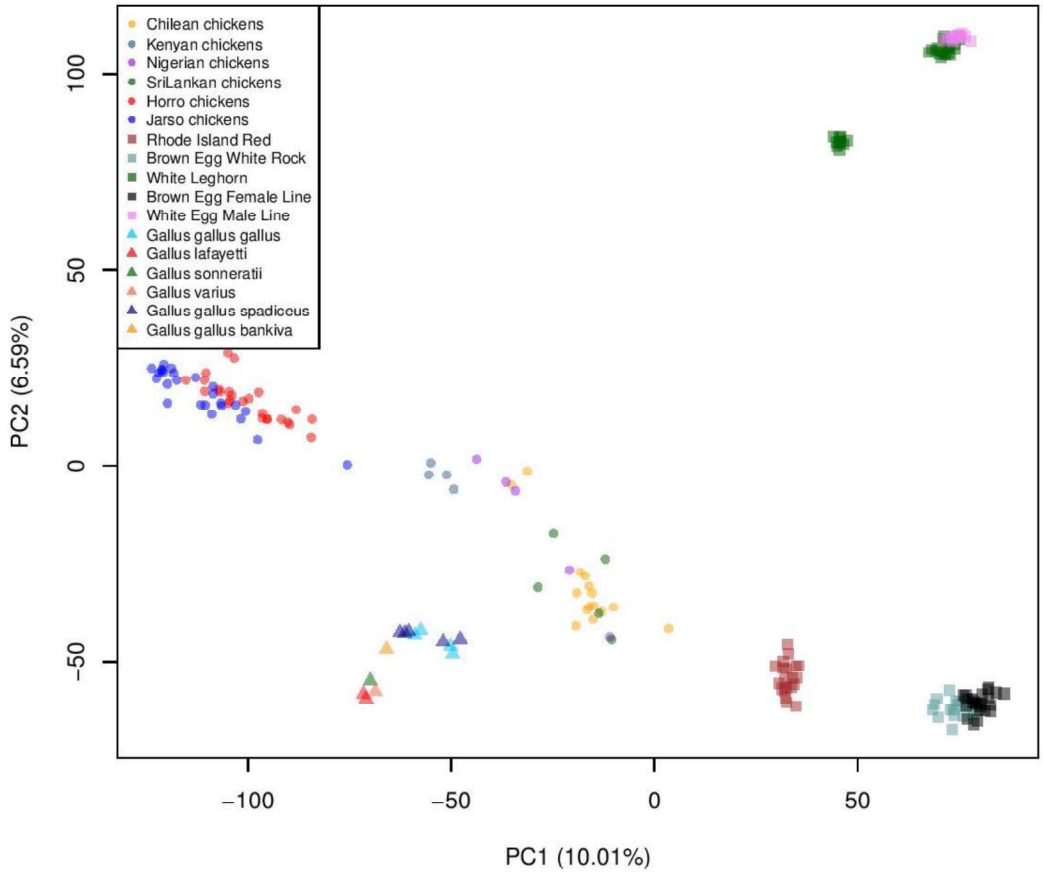


Figure 5.1 Principal component analysis of domestic chickens and the junglefowl species.

Principal component analysis of all Ethiopian chickens and the remaining populations

To investigate the impact of sample size variation on a genetic structure, the entire dataset of Horro (n = 383) and Jarso chickens (n = 373) were analysed with the remaining populations. Unlike the previous PCA, this analysis identifies two independent clusters peculiar to the Ethiopian Horro and Jarso chickens (Figure 5.2). Jarso chickens are more scattered along the PCA axis while Horro chickens are tended to form a tight cluster, which may indicate the genetic heterogeneity of the Jarso chickens. More interestingly, PC1 separates Jarso chickens from the remaining populations, which indicates the unique genetic structure of the Jarso chickens. This analysis made four main clusters and has grouped other village chickens with the junglefowl species.

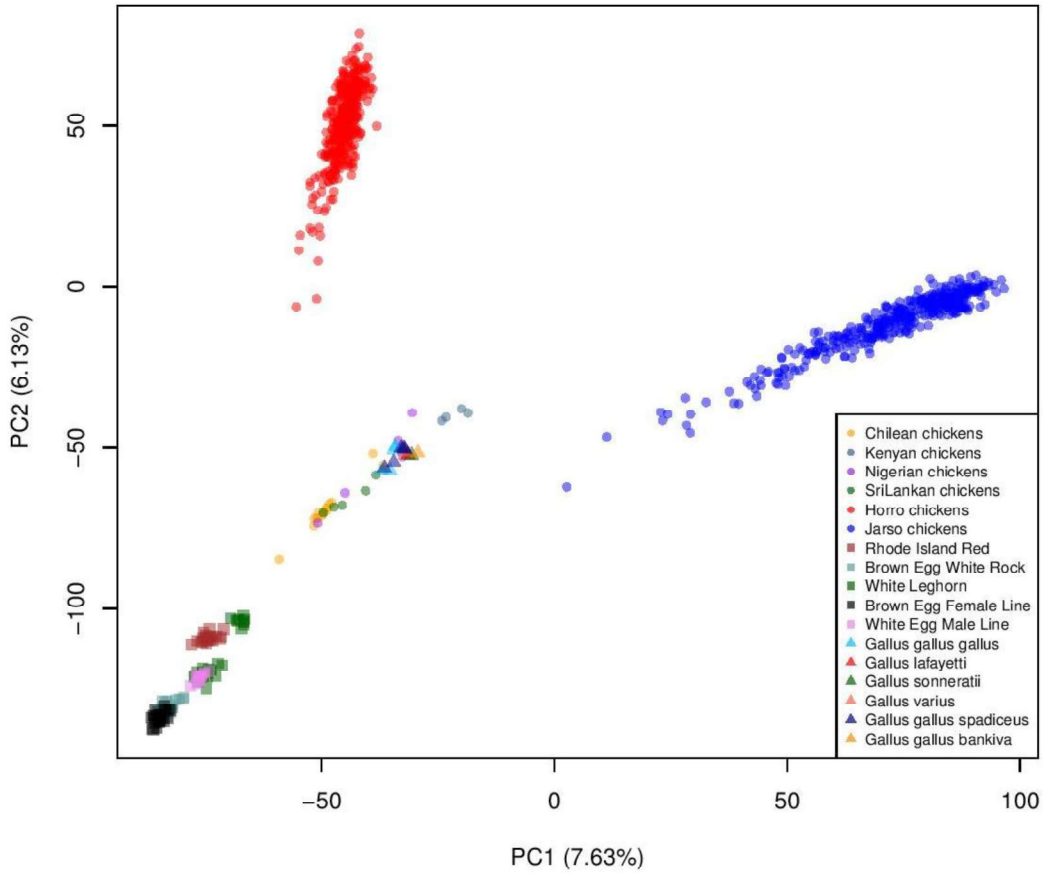


Figure 5.2 Principal component analysis of all Ethiopian chickens and the remaining genepools.

Phylogenetic analysis

The neighbour-joining tree rooted on *G. varius* (Figure 5.4) supports the results of PCA (Figure 5.1). African village chickens share a common node as the commercial lines did, while the Chilean and Sri Lankan chickens were clustered close to the junglefowl. The phylogenetic tree indicates that Horro and Jarso chickens are distantly related to each other than the Kenyan and Nigerian chickens are which suggests a common origin for the latter two. Unexpectedly, a chicken from Nigeria was clustered along with the Chilean chickens, which may reflect genetic introgression of Chilean genepool to some Nigerian chickens. In line with PCA (Figure 5.1 and 5.2), a clearly defined genetic divergence was observed between white and brown egg layers attributable to their divergent selection history (Figure 5.3). The Chilean chickens are closely clustered with the junglefowl species despite their distant geographical location. A recent maritime trade connection between the pacific region of the south East Asia and the new world may partly contribute to this genetic relatedness. In line with our findings, Gongora et al. (2008) also reported Indo-European and Asian origin of the Chilean chickens and their genetic relationship with chickens of the pacific region. Chickens of the pacific region may closely relate with the junglefowl species, as the home range of the junglefowls may extend to the pacific islands (Peterson and Brisbin, 1999). The two major genetic clusters found from the phylogenetic analysis are consistent with the genetic cluster made by PC2 axis in the PCA (Figure 5.1).

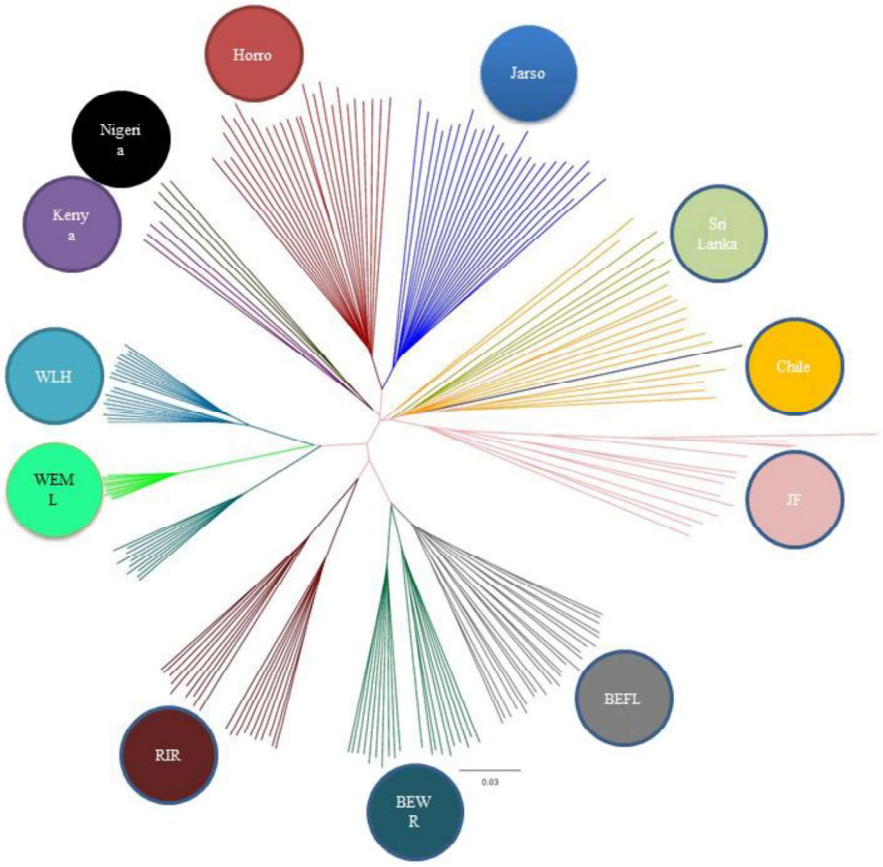


Figure 5.3 Phylogenetic relationship among domestic chickens and the junglefowl species.

Discussion

We analysed the genetic structure of heterogeneous domestic chicken and junglefowl populations consisting of a mildly selected village chickens, intensively selected commercial chickens and the non-selected junglefowls using a genome-wide HD (600K) SNPs array. Our results from PCA and phylogenetic relationship analyses reveal the impact of demographic and breeding history on the genetic structure of domestic chickens and the junglefowls. Differentially acting evolutionary forces could have resulted in allelic frequency variations among populations and this may have shaped the genetic structures of domestic and wild populations. Variation in their production environments may have also shaped the genetic structure of domestic chickens and the junglefowls.

Dispersion of chickens from their putative centre(s) of domestication during a prehistoric period may have been performed following migration and trading-network of an ancient man (Blench & Macdonald, 2000; Mwacharo et al., 2011; Storey et al., 2012). The chickens may have arrived in different continents and/or countries at different times and/or via different routes (Blench & Macdonald, 2000). Attempts have been made to reconstruct the dispersion pattern of domestic chickens and their relationship with the junglefowl species typically from matrilineal lineage perspective using the mtDNA D-Loop sequence (Fumihito et al., 1994 & 1996; Nishibori et al., 2005; Liu et al., 2006; Oka et al., 2007; Kanginakudru et al., 2008; Silva et al., 2009; Mwacharo et al., 2011; Miao et al., 2012). However, controversies have been introduced by different studies (Blench & Macdonald, 2000) and this has created uncertainties on centre of domestication and dispersion pattern of the domestic fowl. Moreover, extensive genetic dilution of village chickens has occurred following a massive distribution of commercial stocks across the the least developed world. This may reduce the resolution power of genetic markers used to reconstruct the phylogeographic structure of the domestic chickens (Storey et al., 2012).

Intricate web of flock movement may have led to extensive genetic admixture in the domestic chickens. Particularly, flocks that have been evolved under uncontrolled breeding could have been experiencing extensive gene flow, although gene flow in some instances might be constrained by physical, biological and anthropogenic barriers. A long history of chicken rearing and their prolific reproductive characteristics, portability and a high recombination rate in the chicken autosomes could have also resulted in a high genetic diversity. In spite of a high genetic diversity of the chicken, populations that share a particular geographic region may considerably share a common genetic background and this has enabled the evolutionary geneticists to trace to the origin of populations.

Besides the relationship established between *G. g. gallus* and the domestic fowl (*e.g.* Fumihito et al., 1994 & 1996); genetic relationship analyses performed using the mitochondrial D-Loop sequence show a similarity between *G. g. spadiceus* and native chickens of Japan (Oka et al., 2007), India (Kanginakudru et al., 2008), China (Liu et al., 2004) and diverse genepools of domestic chickens from Asia and Europe (Fumihito et al., 1996). Moreover, Pramual et al. (2013) reported *G.g. murghi* as the

contributor of Thai chickens genepool. Although Sri Lanka is the native place for Ceylon junglefowl, our analyses show that Sri Lankan chickens are not genetically closer to Ceylon junglefowl as other village chickens are, which may show lack of genetic contribution from Ceylon junglefowl to the domestic chickens genepool. In line with our findings, Sri Lankan village chickens are genetically closer to RIR than the Ceylon junglefowl (see Silva et al., 2009 for a review). Although chickens may have a polyphyletic origin, as our preliminary analyses indicate they may be genetically closer to *G. g. gallus* and *G.g. spadiceus* than other species and subspecies of the junglefowl.

The two separate clades produced by Ethiopian Horro and Jarso chickens on the PCA axis (Figure 5.2) and phylogenetic tree (Figure 5.3) may associate with two entry points of domestic chicken to East Africa – through the Northeast Africa (Egypt) and the Red Sea coast (Blench & Macdonald, 2000). Horro is located near to the course of the Blue Nile River in the western Ethiopia while Jarso is located relatively closer to the Red Sea coast. The two populations may have also originated from different ancestral genepools. Despite the nondescript morphological structure of village chickens, considerable differences were also observed between Horro and Jarso chickens for some morphological traits (Desta et al., 2013).

A variation in anthropogenic effects and production environments may have also shaped the genetic structure of the chicken populations. Following their domestication, chickens have been evolved at least under two mainstream management systems — the subsistence-oriented smallholder farmers and a highly specialized commercial and show breeders management systems. The management system adopted by the flock owners determines the degree at which domestic chickens interact with local environment, the intensity of selection and the extent of gene flow. In commercial and fancy chickens, selection for traits of speciality traits is intense and this ascends the frequency of the allelic variants preferred, which then reduces genetic variation. Moreover, gene flow among commercial chickens is systematically controlled and as a result there is a little (if any) genetic admixture. On the other side, due to extensive gene flow and genetic admixture, village chickens display a high genetic diversity. Although each individual's genome harbours mosaics of chromosomal segments inherited from different populations,

genetically admixed populations like the village chickens are well known for their genetic homogeneity (Muchadeyi et al., 2007).

Ethiopian chickens are genetically disconnected from other village chickens in all the analyses performed. This genetic disconnectedness may reflect Ethiopian chickens' unique demographic history and they may have been developed under considerably contrasting anthropogenic events, production environments and breeding histories. Ethiopia is characterized by a hypervariable agro-ecological landscape (MOA, 1998), and it is the cradle of an ancient agriculture (McCann, 1995). It is however not clear whether ecological diversity and/or the high ethnographic diversity of Ethiopia and/or an ancient practice of agriculture have been dictating the evolution of this unique genetic structure. Despite limited archaeological studies, there are evidences for an ancient use of chicken as a source of food in Ethiopia (D'Andrea et al., 2011). An ancient practice of livestock rearing has been also corroborated from prehistoric rock paintings made on cave and rock shelters with majority of them depicting a range of livestock species and a number of livestock remain assemblages were also discovered in the archaeological sites (Clark & Williams, 1978; Lofrumento et al., 2012; Assefa et al., 2014; Gomes et al., 2013). The archaeological sites that have been discovered in eastern Ethiopia are located close to Jarso. However, there is no reported evidence for independent domestication of chickens in Africa or for genetic admixture between extinct native African junglefowl (if any) and the domestic chickens introduced from Asia. The genetic disconnectedness of Ethiopian village chickens may also partly associate with the sampling strategies we adopt. We sampled indigenous chickens with no known recent history of genetic introgression from commercial chickens, though; Ethiopians chickens have been noted for their remote genetic relatedness to commercial populations (*e.g.* Hassen et al., 2009).

Conclusion

Although the number of samples from other village chickens and the junglefowl species included in this analysis are insufficient to draw a strong conclusion, we found suggestive results that elicit research questions for future studies. The genetic uniqueness of Ethiopian chickens demands further studies that involve well-

represented and large number of genepools from both domestic chickens and the junglefowls. The unique genetic background of Ethiopian chickens particularly evident in Jarso chickens and the discovery of a number of archaeological sites close to Jarso may indicate a prehistoric livestock rearing practice in this area. A comprehensive study that combines ancient DNA analysis with historical, linguistic, archaeological and anthropological evidences should have to be conducted to uncover the unique genetic structure of Ethiopian chickens.

References

ASSEFA, Z., PLEURDEAU, D., DUQUESNOY, F., HOVERS, E., PEARSON, O., ASRAT, A., & LAM, Y. M. Survey and explorations of caves in south eastern Ethiopia: Middle Stone Age and Later Stone Age archaeology and Holocene rock art. Quaternary International, 2014, 343, 136-147.

AULCHENKO, Y. S., RIPKE, S., ISAACS, A. & VAN DUIJN, and C. M. Enable: an R package for genome-wide association analysis. Bioinformatics, 2007, 23(10), 1294-1296.

BLENCH, R. M. & MACDONALD, K. C. Chickens. In F.K. Kenneth & C.O. Kriemhild, Eds, The Cambridge world history of food, Volume I, 496–499. Cambridge, Cambridge University Press, 2000.

CLARK, J. D. & WILLIAMS, M. A. J. Recent archaeological research in Southeastern Ethiopia. 1974 - 1975. Annales d'Ethiopie, 1978, 19-44.

COLLIAS, N. E. & SAICHUAE, P. Ecology of the red junglefowl in Thailand and Malaya with reference to the origin of domestication. Nat. Hist. Bull Siam Soc., 1967, 22, 189-209.

CRAWFORD, R. D. Poultry Breeding and Genetics. Elsevier B.V, 1990.

D'ANDREA, A. C., RICHARDS, M. P., PAVLISH, L. A., WOOD, S., MANZO, A. & WOLDE-KIROS, H. S. Stable isotopic analysis of human and animal diets from two pre-Aksumite/Proto-Aksumite archaeological sites in northern Ethiopia. Journal of Archaeological Science, 2011, 38(2), 367-374.

DARWIN, C. R. The variation of animals and plants under domestication. London: John Murray, 1868.

DESTA, T. T., DESSIE, T., BETTRIDGE, J., LYNCH, S. E., MELESE, K., COLLINS, M., CHRISTLEY, R. M., WIGLEY, P., KAISER, P., TERFA, Z., MWACHARO, J. M. & HANOTTE, O. Signature of artificial selection and ecological landscape on morphological structures of Ethiopian village chickens. Animal Genetic Resources, 2013, 52, 1-13.

DRAY, S. & DUFOUR, A. B. The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software, 2007, 22(4), 1-20.

ERIKSSON, J., LARSON, G., GUNNARSSON, U., BED'HOM, B., TIXIER-BOICHARD, M., STRÖMSTEDT, L., ... & ANDERSSON, L. Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. PLoS Genetics, 2008, 4(2), e1000010.

FAO. World Watch List for Domestic Animal Diversity. 3rd edn. (BD Scherf Ed.), FAO, Rome, Italy, 2000.

FLINK, L. G., ALLEN, R., BARNETT, R., MALMSTRÖM, H., PETERS, J., ERIKSSON, J., ... & LARSON, G. Establishing the validity of domestication genes using DNA from ancient chickens. Proceedings of the National Academy of Sciences, 2014, 111(17), 6184-6189.

FUMIHITO, A., MIYAKE, T., SUMI, S., TAKADA, M., OHNO, S. & KONDO, N. One subspecies of the red junglefowl (*Gallus gallus gallus*) suffices as the matriarchic ancestor of all domestic breeds. Proceedings of the National Academy of Sciences, 1994, 91(26), 12505-12509.

FUMIHITO, A., MIYAKE, T., TAKADA, M., SHINGU, R., ENDO, T., GOJOBORI, T., KONDO, N. & OHNO, S. Monophyletic origin and unique dispersal patterns of domestic fowls. Proceedings of the National Academy of Sciences, 1996, 93(13), 6792-6795.

GOMES, H., ROSINA, P., HOLAKOOEI, P., SOLOMON, T. & VACCARO, C. Identification of pigments used in rock art paintings in Gode Roriso-Ethiopia using Micro-Raman spectroscopy. Journal of Archaeological Science, 2013, 40(11), 4073-4082.

GONGORA, J., RAWLENCE, N. J., MOBEGI, V. A., JIANLIN, H., ALCALDE, J. A., MATUS, J. T., ... & COOPER, A. Indo-European and Asian origins for Chilean and Pacific chickens revealed by mtDNA. Proceedings of the National Academy of Sciences, 2008, 105(30), 10308-10313.

GRANEVITZE, Z., HILLEL, J., CHEN, G. H., CUC, N. T. K., FELDMAN, M., EDING, H., & WEIGEND, S. Genetic diversity within chicken populations from different continents and management histories. Animal Genetics, 2007, 38(6), 576-583.

HASSEN, H., NESER, F. W. C., DE KOCK, A. & VAN MARLE-KÖSTER, E. Study on the genetic diversity of native chickens in northwest Ethiopia using microsatellite markers. African Journal of Biotechnology, 2009, 8(7), 1347-1353.

HILLEL, J., GROENEN, M. A., TIXIER-BOICHARD, M., KOROL, A. B., DAVID, L., KIRZHNER, V. M., ... & WEIGEND, S. Biodiversity of 52 chicken populations assessed by microsatellite typing of DNA pools. Genetics Selection Evolution, 2003, 35(5), 533-558.

JOMBART, T. adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics, 2008, 24, 1403-1405.

KANGINAKUDRU, S., METTA, M., JAKATI, R. D., & NAGARAJU, J. Genetic evidence from Indian red jungle fowl corroborates multiple domestication of modern day chicken. BMC Evolutionary Biology, 2008, 8(1), 174.

KRANIS, A., GHEYAS, A. A., BOSCHIERO, C., TURNER, F., YU, L., SMITH, S., ... & BURT, D. W. Development of a high density 600K SNP genotyping array for chicken. BMC Genomics, 2013, 14(1), 59.

LIU, Y. P., WU, G. S., YAO, Y. G., MIAO, Y. W., LUIKART, G., BAIG, M., ... & ZHANG, Y. P. Multiple maternal origins of chickens: out of the Asian jungles. Molecular Phylogenetics and Evolution, 2006, 38(1), 12-19.

LIU, Z. G., LEI, C. Z., LUO, J., DING, C., CHEN, G. H., CHANG, H., ... & WU, S. L. Genetic variability of mtDNA sequences in Chinese native chicken breeds. Asian Australasian Journal of Animal Sciences, 2004, 17(7), 903-909.

LOFRUMENTO, C., RICCI, M., BACHECHI, L., DE FEO, D., & CASTELLUCCI, E. M. The first spectroscopic analysis of Ethiopian prehistoric rock painting. Journal of Raman Spectroscopy, 2012, 43(6), 809-816.

McCann, J. C. People of the plow: An agricultural history of Ethiopia, 1800–1990, University of Wisconsin Press, 1995.

MIAO, Y. W., PENG, M. S., WU, G. S., OUYANG, Y. N., YANG, Z. Y., YU, N., ... & ZHANG, Y. P. Chicken domestication: an updated perspective based on mitochondrial genomes. Heredity, 2012, 110(3), 277-282.

MOA. Agroecological Zones of Ethiopia. Natural Resources Management and Regulatory Department, Ministry of Agriculture, March 1998, Addis Ababa, Ethiopia.

MUCHADEYI, F. C., EDING, H., WOLLNY, C. B. A., GROENEVELD, E., MAKUZA, S. M., SHAMSELDIN, R., ... & WEIGEND, S. Absence of population substructuring in Zimbabwe chicken ecotypes inferred using microsatellite analysis. Animal Genetics, 2007, 38(4), 332-339.

MWACHARO, J. M., BJØRNSTAD, G., MOBEGI, V., NOMURA, K., HANADA, H., AMANO, T., ... & HANOTTE, O. Mitochondrial DNA reveals multiple introductions of domestic chicken in East Africa. Molecular Phylogenetics and Evolution, 2011, 58(2), 374-382.

MWACHARO, J. M., NOMURA, K., HANADA, H., JIANLIN, H., HANOTTE, O., & AMANO, T. Genetic relationships among Kenyan and other East African indigenous chickens. Animal Genetics, 2007, 38(5), 485-490.

NISHIBORI, M., SHIMOGIRI, T., HAYASHI, T. & YASUE, H. Molecular evidence for hybridization of species in the genus Gallus except for *Gallus varius*. Animal Genetics, 2005, 36, 367-375.

OKA, T., INO, Y., NOMURA, K., KAWASHIMA, S., KUWAYAMA, T., HANADA, H., ... & AKISHINONOMIYA, F. Analysis of mtDNA sequences shows Japanese native chickens have multiple origins. Animal Genetics, 2007, 38(3), 287-293.

PETERSON, A. T., & BRISBIN, I. L. Genetic endangerment of wild Red Junglefowl *Gallus gallus*? Bird Conservation International, 1999, 8(04), 387-394.

PRAMUAL, P., MEEYEN, K. & WONGPAKAM, K. Genetic diversity of Thai native chicken inferred from mitochondrial DNA sequences. Tropical Natural History, 2013, 13(2), 97-106.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013, ISBN 3-900051-07-0. http://www.R-project.org/.

SILVA, P., GUAN, X., HO-SHING, O., JONES, J., XU, J., HUI, D., NOTTER D & SMITH, E. Mitochondrial DNA–based analysis of genetic variation and relatedness among Sri Lankan indigenous chickens and the Ceylon junglefowl (*Gallus lafayetti*). Animal Genetics, 2009, 40(1), 1-9.

SMITH, L. M. & BURGOYNE, L. A. Collecting, archiving and processing DNA from wildlife samples using FTA® databasing paper. BMC Ecology, 2004, 4(1), 4.

STEVENS, L. Genetics and Evolution of the Domestic Fowl. Cambridge University Press, 1991.

STOREY, A. A., ATHENS, J. S., BRYANT, D., CARSON, M., EMERY, K., HIGHAM, C., ... & MATISOO-SMITH, E. Investigating the global dispersal of chickens in prehistory using ancient mitochondrial DNA signatures. Plos One, 2012, 7(7), e39171.

TAMURA, K., PETERSON, D., PETERSON, N., STECHER, G., NEI, M. & KUMAR, S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution, 2011, 28, 2731-2739.

Chapter 6

Runs of homozygosity and uniparental disomy in Ethiopian village chickens

Abstract

Genetic structure of village chickens is mainly shaped by natural selection. Mating is commonly uncontrolled, flock size is small and the impact of anthropogenic effect is mild. A high genetic diversity of village chickens is attributable to their long history of uncontrolled breeding and extensive gene flow and genetic admixture. We map runs of homozygosity (ROH) using a high density (600K) SNP array in large number of two Ethiopian village chicken landraces: Horro (n = 383) and Jarso (n = 373). The two chicken populations have been evolved under considerably contrasting production environments. A considerable portion of the two chicken populations' genome harbours ROH. The frequency and extent of ROH however considerably vary among the autosomes and between the two chicken populations and hens and cocks. The extensive ROH mapped in some chickens implicates the impact of a recent consanguineous mating. Overlapping ROH regions that share at least one SNP among a minimum of fifty chickens were mapped to 40491 genomic regions across the twenty eight autosomes (GGA1–28). Islands of ROH were commonly found in genic regions regulating a number of vital biological processes. For the first time we map twenty three chromosomal uniparental disomy (UPD) in twenty two chickens (two of the UPDs belong to a Horro cock) and to fourteen autosomes (GGA5, 7, 11–16, 20–22 and 26–28). The ROH mapped was intense for some chickens; therefore a breeding program that avoids mating of closely related chickens would have to be initiated. Moreover, a breeding plan that is intended at increasing size of the family flock needs to be implemented to maintain the standing genetic variation.

Keywords: consanguineous mating, selection pressure, extended homozygosity, linkage disequilibrium, autozygosity

Introduction

Village chickens of the least developed world are kept for millenia in areas where livestock farming significantly contributes to the livelihood of subsistence farmers. In village chicken production system, despite small family flock size, mating is commonly random and uncontrolled. Village chickens have been mainly evolved under natural selection and they have been developed in a hypervariable production system. In the breeding management of village chickens intereference of human is less intense. Variability in genetic structure of village chickens is therefore largely attributed to natural selection and their demographic history. Presumably, village chickens can be placed in an intermediate position between the junglefowl and the commercial chickens with respect to their management and breeding history.

Although villge chickens represent outbred populations, their genome may harbour extended homozygosity in regions to which polymorpism has been known to exist from whole genome sequence analyses. Runs of homozygosity (ROH) refers to a contiguous segment of homozygous genotypes that are located on a pair of homologous chromosomes. ROH (tracts of contiguous homozygote genotypes) may arise from a consanguineous mating, extended linkage disequlibrium (particularly in recombination coldspot regions), natural selection, genetic drift and demographic history of the population in question (Gibson et al., 2006, Bosse et al., 2012; Purfield et al., 2012; MacLeod et al., 2013). In rare case, ROH may arise from isodisomy (Gibson et al., 2006). An intensive human-driven selection for speciality traits increases the extent of autozygosity in genomic regions underlying the genetic control of the trait subjected to selection. ROH burden may vary across the genome depending on the demgraphic events and recombination rates (McQuillan et al., 2008; Bosse et al., 2012; Pemberton et al., 2012; Purfield et al., 2012). Besides consanguineous mating, low recombination rate and/or strong linkage diseqilibrium (LD) can lead to formation of a ROH.

A consanguineous mating transmits identical by descent (IBD) haplotypes from mates (Kirin et al., 2010; Bosse et al., 2012), which can then form long tracts of homozygous regions in the offspring's genome. Extensive ROH could also arise from a recent reduction in effective population size and due to inbreeding and an on

going strong artificial selection (Kim et al., 2013; MacLeod et al., 2013). Non-consanguineous individuals could also show ROH due to shared common ancestor in the distant past and in genomic regions exhibiting low recombination rate and experiencing a high selection pressure. Conserved regions of a genome most likely display homozygous genotypes among individuals within a population and even among species, which could be considred as ROH in case the threshold size set while defining a ROH is small.

Innovation of an automated genotyping for high density SNPs and the quickly evolving sequencing technologies provide a huge resource to map ROH in domesticates genome. ROH is mapped at individual level and this is a plausible strategy to assess the level of inbreeding in village chickens that have been characterized by absence of documented pedigree records. ROH status can be used as a resource to select and bred genetically distant and less inbred individuals. To the best of our knowledge, most of the studies that have been intended to map ROH in the livestock species (Ferenčaković et al., 2011; Bosse et al., 2012; Purfield et al., 2012; Ferenčaković et al., 2013a; Ferenčaković et al., 2013b; Kim et al., 2013) have been largely using commercial populations that have been developed under a considerable impact of human-driven selection. Using a high density SNPs array (600K) we map ROH burdens that significantly differ between the two chicken populations and sex groups and among the twenty eight autosomes (GGA1–28). The ROH islands mapped are mostly located in genic regions. For the first time, we map twenty three chromosomal isodisomy in twenty two Ethiopian village chickens and to fouteen autosomes using a ROH analysis.

Materials and methods

The study populations

This study involves two village chicken populations (Horro and Jarso) sampled from two geographical regions (~ 870 km apart) of the Oromia regional state in western and eastern Ethiopia, respectively. The two chicken populations have been maintained and developed in considerably contrasting production systems with

respect to their agroecological setup and demographic structure of the farmers (Desta et al., 2013).

Defining runs of homozygosity and data analysis

The DNA library used for genotyping was extracted from a whole blood sample spotted on FTA® card using Smith & Burgoyne (2004) method 4. For quality control (QC) of SNPs and samples we used GenABEL package (Aulchenko et al., 2007) of the R (R Core Team, 2013). The QC criteria adopted were per indivdual and SNP calling rate (≥ 90%), minor allelic frequency (MAF, ≥ 1%,), ibs.mrk (All), ibs.treshold (≤ 95%) and ibs.exclude (lower). From a raw data that contains 546120 autosomal SNPs (GGA1–28) and 760 chickens, four chickens (one from Horro and three from Jarso) were excluded due to high IBS (identity by state) and 29012 and 1552 SNPs were excluded due to low MAF and low call rate respectively. The QC pruned in 515558 SNPs and 756 chickens (383 from Horro and 373 from Jarso). We havn't accounted for linkage disequilibrium (LD), because from an exploratory analysis we perform (data not shown), a mild LD ($r^2 \sim 0.2$) rarely exceeds 100Kb.

Calling of ROH was performed using PLINK software (Purcell et al., 2007). We set the minimum threshold of ROH tract to 200Kb contiguous homozygous genotypes because the chicken population studied are characterized by a weak LD. For other parameters we used the default setting (a sliding window of 5Mb and 50 SNPs, the number of heterozygote allowed per window is 1, the number of missing genotype allowed per window is 5 and the proportion of overlapping windows that must be called homozygous is 0.05). We used --homozyg function of PLINK to call a ROH. ROH islands (ROH hotspots) were defined as those representing 0.1 percentile of the upper tail region of the hom.overlap, while ROH coldspot regions represent 0.1percentile of the lower tail region of hom.summary file. A putative uniparental disomy was defined when a single ROH segement includes all the SNPs located on a chromosome. Genes that are located in genomic regions showing a considerable overlap among samples were mapped to Galgal4 built of BioMart portal of the Ensembl genome browser. Functional annotation of putative genes was performed using DAVID (http://david.abcc.ncifcrf.gov/, Huang et al., 2009a & b). We used a KEEG pathway as implemented in DAVID to identify genes involving in a pathway

using the list of putative genes identified by BioMart. We used the classification strigency “high” option in functional annotation cluster analysis of DAVID and we select a minimum enrichment score of 1.3 as a cut-off value. ROH was mapped at each autosome and individual chicken level. Summary statistic of the ROH parameters was performed using SPSS 21.0.0.0 (IBM Corp., 2012) and R, while plotting was performed using the base R, ggplot2 (Wickham, 2011) and doBy (Højsgaard et al., 2012) packages of the R. The function summarySE was retrived from http://www.cookbook-r.com/Manipulating_data/Summarizing_data/ and was used to plot mean values of the ROH parameters and the corresponding standard errors.

Individual’s level autozygosity (inbreeding) which is also referred as F_{ROH} was calculated using equation 1. We calculate F_{ROH} for three categories of ROH tracts: ≥ 200Kb ($F_{ROH0.2}$), ≥ 500Kb ($F_{ROH0.5}$) and ≥ 1000Kb ($F_{ROH1.0}$). Moreover, we used --het function of PLINK to calculate inbreeding (F_{PLINK}), which then compared with the three classes of the F_{ROH} defined.

$$F_{ROH} = \frac{KB_{TOTAL}}{L} \qquad (1)$$

where KB_{TOTAL} is the sum of the ROH tracts mapped in each chicken and L the total length of the genome covered by the SNPs included in the ROH analysis. Welch t.test (Welch, 1947) as implemented in R and SPSS 21.0.0.0 were used to perform a comparative statistical test between the two chicken populatons and hen and cock groups.

Analysis of variance for sum of ROH tracts (KB_{TOTAL}) was performed using equation 2.

$$y = \mu + a_i + b_j + c_k + d_l + abi_j + ac_{ik} + ad_{il} + bc_{jk} + bd_{jl} + abci_{jk} + abdi_{jl} + acd_{ikl} + bcd_{jkl} + abcd_{ijkl} + \varepsilon \qquad (2)$$

Where y is KB_{TOTAL}; μ the common mean, a the number of ROH segments (NSEG); b the autosomes (twenty eight levels); c the two chicken populations; d the two sex groups and ε unexpained error term. Two-, three- and four-way

interactions were performed among the explanatory variables to assess their combined effect.

We map protein coding genes that are located in the ROH island regions using the BioMart portal of the Ensembl genome browser built for the chicken (http://www.ensembl.org/Gallus_gallus/Location/Genome).

Definition of terms

Here we define the ROH terms used in the analysis. KB_{IND} refers to indivudal ROH segment that was mapped at chromosome and individual level. KB_{CHR} is the sum length of ROH segments in Kb mapped for each autosome and in each chicken. KB_{TOTAL} refers to the total sum length of ROH segments mapped across the twenty eight autosomes in each chicken. KB_{AVG} refers to the ratio of KB_{TOTAL} to the total number of ROH segments mapped in each chicken. $KB_{OVERLAP}$ refers to an overlapping region of KB_{IND} produced by a group of chickens.

Results

We map ROH burden and isodisomy at autosomal and individual chicken level. Statistical tests were performed between the two chicken populations and sex groups, and among the twenty eight autosomes (GGA1–28) to assess variabilty in ROH burden. We define ROH as runs of at least 200Kb homozygous genotypes to reduce the proportion of short ROHs that may commonly arise due to LD.

Genome-wide runs of homozygosity

A summary statistics of ROH analysis performed on twenty eight autosomes (GGA1–28) at population and sex group level is presented in Table 6.1. Congurent to chromosome-wide analysis (see the supplementary information in the Appendix 6), analysis of the total ROH burden of the twenty eight autosmes shows statistically significant difference between the two chicken populations and for some of the ROH parameters between sex groups (Table 6.1). However, a high intrapopulation variation in ROH burden may indicate variation in extent of inbreeding even down to

a family flock level. At autosomal genome level, number of ROH segments (NSEG) and the total sum of ROH segments (KB_{TOTAL}) show a significant positive correlation (r = 0.122, $P < 0.0001$), however, this correlation is lower than the chromosome-wide result (Supplementary Information). A high variation in ROH burden among the autosomes may have contributed for low correlation at genome-wide level.

Analysis of variance including main and interaction effects of four explanatory variables (NSEG, chromosome, population and sex) shows a significant impact of the main and interaction effects except the three-way interaction among chromosome, sex and population ($P = 0.199$). A model fitted in the ANOVA expalins a large proportion of the variation (adjusted R-square = 74.2%). Both NSEG ($t_{754} = 15.372$, $P < 0.0001$) and KB_{TOTAL} ($t_{754} = 6.463$, $P < 0.0001$) are signficantly lower in Horro than Jarso chickens.

Table 6.1 Genome-wide ROH parameters calculated for Horro and Jarso chickens and hen and cock groups.

Population	Statistics	NSEG	KB_{TOTAL}	KB_{AVG}
Horro(n = 383)	Mean (SD)	230.37(34.318)	181852.27(87559.28)	848.800(567.564)
	Median	231	137084.16	571.881
	Range	56–419	24287.19–422751.62	351.451–4295.526
Jarso(n = 373)	Mean (SD)	279.04(51.286)	224345.65(93182.313)	809.024(475.841)
	Median	284	190156.00	561.920
	Range	138–648	76920.75–533202.67	360.220–2628.949
Total(n = 756)	Mean (SD)	254.38(49.847)	202817.92(92785.60)	829.165(524.350)
	Median	249	173476.16	566.777
	Range	56–648	24287.19–533202.67	351.451–4295.526

Scatterplot and regression analyses performed on KB_{TOTAL} and NSEG at population level show moderate relationship (Figure 6.1 & 6.2). NSEG shows low variability than KB_{TOTAL} (Figure S6.1 & S6.2). A high varibility was observed for both NSEG and KB_{TOTAL} in Jarso than Horro chickens (Figure S6.1 & S6.2).

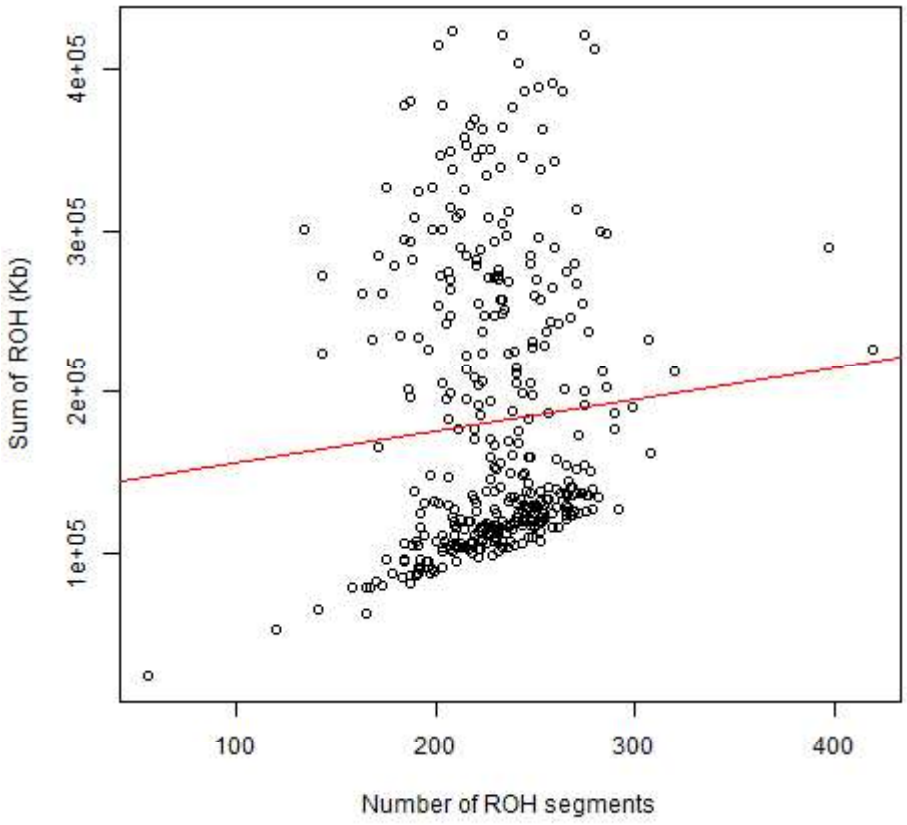


Figure 6.1 Scatterplot and linear regression analyses for sum of ROH tracts (KB_{TOTAL}) versus number of ROH segments (NSEG) in Horro chickens.

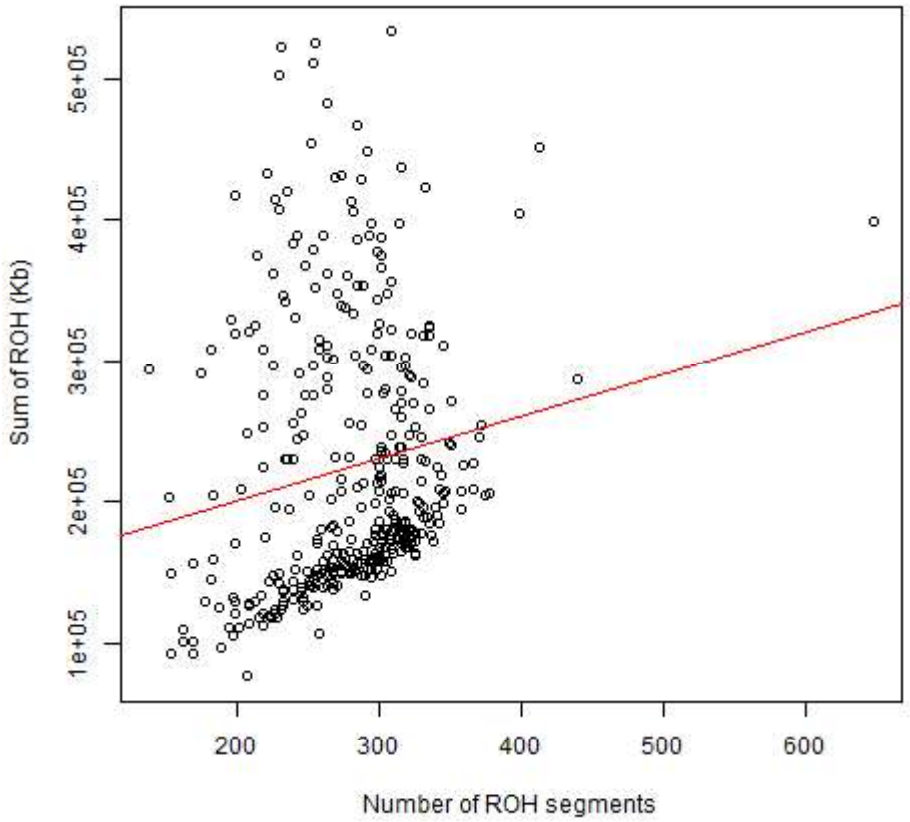


Figure 6.2 Scatterplot and linear regression analyses for sum of ROH tracts (KB_{TOTAL}) versus number of ROH segments (NSEG) in Jarso chickens.

Scatterplot and regression analyses performed for the number of SNP (NSNP) making a ROH tract and physical size of the individual ROH (KB_{IND}) at population level show strong relationship and are presented in Figure 6.3 & 6.4. This shows that as the physical size of the individual ROH segment increases, the number of SNPs forming a ROH segment increase in a sort of linear fashion.

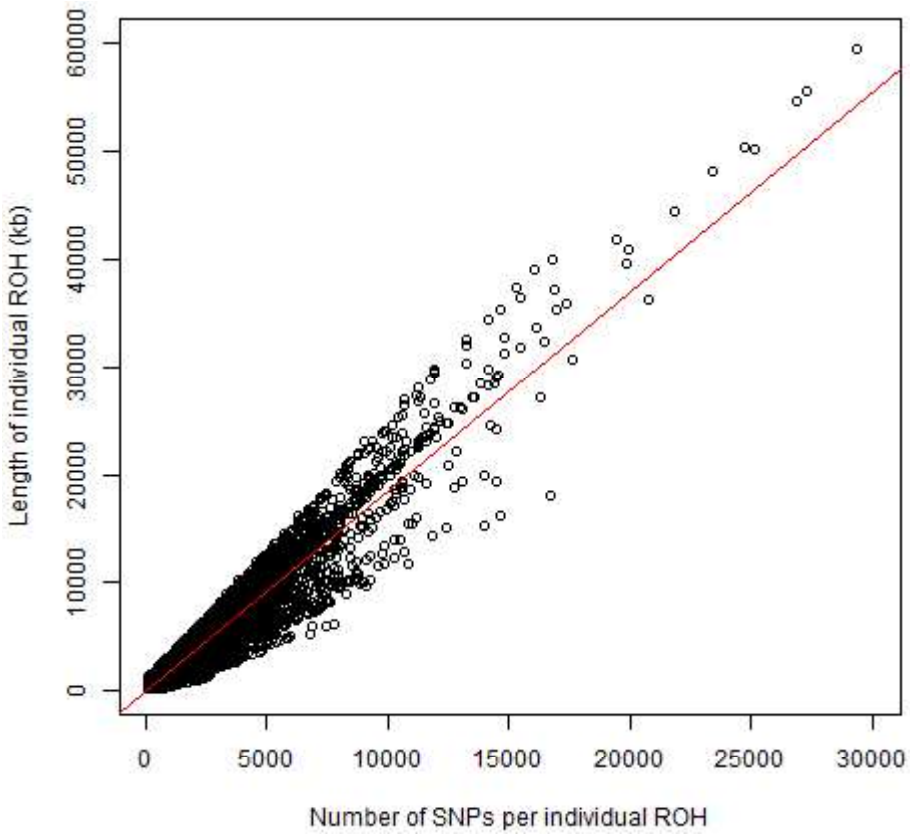


Figure 6.3 Scatterplot for length of individual ROH tract (KBIND) versus the number of SNPs forming each individual ROH tract (NSNP) in Horro chickens.

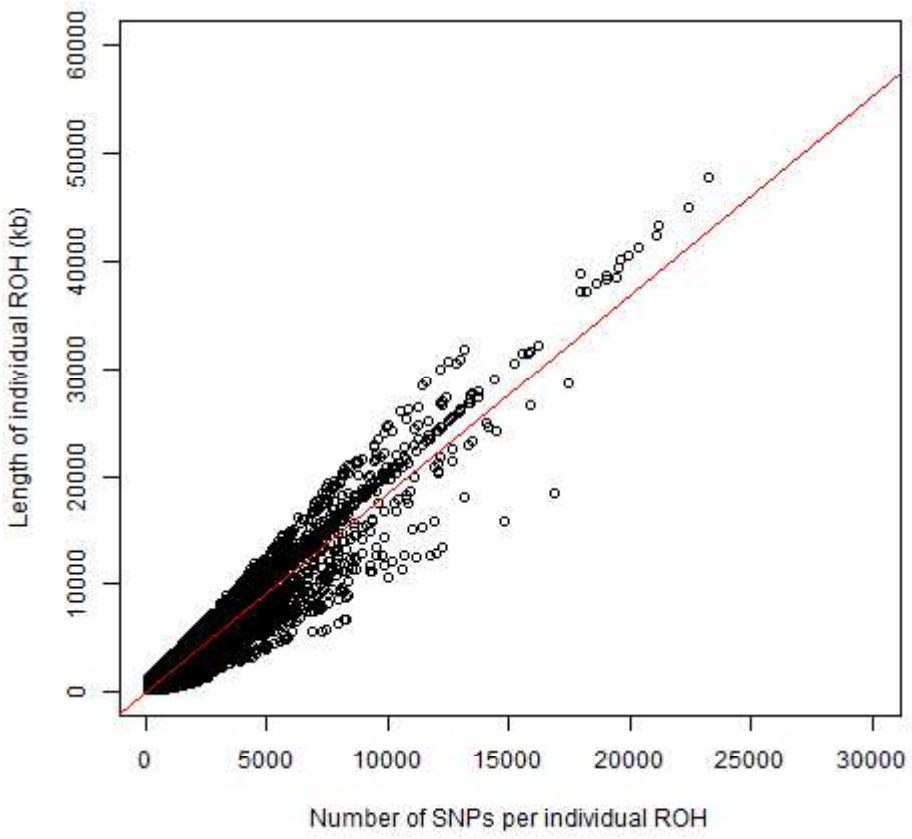


Figure 6.4 Scatterplot for length of individual ROH tract (KBIND) versus the number of SNPs forming each individual ROH tract (NSNP) in Jarso chickens.

A high variability observed in the NSEG mapped across the twenty eight autosomes (Figure 6.5) may be partly explained by a wide variation in recombination rate and physical size among the chicken autosomes (Hillier et al., 2004). The macrochromosomes (GGA1–5) show high ROH burden than the remaining twenty three autosomes (Figure 6.5). The NSEG mapped commonly ranges from 150 to 350 at individual chicken level, while the KB_{TOTAL} shows a wide range (25 to 533Mb). The average size of KB_{IND} was 797.288Kb (median = 378.709Kb and range: 200.005 – 59545.19Kb) with a corresponding standard deviation of 1709.307Kb, which exhibits a high coefficient of variation (214.39%). At the entire population level 2.18% of the KB_{IND} are ≥ 5Mb (this includes 2.3% (2068/88232) in Horro chickens and 2% (2130/104083) in Jarso chickens) and these large ROH tracts may occur due to the impact of a recent consanguineous mating. Genome of the two chicken poulations is mostly populated by short ROHs (200–500Kb). For exmaple, short ROH represent 68% (59998/88232) and 62% (64762/104083) of the ROH segments mapped in Horro and Jarso chickens respectively. Both the most frequent

NSEG (n = 648) and the extensive ROH burden ~ 533Mb (~ 58% of the autosomal genome) were mapped to two Jarso chickens. The longest KB_{IND} was found in a Horro chicken on GGA5 and it covers ~ 59.5Mb, *i.e.*, ~ 100% of the GGA5 and it contains 29394 contiguous homozygous SNPs (*i.e.*, all the SNPs included in ROH analysis of GGA5). At autosomal level, the most frequent NSEG (n = 131) was mapped to a Jarso chicken on GGA1 covering ~ 90Mb (46.2% of GGA1). An extensive ROH burden found in some chickens may be a consquence of an ongoing mating that involves closely related individuals. The commoness of short ROHs (200-500Kb) may be partly associate with LD and impact of selective sweep and may also indicate the panmictic nature of the two chicken populations.

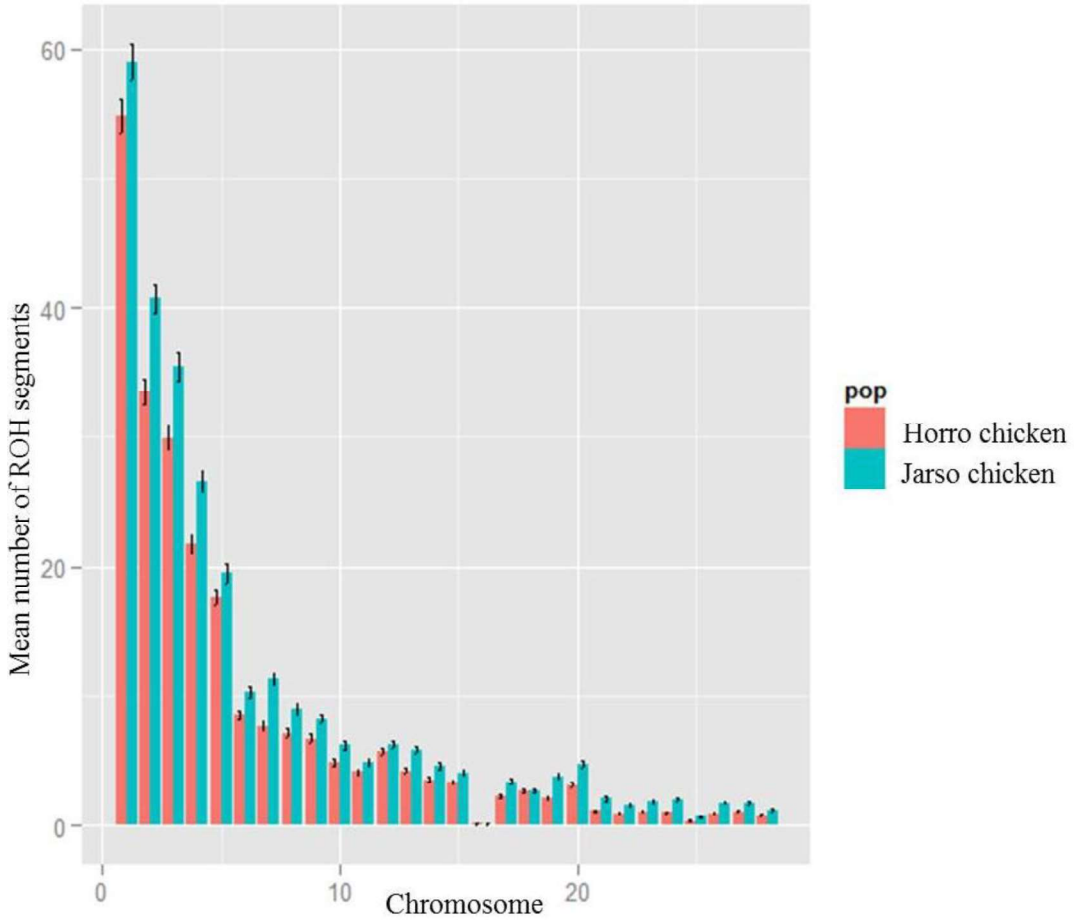


Figure 6.5 The mean number of ROH segments (NSEG) mapped in each autosome and population. The x-axis represents the twenty eight autosomes (GGA1–28) and the y-axis represents mean number of ROH segments and their associated standard error at chromosome and population level.

Inbreeding

Inbreeding coefficient was calculated at population and sex group level using the three classes of ROH defined based on the physical size of the ROH tract and the summary statistics is presented in Table 6.2. F_{ROH} is significantly differ both between the two chicken populations and the sex groups (Table 6.2). The inbreeding coefficient calculated using PLINK (F_{PLINK}) was significantly different from the F_{ROH} calculated for each of the three datasets and its value falls between $F_{ROH0.5}$ and $F_{ROH1.0}$. On average ~ 19.8 and ~ 24.4% of the autosomal genome (~ 919Mb) was covered by $ROH_{0.2}$ (a ROH tract of ≥ 0.2Mb) in Horro and Jarso chickens respectively and $F_{ROH0.2}$ accounted for ~ 18.32% of the autosomes physical size at the entire dataset level. However, these proportions are lower in the intermediate ($F_{ROH0.5}$) and long range ROHs ($F_{ROH1.0}$) (Table 6.2). Similarily, in hen and cock populations ~ 21.5 and 23.2% of the genome was covered by $F_{ROH0.2}$ (Table 6.2). As it has been found in hen and cock populations at the entire dataset level, there is a significant difference in $F_{ROH0.2}$ between Jarso hens and cocks (t_{371}= 2.484, P = 0.013), however, it is not visible (t_{381}= 1.318, P = 0.188) between Horro hens and cocks.

Table 6.2 The three classes of FROH calculated at population and sex group level.

Population	F_{ROH}	Mean(SD)	Median	Range	t_{754}	P
Horro (n = 383)	$F_{ROH0.2}$	0.198(0.095)	0.150	(0.026, 0.460	6.463	< 0.0001
	$F_{ROH0.5}$	0.145(0.099)	0.093	(0.011, 0.426)	5.251	< 0.0001
	$F_{ROH1.0}$	0.113(0.101)	0.059	(0.007, 0.401)	3.284	0.001
Jarso (n = 373)	$F_{ROH0.2}$	0.244(0.101)	0.207	(0.084, 0.580)		
	$F_{ROH0.5}$	0.184(0.105)	0.138	(0.024, 0.536)		
	$F_{ROH1.0}$	0.138(0.107)	0.086	(0.004, 0.501)		

The probability values indicate the statistically significant difference level observed between Horro and Jarso chickens and hens and cock populations for the three classes of the F_{ROH} defined.

Overlapping regions of the runs of homozygosity

In total, 6555247 overlapping ROH segments ($KB_{OVERLAP}$) were mapped (2959002 belong to Horro while 3596245 represent Jarso chickens), however the proportions are not significantly different (χ^2_1 = 3.1178, p-value = 0.9226). The average NSNP forming the $KB_{OVERLAP}$ and NSIM (number of ROH tracts that show similarity with a ROH tract in the same pool) matching with $KB_{OVERLAP}$ and their corresponding standard error at autosome and population level are presented in Figure 6.6 & 6.7. Moreover, a summary statistic is presented in Table 6.3. Consensus pools consist of 50 to 572 chickens and are made by 1 to 105 overlapping SNPs with a corresponding $KB_{OVERLAP}$ of 0 to 281.124Kb. The $KB_{OVERLAP}$ mapped produced 40491 genomic regions representing consensus pools of the $KB_{OVERLAP}$.

Among the 15004 autosomal genes (GGA1–28) located in ENSEMBL Gallus gallus genes (Galgal4) (http://www.ensembl.org/Gallus_gallus/Location/Genome), 8271 (55.13%) are located in the $ROH_{OVERLAP}$ regions. Large numbers of genes mapped in $KB_{OVERLAP}$ regions indicate their vital role in a range of biological functions, which then may have subjected them to a high selection pressure that enables them to maintain their sequence identity.

Table 6.3 Number of SNP forming each an overlapping ROH segment and the number of similar segments mapped for each overlapped ROH segment.

Population	Statistics	NSNP	KBoverlap	NSIM
Horro(*n* = 383)	Mean (SD)	2570 (3772.185)	4958.3 (7647.561)	48.97 (72.09047)
	Median	760	1441	15
	Range	100–29394	200–59545.2	0–536
Jarso(*n* = 373)	Mean (SD)	2143 (3240.651)	4121.6 (6474.736)	54.35 (67.34637)
	Median	656	1237.5	31
	Range	100–23280	200–47854.8	0–536
Total(*n* = 756)	Mean (SD)	2335 (3497.07)	4499.3 (7040.741)	51.92 (69.57947)
	Median	693	1310.6	24
	Range	100–29394	200–59545.2	0–536

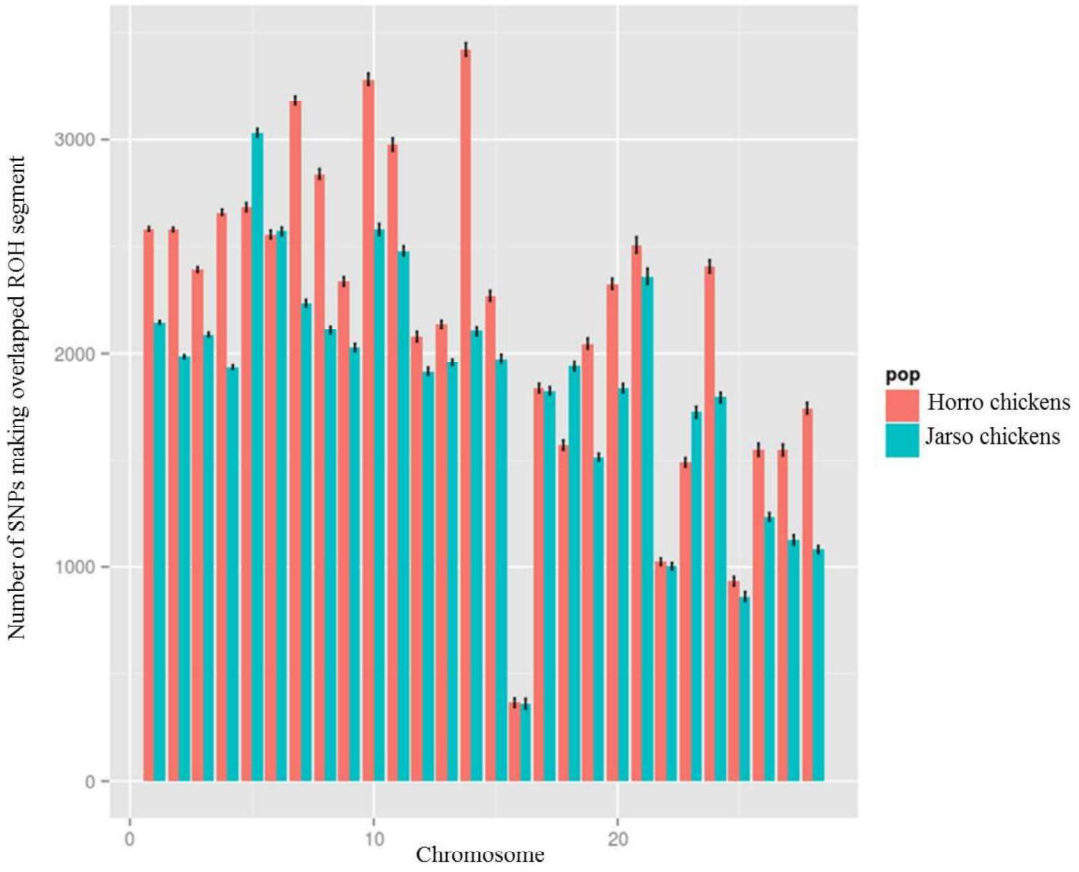


Figure 6.6 Number of SNPs forming overlapped ROH segment. The x-axis represents the twenty eight autosomes, while the y-axis refers to average number of SNPs forming each overlapping ROH segment.

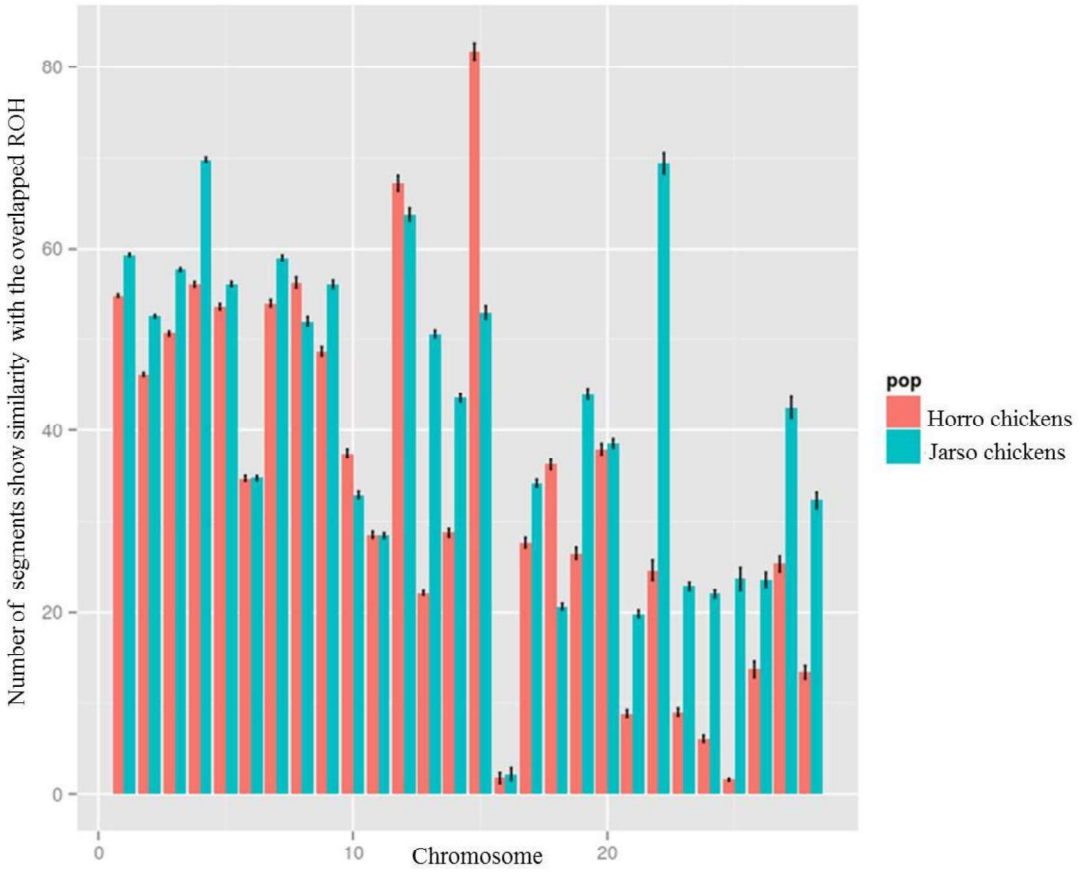


Figure 6.7 Number of ROH segments showing similarity with an overlapped ROH segment. The x-axis represents the twenty eight autosomes, while the y-axis refers to average number of ROH segments showing similarity with overlapping ROH segment. The vertical line on each bar represents the standard error associated with mean value of NSIM.

We defined forty one ROH islands (representing 0.1th percentile of the upper tail region of hom.overlap consensus region) based on the size of an overlapping ROH segment and number of chickens making the the overlapping ROH region. Fifty four genes are located in the ROH island regions mapped to GGA1–3, 6–12, 14–16, 18, 21, 23, 24 and 26–28. These genes make nine functional annotation clusters two out of them have enrichment score of 1.42 and 2.92. These genes involve in four pathways– GnRH, VEGF, MAPK and Toll-like receptor signalling pathways. The top ten ROH island regions are presented in Table 6.4.

Table 6.4 Top ten runs of homozygosity hotspot regions.

NIND	Chr	Start	End	Length	NSNP	Associated genes
570	8	432469	528393	95.9240	36	
482	1	160260238	160334977	74.7390	24	
480	15	2919680	3020111	100.4310	67	HIC2, STX2
421	4	27389741	27461251	71.5100	28	
350	8	9575696	9805933	230.2370	100	PLA2G4A, ENSGALG00000028389
337	26	95603	225094	129.4910	75	SRPK1, SLC26A8, MAPK14, MAPK13, BRPF3, TBC1D22B, PIM1
337	20	375769	464649	88.8800	22	
336	5	3295508	3458052	162.5440	29	
334	2	52242889	52524013	281.1240	105	ENSGALG00000012353, VOPP1, LANCL2, EGFR
334	2	52852275	53055039	202.7640	67	TPK1

NIND refers to the number of indidvudals while NSNP represents the number of SNPs.

We also identified ROH coldspot regions. Most of the SNPs that form a ROH segment in a few individuals are located on GGA16 & 25 (84.7%); however, since these two autosomes are under-represented by the SNP chip, we decided to make our discussion on the remaining autosomes (GGA1, 3, 4, 7, 19 & 28) to which a ROH coldspot region was identified. The five genes that are located in ROH coldspot regions may represent fastly evolving loci. These genes produce neither functinonal annotation cluster nor pathway. Moreover, these colsopt regions may locate in recombination hotspot regions. Ten genomic regions that show a substantial level of ROH coldspot are presented in Table 6.5.

Table 6.5 ROH coldspot regions mapped in Horro and Jarso chickens.

Chr	**Start**	**End**	**Length**	**NIND**	**Associated gene**
19	2682769	2685905	3136	34	GTF2I
4	43422521	43425044	2523	34	
28	4738768	4742489	3721	38	R3HDM4
28	4738025	4738375	350	39	
1	195272878	195273389	511	47	
1	195266168	195271650	5482	48	ENSGALG00000017301
1	184908786	184909730	944	49	PANX1
1	195264838	195265732	894	49	SLCO2B1
19	2688651	2689517	866	50	
1	184913552	184914376	824	50	

NIND refers to the number of individuals to which ROH coldspot regions are identified.

Uniparental disomy

A ROH analysis identified twenty three putative uniparental disomy (UPD) – a meiotic mis-segregation (cytological abnormality). Putative uniparental disomy was mapped in twenty two chickens including thirteen from Horro (3.39%) and nine from Jarso (2.41%) and to fourteen autosomes (GGA5, 7, 11–16, 20–22 and 26–28). Among the twenty two chickens to which putative isodisomy was mapped, nine (four from Horro and five from Jarso) are cock while thirteen represent hen (nine from Horro and four from Jarso). Two of the putative UPDs that belong to GGA11 & 14 were mapped to a Horro cock. Putative UPD is most frequently mapped to GGA15 (n = 4) followed by GGA16 (n = 3), however, due to low genomic coverge of GGA16 by the SNP chip, putative UPDs mapped to GGA16 rather comfortably considered as suggestive. GGA20, 21, 26 and 28 each has two UPDs, while GGA5, 7, 12–14, 22 and 27 each represented by a single putative UPD. The percentage of heterozygote genotypes mapped in the fourteen autosomes and the twenty two chickens showing putative UPD ranges from 0.002 to 0.014, which represents a small proportion , which then validates the accuracy of our UPD mapping approach. Moreover, the DENSITY (the number of SNPs forming each KB of a ROH tract) of SNPs in the fourteeen autosomes showing putative UPD ranges from 0.8 to 2, which substantiates a dense coverage of the genome. However, re-sequencing the entire genome of the twenty two chickens will provide undisputable evidence.

Discussion

Landraces of village chickens that populate the least developed world have been developed under uncontrolled breeding and through a mild impact of artificial selection mainly intended for traits of visual significance. Uncontrolled breeding in village chickens results in an outbred population. Inbreeding is thought to be less intense in outbred populations. However, consanguineous mating, small family flock size, infrequent intermix among scavenging chickens due to local physical and/or socio-cultural barriers may lead to formation of ROH even in outbred village chickens. Morover, an extended selective sweep in genomic regions harbouring major loci that confer local adaptation may also lead to formation of ROH. We perform mapping of ROH using large number of informative SNPs assayed to the

autosomal genome of two Ethiopian village chicken populations that have been maintained under considerably contrasting production environments (Desta et al., 2013). Although it is more appropriate to map ROH using a whole genome resequence data, a high density SNP array can also perform a comparable job in absence of a sequence data.

ROH may represent a genomic tract identical by state, which then not necessarily implicates identical by descent (IBD) (MacLeod et al., 2013). ROH is not uniformly distributed across the autosomes, between the two chicken populations and even among individuals at intrapopulation level. As the ROH burden mapped for some chickens thought to be extensive for panmictic populations and for chickens exhibting a high rate of chromosomal recombination, extensive ROH is possibly the consquence of a recent consanguineous mating.

Extensive old ROHs (long ROHs that are shared by large nummber of individuals) may commonly found in recombination coldspot regions. However, when the chickens have a common ancestor in the recent past, they may share a considerable part of their genome as identical by descent (IBD). Moreover, if mating is occurred continiously within a family flock or among few flocks with a considerable absence of (minimum) external gene flow, there is a high chance for formation of ROH even in outbred populations (Purfield et al., 2012) due to the impact of a localised genetic drift. An extended isolation and reduction in effective population size in the course of the breeding history of a population may also contribute to formation of ROH (Nothnagel et al., 2010). Populations that have large effective population size and exhibiting genetic homogenity (due to intricate web of gene flow) likely have smaller proportion of ROH than isolated populations that are under considerable impact of genetic drift and inbreeding.

Individuals that share a recent common demographic history and/or ancestor tend to show a similar pattern of ROH burden. ROH burden is considerably impacted by the demographic history of a population (Bosse et al., 2012 and the references therein), mating pattern and prolificacy of a species. Distance from center of domestication and anthropogenic effects may also alter the extent of a ROH burden. Founder effect, genetic bottleneck, reduction in effective population size, long generation interval

and lifetime production of a few offspring, geographical isolation and inbreeding likely increase ROH burden due to the associated impact of genetic drift. Neverthless, demographic factors like population expansion (increase in effective population size), a standing gene flow, a similar proportion of breeding males and females tend to increase genetic diversity, which then reduce ROH burden.

Burden of ROH unfavourably correlates with time to common ancestor. ROH of recent origin are extensive as they have been subjected to few meioses, while short ROHs likely be of an ancient origin (Purfield et al., 2012; Ferenčaković et al., 2013a & b). Short ROHs may also reflect a background relatedness that survive the impact of histrorical recombination and genetic admixture. Short ROH is common in panmictic populations experiencing a considerable outbreeding and that have high genetic diversity. In genetically diverse populations like village chickens, the common ancestor may trace back to ancient time and historical recombination could have had ample chance to breakdown long range ROHs. Island of ROH is characterized by a low genetic diversity and may occur due to recombination coldspot, genetic drift and strong selective sweep. We scan the entire autosomal genome to map overlapping ROH regions and to investigate the relationship between ROH burden and density of genes and we found that overlapping ROH is more frequent in genic regions. ROH islands have been commonly mapped in genomic regions harbouring causative variants underlying the genetic basis of economically important traits (Kim et al., 2013). However, Bosse et al. (2012) reported absence of correlation between ROH hotspot and gene density although they report as ROH hotspot regions represent loci that are subjected to a positive selection.

ROH is infrequent in large genetically homogenous outbred populations (see Bosse et al., 2012 for a review). In randomly mating populations, ROH decays quickly due to meiotic recombination and extensive genetic admixture (Bosse et al., 2012). Formation of ROH is the function of the demographic history of a population and recombination rate. Effective population size in village chickens that haven't been selected for any traits of particular interest is high, then the impact of genetic drift that otherwise leads to fixation of alleles and formation of ROH is most likely less evident. The ROH mapped in the two chicken populations shows uneven distribution. Uneven distrbution of ROH could occur due to variation in

recombination rate across the chicken genome (Hillier et al., 2004), among populations and even among individuals within a population (Dumont & Payseur, 2011). A selection pressure that ascends the advantageous allele could also result in ROH particularly in QTL regions with a large effect, which may also contribute to uneven distribution of ROH (Wang et al., 2009; Pemberton et al., 2012). Although the main driving force behind ROH formation thought to be inbreeding (Ferenčaković et al., 2013a & b), the ROH islands mapped by our analysis indicate that functional diversity may also has a considerable impact.

A low proportion of long range ROHs mapped by our analyses is most likely the impact of a random mating and an extensive genetic admixture going on in the two chicken populations. The real ROH burden could be even less intense, because homozygous SNPs may also arise due to ascertainment bias of a SNP chip, which was developed using genomic information of commerical populations (Kranis et al., 2013). We map a large proportion of ROHs in the macrochromosomes and this may associate with favourable relationship between burden of ROH and physical size of a chromosome (Nothnagel et al., 2010) and due to low recombination rate in the macrochromosomes (Hillier et al., 2004). Variation in ROH burden across chromosomes may be more evident in chickens compared to other livestock species as the former have a wide variation in recombination rate among the autosomes (Hillier et al., 2004). ROH burden may also vary between sex groups. Difference in ROH burden may arise from sex-biased recombination rate. For example, in mammals genetic distances (recombination size) in males are less than females (Burt et al., 1995), which could result in a long range ROH in males.

ROH could also rarely arise from imbalance in meiotic segregation including the UPD (Figure 6.8). UPD refers to inheritance of a pair of homologous chromosomes from a single parent in absence of a copy from a partner (Engel, 1980; 2006). UPD is however a rare event, because it requires two independent chromosomal non-disjunctions to occur (Dawson et al., 2011). The putative UPD mapped by our analysis represents duplicated copies of a homologous chromosome inherited from one of the parents. The proportion of chickens having putative UPD (2.9%) is higher than human (0.2%, 1 in 500) (Schinzel & Baumer, 2011). UPD can mainly occur due to trisomy correction; monosomy duplication and gamete complementation (consult

Kotzot & Utermann, 2005; Engel, 2006; Dawson et al., 2011; Schinzel and Baumer, 2011 and the references therein). Improper segregation of a pair of homologous chromosome into two daughter cells in meiosis I and II could form UPD (Dawson et al., 2011). Mitotic nondisjunction could also lead to isodisomy (Schinzel and Baumer, 2011). Given the high proportion of a chromosomal putative UPD found in the chicken populations studied, some of the long range ROH tract mapped may also partly arise from another form of UPD – segmental UPD, which involves duplication of a part of a chromosome. As it has been explained in the RESULT section, analysis of a whole genome resequence data from chickens that have isodisomy will provide better evidence. Studying the segregation pattern of the putative UPDs mapped would help to classify the isodisomic chromosomes to maternal or paternal origin. However, due to absence of documented record on familial relationship, high flock turnover and short lifespan of the chicken, this would be hardly possible. In humans, maternal UPD is more frequent than paternal (Kotzot & Utermann, 2005). The impact of UPD on the performance of chicken would be the subject of a future study.

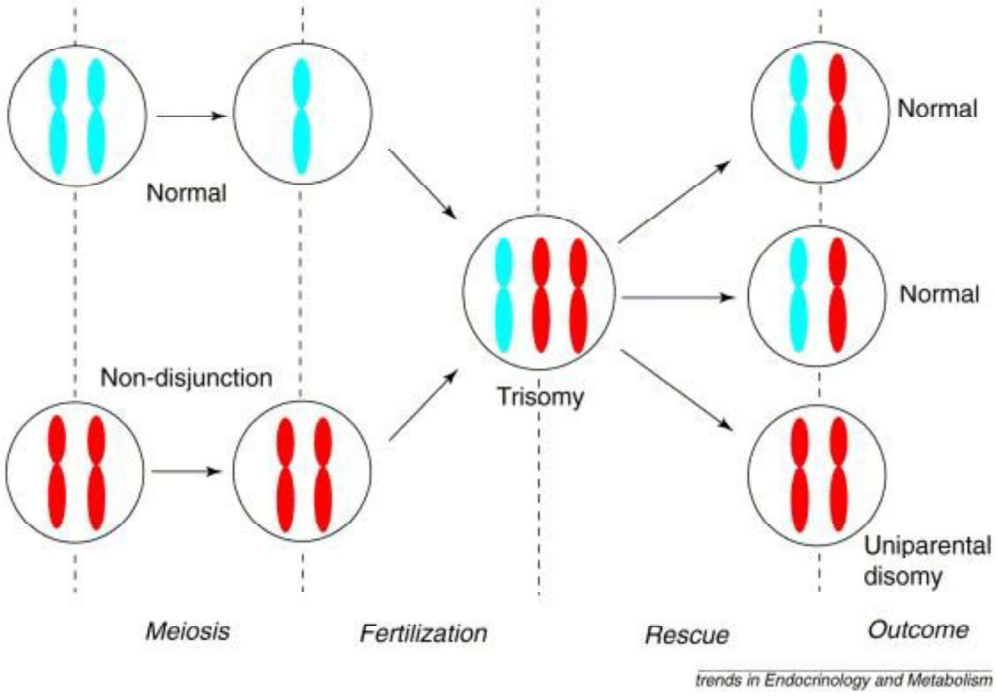


Figure 6.8 The schematic presentation of uniparental disomy (source: Preece & Moore, 2000).

Conclusion

The ROH mapped in some chickens could be considered as extensive for populations that are thought to be outbred, although the proportion of long range ROH is low. There is significant variation in ROH burden between population and among the autosomes. Moreover, there is a considerable variation in ROH burden even at autosomal and intrapopulation level. ROH burden is more frequent in genic regions, which implicates the impact of natural selection in the formation of ROH. Moreover, the demographic history of a population could result in variation of ROH burden. The putative uniparental disomy mapped requires further study to uncover the segregation pattern of this chromosomal non-disjunction. A high proportion of ROH burden mapped in some chickens implies the need to revise the existing village chicken breeding system.

References

AULCHENKO, Y. S., RIPKE, S., ISAACS, A. & VAN DUIJN, C. M. GenABEL: an R package for genome-wide association analysis. Bioinformatics 2007, 23(10), 1294-1296.

BOSSE, M., MEGENS, H. J., MADSEN, O., PAUDEL, Y., FRANTZ, L. A., SCHOOK, L. B., ... & GROENEN, M. A. Regions of homozygosity in the porcine genome: consequence of demography and the recombination landscape. PLoS Genetics, 2012, 8(11), e1003100.

BURT, D. W., BUMSTEAD, N., BITGOOD, J. J., DE LEON, F., & CRITTENDEN, L. B. Chicken genome mapping: a new era in avian genetics. Trends in Genetics, 1995, 11(5), 190-194.

DAWSON, A. J., CHERNOS, J., MCGOWAN-JORDAN, J., LAVOIE, J., SHETTY, S., STEINRATHS, M., ... & XU, J. CCMG guidelines: prenatal and postnatal diagnostic testing for uniparental disomy. Clinical Genetics, 2011, 79(2), 118-124.

DESTA, T. T., DESSIE, T., BETTRIDGE, J., LYNCH, S. E., MELESE, K., COLLINS, M., CHRISTLEY, R. M., WIGLEY, P., KAISER, P., TERFA, Z., MWACHARO, J. M. & HANOTTE, O. Signature of artificial selection and ecological landscape on morphological structures of Ethiopian village chickens. Animal Genetic Resources, 2013, 52, 1-13.

DUMONT, B. L. & PAYSEUR, B. A. Genetic analysis of genome-scale recombination rate evolution in house mice. PLoS Genetics, 2011, 7(6), e1002116.

ENGEL, E. A fascination with chromosome rescue in uniparental disomy: Mendelian recessive outlaws and imprinting copyrights infringements. European Journal of Human Genetics, 2006, 14, 1158-1169.

ENGEL, E. A new genetic concept: uniparental disomy and its potential effect: isodisomy. American Journal of Medical Genetics, 1980, 6, 137-142.

FERENČAKOVIĆ, M., HAMZIC, E., GREDLER, B., CURIK, I. & SÖLKNER, J. Runs of homozygosity reveal genome-wideautozygosity in the Austrian fleckvieh cattle. Agriculturae Conspectus Scientificus, 2011, 76, 325-328.

FERENČAKOVIĆ, M., HAMZIĆ, E., GREDLER, B., SOLBERG, T. R., KLEMETSDAL, G., CURIK, I. & SÖLKNER, J. Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. Journal of Animal Breeding and Genetics, 2013a, 130, 286-293.

FERENČAKOVIĆ, M., SÖLKNER, J., & CURIK, I. Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. Genetics Selection Evolution, 2013b, 45(1), 42.

GIBSON, J., MORTON, N. E., & COLLINS, A. Extended tracts of homozygosity in outbred human populations. Human Molecular Genetics, 2006, 15(5), 789-795.

HILLIER, L. W., MILLER, W., BIRNEY, E., WARREN, W., HARDISON, R. C., PONTING, C. P., ... & DODGSON, J. B. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature, 2004, 432(7018), 695-716.

HØJSGAARD, S., HALEKOH, U., ROBISON-COX, J., WRIGHT, K., & LEIDI, A. A. doBy: doBy-Groupwise summary statistics, general linear contrasts, population means (least-squares-means), and other utilities. R package version, 4-5, 2012.

HUANG, D. W., SHERMAN, B. T. & LEMPICKI, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research, 2009a, 37(1), 1-13.

HUANG, D. W., SHERMAN, B. T. & LEMPICKI, R. A. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protocol, 2009b, 4(1), 44-57.

IBM CORP. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp, 2012.

KIM, E. S., COLE, J. B., HUSON, H., WIGGANS, G. R., VAN TASSELL, C. P., CROOKER, B. A., ... & SONSTEGARD, T. S. Effect of artificial selection on runs of homozygosity in US Holstein cattle. Plos One, 2013, 8(11), e80813.

KIRIN, M., MCQUILLAN, R., FRANKLIN, C. S., CAMPBELL, H., MCKEIGUE, P. M., & WILSON, J. F. Genomic runs of homozygosity record population history and consanguinity. Plos One, 2010, 5(11), e13996.

KOTZOT, D & UTERMANN, G. Uniparental disomy (UPD) other than 15: Phenotypes and bibliography updated. American Journal of Medical Genetics, 2005, 136A, 287–305.

KRANIS, A., GHEYAS, A. A., BOSCHIERO, C., TURNER, F., YU, L., SMITH, S., ... & BURT, D. W. Development of a high density 600K SNP genotyping array for chicken. BMC Genomics, 2013, 14(1), 59.

MACLEOD, I. M., LARKIN, D., LEWIN, H., HAYES, B. J. & GODDARD, M. E. Inferring demography from runs of homozygosity in whole genome sequence, with correction for sequence errors. Molecular Biology and Evolution, 2013, 30(09), 2209-2223.

MCQUILLAN, R., LEUTENEGGER, A. L., ABDEL-RAHMAN, R., FRANKLIN, C. S., PERICIC, M., BARAC-LAUC, L., ... & WILSON, J. F. Runs of homozygosity in European populations. The American Journal of Human Genetics, 2008, 83(3), 359-372.

NOTHNAGEL, M., LU, T. T., KAYSER, M., & KRAWCZAK, M. Genomic and geographic distribution of SNP-defined runs of homozygosity in Europeans. Human Molecular Genetics, 2010, 19(15), 2927-2935.

PEMBERTON, T. J., ABSHER, D., FELDMAN, M. W., MYERS, R. M., ROSENBERG, N. A., & LI, J. Z. Genomic patterns of homozygosity in worldwide human populations. The American Journal of Human Genetics, 2012, 91(2), 275-292.

PREECE, M. A., & MOORE, G. E. Genomic imprinting, uniparental disomy and foetal growth. Trends in Endocrinology & Metabolism, 2000, 11(7), 270-275.

PURCELL, S., NEALE, B., TODD-BROWN, K., THOMAS, L., FERREIRA, M. A., BENDER, D., ... & SHAM, P. C. PLINK: a tool set for whole-genome association and population-based linkage analyses. The American Journal of Human Genetics, 2007, 81(3), 559-575.

PURFIELD, D. C., BERRY, D. P., MCPARLAND, S. & BRADLEY, D. G. Runs of homozygosity and population history in cattle. BMC Genetics, 2012, 13, 70.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, 2013. http://www.R-project.org/.

SCHINZEL, A., & BAUMER, A. Uniparental Disomy. Encyclopedia of Life Sciences, 2011.

SMITH, L. M., & BURGOYNE, L. A. Collecting, archiving and processing DNA from wildlife samples using FTA® databasing paper. BMC Ecology, 2004, 4(1), 4.

WANG, S., HAYNES, C., BARANY, F., & OTT, J. Genome-wide autozygosity mapping in human populations. Genetic Epidemiology, 2009, 33(2), 172-180.

WELCH, B. L. The generalization of Student's problem when several different population variances are involved. Biometrika, 1947, 34(1/2), 28-35.

WICKHAM, H. ggplot2. Wiley Interdisciplinary Reviews: Computational Statistics, 2011, 3(2), 180-185.

Chapter 7

The geography of selective sweep in Ethiopian village chickens

Abstract

Selection is ubiquitous evolutionary force that makes advantageous genomic variants to increase in frequency. Strong selection may have occurred during domestication. Selective sweep can be a consequence of natural and/or artificial selection. We map a number of loci that are under recent positive selection in two Ethiopian village chickens (Horro and Jarso) using a genome-wide high density (600K) SNP array. Our results show that in most instances the two chicken populations are divergently selected. Divergently selected loci among others involve in melanogenesis pathway. In most instances the genomic regions mapped were not overlapped among the selective sweep statistical tests and the two chicken populations. Variation in demography history, local adaptation and divergent selection pressure may have contributed to variation in selective sweep landscape. Selective sweeps were distributed across the autosomes with frequent and strong signals on GGA6 & 7. Functional annotation of the putative candidate genes shows their role in local adaptation and vital biological processes.

Introduction

A transient genetic bottleneck was imposed on domestic fowl during domestication (Rubin et al., 2010; Wiener & Wilkinson, 2011) following an abrupt change in habitat, exposure to novel selective pressures and a domestication process that presumably involved few captive fowls. However, the report of Rubin et al. (2010) on domestication genes was questioned by Flink et al. (2014) study that is based on ancient DNA analysis and they suggest that the evolution of some of the so-called domestication genes postdates domestication. Domestic fowl may have been largely reinstated its genetic diversity lost due to genetic bottlenecks introduced from domestication associated events by adapting to a wide range of environments and management systems. Habitat and population expansion and human preference for diverse phenotypes have resulted in a high genetic diversity of domestic fowl. Most

of the genetic variation that might be either fixed or lost due to founder events and an associated genetic drift could have been likely gradually reinstated due to recurrent natural selection and adaptive radiation of the domestic fowl. However, adaptive variation may have left informative imprints in domesticates genome. For example, as genetic characteristic tameness had been gradually developed following captivity, the domestic fowl may have had experienced a considerable loss of those genetic variants conferring alertness (aggression), which is inevitable to survive in a hostile environment. Advantageous *de novo* mutations that have occurred in the course of domesticates evolutionary history may have also swept closely linked loci through genetic hitchhiking.

Village chickens are evolved mainly under natural selection and through a mild impact of artificial selection, which is often associated with demographic structure of subsistence farmers (Desta et al., 2013). Village chickens have been developed under uncontrolled breeding, hypervariable production environment and suboptimal management system. In outbred populations like village chickens, less intense and mild signature of selection is expected. Unlike artificial selection that commonly focuses on visual traits that are thought to be controlled by a few loci, natural selection usually involves many loci; therefore, signature of the latter likely distribute across the genome. Selective sweeps arising from natural selection in some instances may involve major genes (QTL regions) explaining a large proportion of the genetic variation associated with a polygenic trait. This could form detectable footprints of natural selection. Selective sweep that arise from natural selection forms the basis of local adaptation and phenotypic evolution (Kim and Nielsen, 2004).

Selective sweep ascends the frequency of beneficial mutations and creates allelic frequencies distribution peculiar to the selected region and this could change the degree and pattern of genetic variation (Kim and Nielsen, 2004). Marker density, genome architecture (recombination rate and linkage disequilibrium) and selection pressure impact the power of detecting footprints of selection. For village chickens having a high genetic and phenotypic diversity and for fowl genome characterized by a high rate of recombination, large number of informative genetic markers are required to map a fine-scale selective sweep. Fixation of a beneficial allelic variant

due to selection creates genomic tracts displaying low polymorphism (Barrett & Schluter, 2008). Populations that have evolved under divergent selection pressure may share small proportion of selective sweep regions (Ramey et al., 2013). However, population that have been mainly evolved under natural selection may share selective sweep regions for loci that are invariably required in local adaptaion.

Pattern and intensity of selective sweep observed in panmictic village chickens may serve as a model to study the landscape of selective sweep in natural populations. We use two Ethiopian village chicken populations (Horro and Jarso) and a high density (600K) SNP array (Kranis et al., 2013) to map putative genomic regions that have been subjected to a recent positive selection. We map large number of putative loci invoving in local adaptation in the two chicken populations. However, in most instances the statistical tests mapped different regions in line with the assumption made by each test. Moreover, in most instances different selective sweep regions were detected in the two populations, which implicates a divergent selection.

Materials and methods

The study populations

This study involves two village chicken populations; Horro and Jarso sampled from two distantly located regions in the Oromia region of western and Eastern Ethiopia, respectively. The two chicken populations have been developed in considerably differing production environments with respect to their ecological setup and demographic structure of the communities (Desta et al., 2013).

Data analysis

The DNA was extracted from FTA® cards using Smith & Burgoyne (2004) method 4. For quality control (QC) of SNPs and samples we used the check.marker function of GenABEL package (Aulchenko et al., 2007) of the R (R Core Team, 2013). The QC criteria used were per individual and SNP calling rate (≥ 90%), minor allele frequency (MAF, ≥ 1%), ibs.mrk (All), ibs.treshold (≤ 95%) and ibs.exclude (lower). From a raw autosomal data that contain 546120 SNPs across GGA1–28 in 760

chickens, four chickens (one from Horro and three from Jarso) were excluded due to high IBS, and 29012 SNPs were excluded due to low MAF, while 1552 SNPs were excluded due to low calling rate. The QC retained 515558 SNPs and 756 chickens (383 from Horro and 373 from Jarso). We assigned the two alleles as ancestral and derived based on outgroup allelic polarity information of the junglefowl (grey (n = 2), green (n = 2) and Ceylon (n = 2)), pheasant (n = 3) and waterfowl (n = 3) genotyped using the same SNP chip. From 515558 autosomal SNPs qualified by the QC, 391384 SNPs that are monomorphic atleast in one of the three outgroup species were subsquently used in selective sweep analysis. We also extracted smaller subset of SNPs (n = 53533) that are invariably monomorphic in all the outgroup species and this smaller dataset has been used in iHS, Rsb ans SweeD analyses along with the bigger dataset (391384 SNPs).

We used a custom R script to prepare an input file for fastPHASE. Phasing of SNP data and imputation of missing genotypes was performed at population level using fastPHASE (Scheet and Stephens, 2006). We used the R package rehh (Gautier and Vitalis, 2012) to perform unstandardized integrated haplotype score (*iHS*) (equation 1) and *Rsb* (equation 2) test following Voight et al. (2006) and Tang et al. (2007) approach respectively. Each SNP was considered as a core SNP in the analyses. The standardized scores of *iHS* and *Rsb* test were plotted using a custom R script.

$$iHS = \ln\left(\frac{iHH_A}{iHH_D}\right) \qquad (1)$$

Where iHH_A and iHH_D refer to the integrated exended haplotype homozygosity in ancestral and dervied allele respectively.

$$Rsb = \ln\left(\frac{iESpop1}{iESpop2}\right) \qquad (2)$$

Where $iESpop1$ and $iESpop2$ refer to the integrated site specific extended haplotype homozygosity in Horro and Jarso chickens respectively.

Two-sided p-values were calculated for iHS and Rsb score using $1-2|\phi(iHS)-0.5|$ Gaussian cumulative density function because both high positive and low negative values indicate selection in ancestral and derived allele or in poulation one and two respectively. We used $-\log_{10}$ of the calculated p-values for plotting.

A chromosome-wide F_{ST} analysis weighted by sample size of each population was performed using a sliding window approach invloving ten contiguous syntenic SNPs with an overlapping window of eight SNPs. The calculated F_{ST} value was then averaged for each sliding window. Due to positively skewed values of F_{ST} (the assumption of normality could not apply), we used the raw value of 0.2 as a cut-off point to declare genetic divergence between the two chicken populations. However, for comparison, we also display the standardized score of F_{ST} value.

We used SweeD software (Pavlidis et al., 2013) to map genomic regions that have been subjected to strong selection pressure at population level. SweeD analysis was performed for each autosome independently. Signature of selection was identified based on alpha value (equation 3) following Pavlidis et al. (2013), the lower the alpha the stronger is the selective sweep. The standardized z-score of $-\log_{10}(\alpha)$ was plotted in R.

$$\alpha = r\ln(2N)/s$$

Where α is the intensity of selective sweep, r recombination rate, N effective population size and s selection coefficient.

We quantify a chromosome-wide variation in LD pattern between the two chicken populations using varLD software (Ong & Teo, 2010). We used the default options to perform a varLD analysis. The standardized varLD score was plotted using R. A Z-score of ≥ 4 of the standardized varLD score was used as a cut-off value to declare variation in LD between the two chicken populations. For intrapopulation selective sweep analysis we used iHS and SweeD, while Rsb, F_{ST} and varLD were used for interpopulation analyses.

Functional annotation of putative genes

We scanned for genomic regions within 50Kb upstream and downstream of SNPs showing significant selective sweep signal to map putative genes that have been subjected to a selection pressure. Genomic coordinates of SNPs showing strong signals of selective sweep ($P < 0.0001$ for iHS and Rsb; $F_{ST} \geq 0.2$, $Z \geq 3$ for SweeD) were mapped to Galgal4 built of the Ensembl genome browser release 74 (http://www.ensembl.org/info/website/tutorials/index.html). BioMart portal of the Ensembl was used to map genes that are closely located with the outlier SNPs. Functional annotation of putative genes was performed using DAVID (http://david.abcc.ncifcrf.gov/, Huang et al., 2009a & b). Additional information on function of putative genes was obtained from GeneCard (http://www.genecards.org/) and NCBI genome browser (http://www.ncbi.nlm.nih.gov/). We used a KEEG pathway as implemented in DAVID to identify genes involving in a pathway using the list of putative genes produced by BioMart. We used DAVID for each test statistic independently as the assumption made by each test while mapping the selective sweep region is different. For iHS and SweeD analyses we run DAVID for each population independenlty. To get a general pattern of intrapopulation selective sweep, we used the combined list of putative genes produced by iHS and SweeD analyses at population level. We also combined the lists of genes obtained from iHS and SweeD in the entire population to investigate the global trend of intrapopulation selection. Similarily, besides independent analysis of F_{ST} and Rsb, we combine the list of putative genes produced by F_{ST} and Rsb analyses. We also combine F_{ST} and Rsb gene lists with varLD to investigate the impact of variation in LD pattern on divergent selection. We used the classification strigency “high” option in functional annotation cluster analysis of DAVID and we select a minimum enrichment score of 1.3 as a cut-off value.

Results

Out of 391384 SNPs that have been found to be informative for allelic polarity and included in this analysis, in 68.2% of the cases (266785/391384), the major allele represents ancestral allele. In absence of out-group information assigning the major allele as ancestral therefore can be considered as alternative option. The Gaussian and the standardized observed distributions of iHS and Rsb plots were overlaid (Figure S7.1a & b and Figure S7.2a), indicating the standardized scores assume a normal distribution. For iHS, Rsb and SweeD analyses we report independent and combined results from the bigger (391384 SNPs) and smaller (53533 SNPs) dataset. Intrapopulation selective sweep was performed to map signature of locally-driven selection, while interpopulation selective sweep analysis was performed to map the geography of a divergent selection.

Intrapopulation selective sweep

Integrated haplotype score

From iHS analysis of Horro chickens using the bigger dataset forty nine SNPs that are located in selective sweep region were mapped to GGA1–8, 10, 15, 18, 19, 21 and 24 (Figure 7.1). Ninety eight putative genes that are closely linked with the SNPs were identified. Among the putative genes mapped, forty one have been characterized in chickens. The iHS test shows a better resolution in Horro than Jarso chickens. Interesting selective sweep peaks were observed on GGA5 & 7 in Horro chickens. The selective sweeps mapped to GGA1 & 6 in Horro chickens also show a clearly visible peak. The strong selective sweep observed on GGA7 in Horro chickens is most likely linked with the combined effect of natural and artificial selection. GGA7 harbours the structural mutation underlying the rose comb mutation (Imsland et al., 2012) and selection in favour of the derived variants of comb shapes (including the rose comb) is common in Ethiopia (e.g. Desta et al., 2013). GGA7 also harbours a gene family (homobox) that involves in morphogenesis (Nelson et al., 1996). Among the putative genes mapped in Horro chickens Gpr143 (Schnur et al., 1998) and brwd (see Bennett & Lamoreux, 2003 for a review) involve in melanogenesis. TMTC2 may involve in endoplasmic reticulum calcium homeostasis

(Sunryd et al., 2014). Moreover, SYNE3 binds with plectin which can then associate with the intermediate filament (IF) system. The connection between the nucleus and the extracellular matrix through the IF cytoskeleton may keep the nucleus in its proper position (Wilhelmsen et al., 2005). DHRS3 involves in retinol metabolism (Haeseleer et al., 1998). The most significant ten SNPs showing a recent positive selection in Horro chickens from iHS analysis of the bigger dataset are presented in Table 7.1.

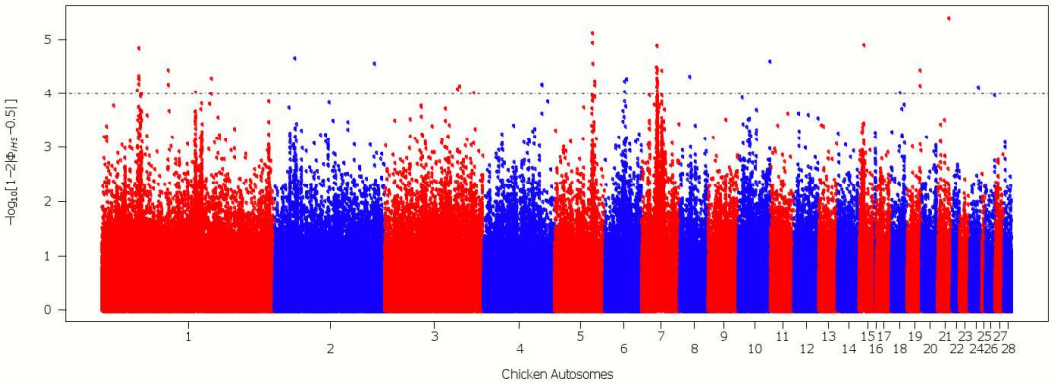


Figure 7.1 The standardized iHS plot for Horro chickens using 391384 SNPs.

Table 7.1 The most significant ten SNPs showing recent positive selection in Horro chickens based on the bigger dataset.

Chr	Pos	iHS score	-log$_{10}$ (P)	Associated genes
21	5341978	-4.61693	5.409545189	ENSGALG00000023667, DHRS3, VPS13D
5	45273889	4.4849	5.136986684	SYNE3
5	45274321	-4.39583	4.957227874	SYNE3
15	4218260	4.374029	4.91373159	TMEM132B
7	15699479	-4.36924	4.904206825	MTX2
7	15752923	4.365185	4.896143802	HOXD3, HOXD4, HOXD8, HOXD9
1	40914733	-4.34445	4.855032315	TMTC2
2	28427172	-4.24998	4.670006064	ISPD
10	19425595	4.21926	4.610646222	BLM, RASGRF1
2	135805308	4.200409	4.574411722	SAMD12

In Jarso chickens nineteen outlier SNPs were found on GGA1, 2, 6, 9–15 and 24. In Jarso chickens, thirty two putative genes that are physically closely linked with the outlier SNPs were mapped and out of these five have been characterized in chicken. Among the putative genes mapped in Jarso chickens using the bigger dataset, IDE involves in degradation of insulin (Authier et al., 1996). The closely associated genes IDE-KIF11-HHEX of GGA6 may associate with diabetics in human (Furukawa et al., 2008). CCNG1 of the GGA13 is associated with p53 that regulates the cell cycle (Seo et al., 2005). NUDCD2 is syntenic with CCNG1 and may regulate the LIS1/dynein pathway by stabilizing LIS1 with Hsp90 chaperone (Yang et al., 2010). ZBTB16 among other may interfere with glucocorticoid-induced apoptosis (Wasim et al., 2010). The most significant ten SNPs showing a recent positive selection in Jarso chickens from iHS analysis of the bigger data are presented in Table 7.2.

Most of the outlier SNPs in the two chicken populations were mapped to the intermediate-sized autosomes (Figure 7.1 & 7.2). Though, the two chicken populations have been mainly evolved under natural selection, an overlapping selective sweep regions are less frequent (Figure 7.1 & 7.2). For example, only two putative genes (UBE4A and ZBTB16) were commonly mapped in Horro and Jarso chickens on GGA24 from iHS analysis of the bigger dataset. UBE4A drives multiubiquitin chain assembly and may also involve in stress tolerance (Koegl et al., 1999) and ZBTB16 may among others involve in skeletal and male germline development (Fischer et al., 2008) and in glucocorticoid-induced apoptosis (Wasim et al., 2010).

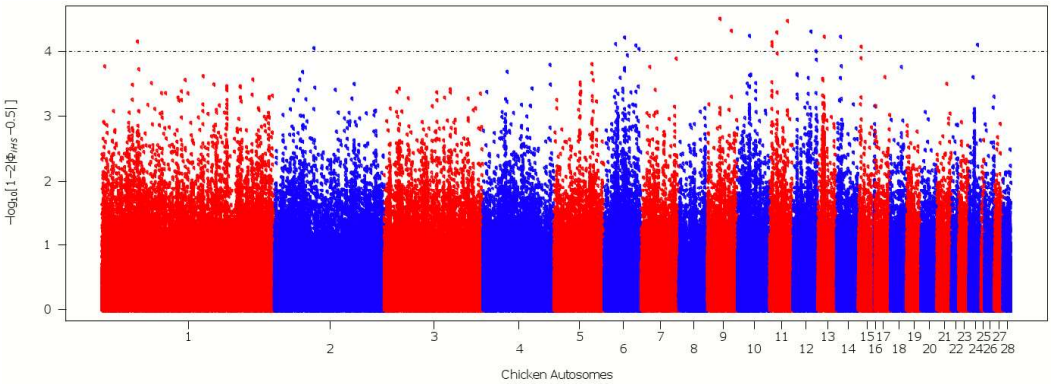


Figure 7.2 The standardized iHS plot for Jarso chickens using 391384 SNPs.

Table 7.2 The most significant ten SNPs showing recent positive selection in Jarso chickens based on the bigger dataset.

Chr	Pos	iHS score	-log$_{10}$ (P)	Associated genes
9	9584218	-4.1701	4.516454766	TRIP12
11	14399775	4.151548	4.481178034	
9	18447672	4.071342	4.330283035	NLGN1
12	14632994	4.064676	4.317861329	KBTBD8, ENSGALG00000007596
11	6298298	4.057155	4.30386958	ZNF423, ENSGALG00000028124
10	7732961	4.029356	4.252356999	UNC13C
13	6089860	4.023074	4.240758886	NUDCD2, CCNG1
14	2933484	-4.02204	4.238849658	MAD1L1
6	20298722	4.015967	4.227660365	HHEX, KIF11, IDE
1	39625678	3.980959	4.163436472	MBSP

An integrated haplotype score analysis performed on the smaller dataset (53533 SNPs) has mapped seventy two SNPs associated with genomic regions subjected to selective sweep on GGA1, 2, 4, 6 – 8 and 18 in Horro chickens. Ninety five putative genes that are closely linked with these SNPs were identified. Among the putative genes, twenty seven have been characterized in chicken. The iHS analysis performed using the smaller dataset shows a prominent peak on GGA7 in Horro chickens (Figure 7.3) clearly visible than the bigger dataset. Twenty five common genes found in Horro chickens from iHS analyses performed on the bigger and smaller dataset. The iHS analyses performed independently on the bigger and smaller dataset produced a total of one hundred sixty eight putative genes in Horro chickens. From putative genes that are located on GGA7, IHH is known to regulate tissue patterning, skeletogenesis and cellular proliferation (see Shimoyama et al., 2007 and the references therein) and it may also involve in pigment differentiation (Moshiri et al., 2004). Homobox family loci (*e.g.* HOXD4, HOXD8, HOXD9, HOXD10, HOXD11, HOXD12 and HOXD13) involve in embryonic development including anterior-posterior axis patterning (Kuraku & Meyer, 2009). LY75 is known to have immune function in chicken (Staines et al., 2013). INHA is known to involve in gonadal sex

differentiation (Ayers et al., 2013). DNPEP plays a sytemic role in the skeletal development (Nakamura et al., 2011). ATG9A is among authophagy-related genes and it may involve in maintaining homeostasis (Piekarski et al., 2014). HDAC4 inhibits chondrocyte hypertrophy (Guan et al., 2011). ATF2 involves in modification of the chromatin structure that enhances transcription (Bruhat et al., 2007). WIPF1 associates with aggressive Neuroblastoma (Angeles Rabadán et al., 2013).The most significant ten SNPs showing a recent positive selection in Horro chickens from iHS analysis of the smaller data are presented in Table 7.3.

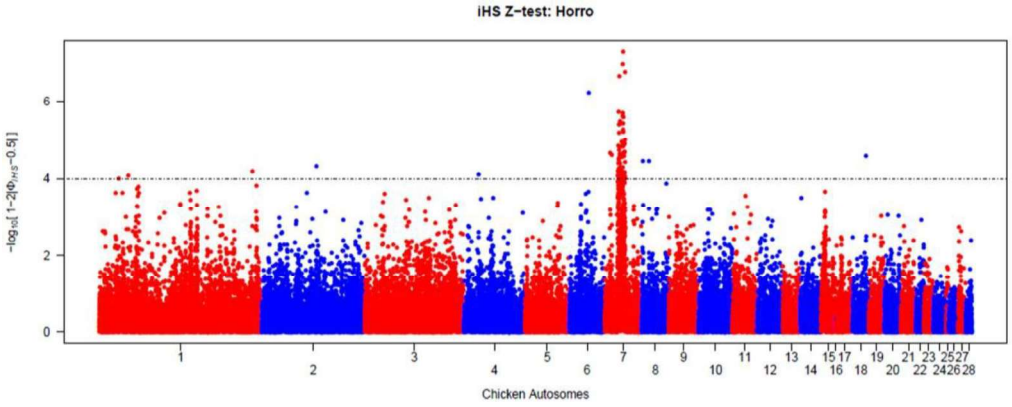


Figure 7.3 The standardized iHS plot for Horro chickens using 53533 SNPs.

Table 7.3 The most significant ten SNPs showing recent positive selection in Horro chickens based on the smaller dataset.

Chr	Pos	iHS score	-log$_{10}$ (P)	Associated genes
7	19991308	5.443488	7.281936587	
7	19779756	5.314097	6.969855189	GRB14
7	21583378	5.220369	6.748199091	SLC4A3, STK11IP, INHA, OBSL1, ENSGALG00000011233, TMEM198, CHPF, ASIC4
7	16311522	5.17897	6.651472284	CHN1, CHRNA1, WIPF1
6	21637946	4.989241	6.217405545	CPN1, DNMBP, ABCC2, CUTC, COX15, ATP6V0E2, ENTPD7, SLC25A28, GOT1, ENSGALG00000020753,
7	15711977	4.759045	5.711055654	MTX2
7	19801640	-4.75002	5.691666197	
7	19894616	-4.71645	5.619788195	
7	20286636	4.707024	5.599704148	KCNH7
7	17038749	4.649004	5.476849601	

Despite a fewer number of SNPs, the smaller dataset has better resolution power even in Jarso chickens, for example, the peaks on GGA6, 7, 13 and 19 (Figure 7.4). In Jarso chickens seventeen SNPs (GGA1–4, 6–8 & 19) were found to associate with signature of selection and twenty one putative genes that are closely linked to these SNPs were mapped. Out of these, six genes have been characterized in chicken. In Jarso chickens the bigger and smaller dataset iHS analyses didn't produce common putative genes. The two iHS analyses have produced fifty three putative genes in Jarso chickens. Among these BCAS3 of the GGA19 is suggested to involve in embryogenesis and tumour angiogenesis in human (Siva et al., 2007). DDC of the GGA2 induces apoptosis (Mehlen et al., 1998) and CPN1 encodes for chaperonin subunits that involve in folding of a subset of newly-synthesized proteins (Yamabe et al., 2010). C10orf11 involves in differentiation of melanocytes (Grønskov et al., 2013) and DNMBP binds the dynamin and actin regulatory proteins (Kuwano et al.,

2006). ABCC2 is involved in trafficking of organic anions (Kast et al., 2002) and CUTC is involved in the homeostasis of copper including uptake, intracellular storage and delivery, and its efflux (Gupta et al., 1995 and the references therein). COX15 is involved in the biosynthesis of heme A (Antonicka et al., 2003) and Atp6v0e2 involves in proton pump (Cort et al., 2010). ENTPD7 is an endo-apyrase (Shi et al., 2001) and SLC25A28 is mitochondrial iron uptake molecule (Galy et al., 2010). GOT1 may regulate the production of the soluble and mitochondrial forms of glutamate oxalate transaminase (DeLorenzo & Ruddle, 1970). The most significant ten SNPs showing a recent positive selection in Jarso chickens from iHS analysis of the smaller data are presented in Table 7.4.

The iHS analyses based on the smaller dataset produced sixteen common putative genes in Horro and Jarso chickens (see the list below), which is higher than two putative genes commonly mapped using the bigger dataset. This shows that despite its porosity, the smaller dataset found to be relatively more informative despite fewer markers included in the analyses. A combined list of putative genes obtained from iHS analyses performed on the bigger and smaller datasets independently produced eighteen common putative genes (UBE4A, ZBTB16, CPN1, DNMBP, ABCC2, CUTC, COX15, SLC25A28, GOT1, ENSGALG00000009344, IKZF1, DDC, ENSGALG00000013233, ENSGALG00000020753, ATP6V0E2, ENSGALG00000021859, ENTPD7 and FIGNL1) in the two chicken populations. A functional annotation cluster analysis shows that these common putative genes involve in ion binding regardless of theit low enrichment score (0.36). Among the eighteen common putative genes, DDC and GOT1 involve in two pathways – phenylalanine and tyrosine metabolism.

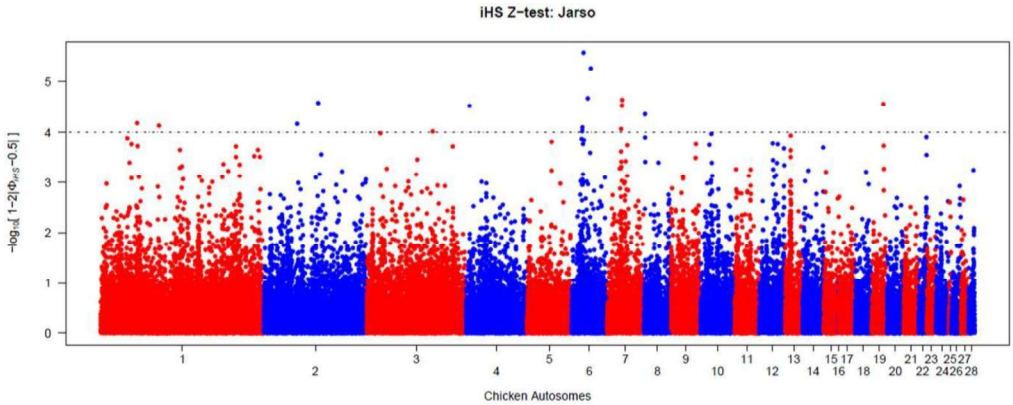


Figure 7.4 The standardized iHS plot for Jarso chickens using 53533 SNPs.

Table 7.4 The most significant ten SNPs showing recent positive selection in Jarso chickens based on the smaller dataset.

Chr	Pos	iHS score	-log$_{10}$ (P)	Associated genes
6	14231336	4.694472	5.573008244	
6	21637946	4.54816	5.266666463	CPN1, DNMBP, ABCC2, CUTC, COX15, ATP6V0E2, ENTPD7, SLC25A28, GOT1, ENSGALG00000020753
6	18439134	4.25085	4.671700793	
7	17116657	-4.23381	4.638719791	ENSGALG00000021859, ENSGALG00000013233
2	80712680	-4.2054	4.583996019	IKZF1, FIGNL1, DDC
19	7716946	-4.19449	4.563060069	BCAS3
7	17027628	4.178452	4.532391788	
7	16960817	-4.17389	4.523678375	ENSGALG00000009344
4	5777580	4.172564	4.521157119	
8	1890515	4.083751	4.353453912	

SweeD analysis

Here in the main chapter, we report a SweeD result from the bigger dataset only because unlike iHS test, SweeD analyses performed using the smaller dataset produce relatively weak selective sweep signals in both populations (Figure S7.4a & b). Despite this, SweeD analyses from the two datasets show a similar trend in the two chicken populations. Unlike iHS test that detects alleles that have been selected to moderate level, SweeD detects loci that have been strongly selected and either fixed or approach to fixation.

One hundred forty one genomic regions with a Z-score value of > 3 (representing ~ 0.27% of the upper tail region) were found in Horro chickens. Thirty nine putative genes were mapped to GGA1, 2, 5 & 20 (Figure 7.5). Despite prominent peaks found on GGA3 & 4, we couldn't detect any gene within ~ 0.4Mb regions bracketing

the selected genomic region in Horro chickens. Among the identified putative loci, Snai2 is one of the lineage specific melanocyte transcriptional factor (Wang et al., 2013) and BCL2 may inhibit apoptosis of the melanocytes (Müller-Röver et al., 1999). TP53 activates tyrosinase and tyrosinase-related protein-1 (TRP-1) biosynthesis (Nylander, et al., 2000). GRB2 activates the Ras guanine nucleotide-binding protein by tyrosine kinases (Rozakis-Adcock et al., 1992) and CDK14 is involved in eukaryotic cell cycle (Kaldis & Pagano, 2009). Hoxa11 and Hoxa13 may regulate muscle patterning in the limb buds (Yamamoto et al., 1998). A homeobox-containing Evx1 may activate the cytotactin promoter in chicken via growth-factor signal transduction pathway (Jones et al., 1992) and ROMO1 may play a key role in innate immune defence (Kasthuri et al., 2013). No common putative genes were mapped by iHS and SweeD analyses in Horro chickens, due to different assumptions made by the two tests, this is not unexpected.

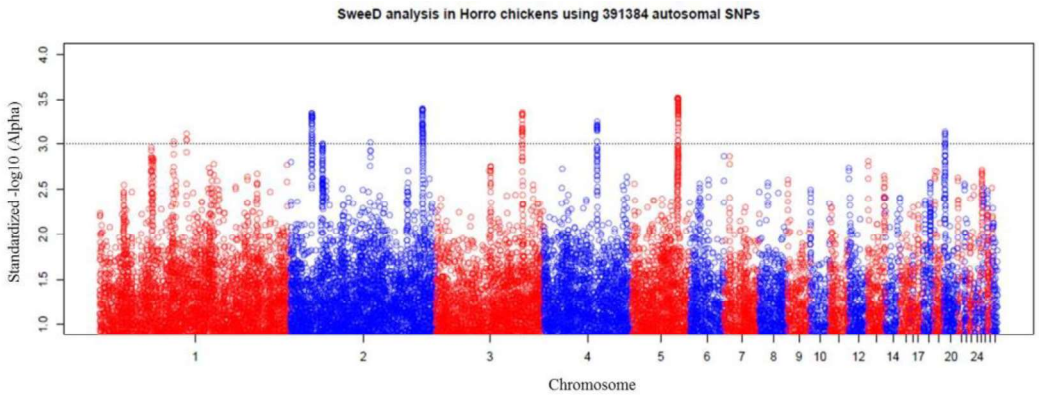


Figure 7.5 The standardized SweeD plot for Horro chickens using 391384 SNPs.

One hundred nineteen regions showing a strong selection were mapped in Jarso chickens using a SweeD analysis. Thirty seven heavily selected putative genes were mapped to GGA1–3, 5, 7, 20 & 27 (Figure 7.6). Among these twenty one genes are coding for uncharacterized proteins. Among the functionally characterized genes MYC may involve in the localization of melanosomes (Yatsu et al., 2013). Gata6 has essential role in extraembryonic development (Koutsourakis et al., 1999) and SH3BP4 is a negative regulator of the Rag GTPase complex and amino acid-dependent mTORC1 signalling (Kim et al., 2012). No common genes were also found between iHS and SweeD analyses in Jarso chickens.

SweeD analyses using the bigger dataset in Horro and Jarso chickens has mapped eleven common putative genes (NFS1, CPNE1, ENSGALG00000001694, ENSGALG00000001710, ENSGALG00000001754, ENSGALG00000011127, ENSGALG00000026151, ENSGALG00000026222, ENSGALG00000027834, ENSGALG00000028278 and gga-mir-1467-1), however, most of these genes are coding for uncharacterized proteins, which has limited us from doing further functional annotation. In the functionally characterized genes, NFS1 serves as a supplier of sulphur to MOCS3 – a protein involving in the biosynthesis of molybdenum cofactor (Marelja et al., 2008) and CPNE1 is thought to be involved in membrane trafficking (Creutz et al., 1998). In both chicken populations a heavily selected region was identified on GGA5 from SweeD analysis; however the closely linked putative genes are coding for uncharacterized proteins.

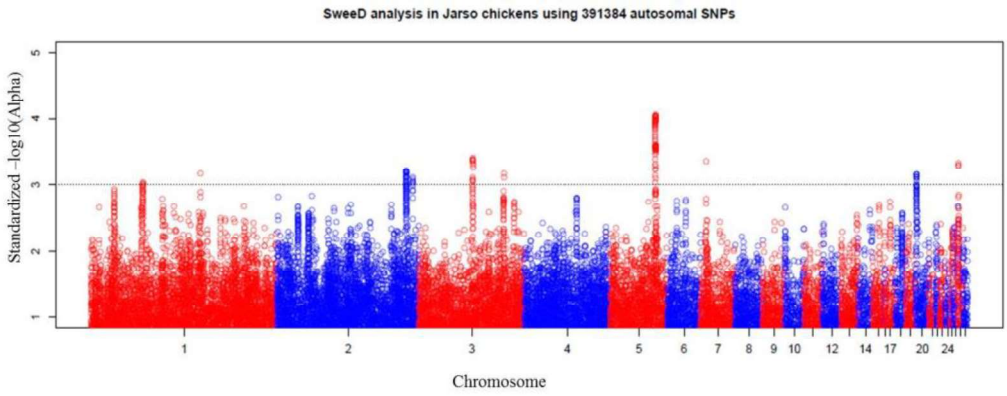


Figure 7.6 The standardized SweeD plot for Jarso chickens using 391384 SNPs.

Interpopulation selective sweep

We map thirty three outlier SNPs associated with a divergent selection ($P < 0.0001$) between Horro and Jarso chickens to nine autosomes (GGA1, 2, 6, 10, 11, 15, 18, 17 and 22) from Rsb analysis performed using the bigger dataset. Thirty two putative genes are located close to the SNPs associating with a divergent selective sweep (Figure 7.7). Among the putative genes, fifteen have been functionally characterized in chickens. This shows a considerable level of divergent selection between Horro and Jarso chickens. Among the putative genes mapped, brwd involves in pigmentation (see Bennett & Lamoreux, 2003 for a review) and BCL2 may inhibit apoptosis of the melanocytes (Müller-Röver et al., 1999). GRB2 activates the Ras

guanine nucleotide-binding protein by tyrosine kinases (Rozakis-Adcock et al., 1992). Sp4 can bind to the GC box or GT motifs and serves as transcriptional activator (Parakati & DiMario, 2002). ALDH6 may involve in retinoic acid synthesis, a molecule used in growth and development (Duester, 2001). ASB7 interacts with Cul5–Rbx2 and has ubiquitin ligase activity (Kohroki et al., 2005). LRKK1 regulates the endosomal trafficking of epidermal growth factor receptor (Hanafusa et al., 2011). GNAO1 may interact with tyrosine, an amino acid involving in melanogenesis (Wang et al., 2012). PIWIL1 interacts with RNA pathway and involves in spermatogenesis (Gu et al., 2010). The most significant ten SNPs showing a divergent selection between Horro and Jarso chickens from Rsb analysis of the bigger dataset are presented in Table 7.5.

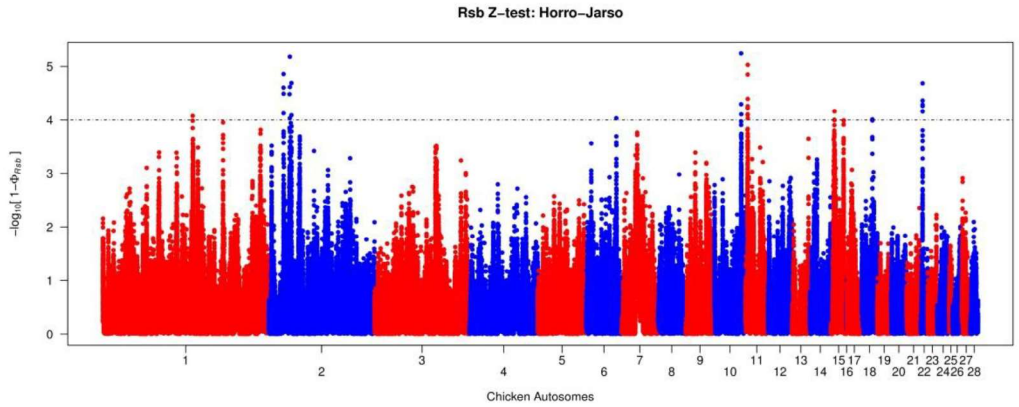


Figure 7.7 Rsb analysis between Horro and Jarso chickens using 391384 SNPs.

Table 7.5 The most significant ten SNPs showing divergent positive selection between Horro and Jarso chickens based on the bigger dataset.

Chr	Pos	Rsb score	-log 10 (P)	Associated genes
10	16916495	4.387239	5.241093477	ENSGALG00000021874, ALDH6
2	30574040	4.356	5.178942696	SP4
11	2205496	4.279781	5.029008682	AMFR, GNAO1
2	21559333	4.191724	4.858797127	ENSGALG00000009006, STEAP2, CFAP69, GTPBP10
11	2125533	4.186872	4.84951274	NUP93, MT4, MT1, BBS2, OGFOD1, NUDT21, AMFR
2	33358439	4.100749	4.686328573	HIBADH
22	1983721	4.097732	4.680666703	KCNU1
2	30625024	4.060897	4.611858783	ENSGALG00000010906
2	21499919	4.054175	4.599361441	ENSGALG00000009006, STEAP2
2	21564728	3.993226	4.486912558	STEAP2, CFAP69, GTPBP10

Fifteen outlier SNPs were detected from Rsb analysis of the smaller dataset (GGA1–3, 7, 11 and 15). Additional selective sweep peaks that show a divergent selection were mapped to GGA3, 7 and 17 using the smaller dataset (Figure 7.8). Thirty putative genes were mapped in the divergently selected regions among which twelve have been characterized in chicken. Among the putative loci mapped Rab1 may involve in melanosomes trafficking (Hume et al., 2001) and Oca2 involves in pigmentation (Sitaram et al., 2009). Mutation in SOX5 gene underlies the genetic control of pea comb (Wright et al., 2009) and is known to involve in chondrocyte differentiation and cartilage formation (Lefebvre et al., 2001). Ufl1 is an E3 ligase that regulates conjugation of C20orf116 with Ufm1 (Tatsumi et al., 2010). WDR59 involves in mTORC1 signalling pathway, which regulates protein synthesis (Bar-Peled et al., 2013). UlK1 as mTORC1 substrate have been suggested to play a major role in initiation of autophagosome formation (Alers et al., 2011). RTN4R regulates axonal growth and involve in regeneration of injured axon (Hsu et al., 2007). The most significant ten SNPs showing a divergent selection between Horro and Jarso chickens from Rsb analysis of the smaller data are presented in Table 7.6.

The combined putative gene list of the two Rsb analyses produced sixty putative genes. Among these, two genes (PIWIL1 and FZD10) were commonly mapped by Rsb analyses independently performed on the bigger and the smaller dataset. PIWIL1 is known to regulate spermatogenesis in chicken (Chen et al., 2013) and FZD10 may regulate the development of limb and central nervous system during embryogenesis (Kawakami et al., 2000).

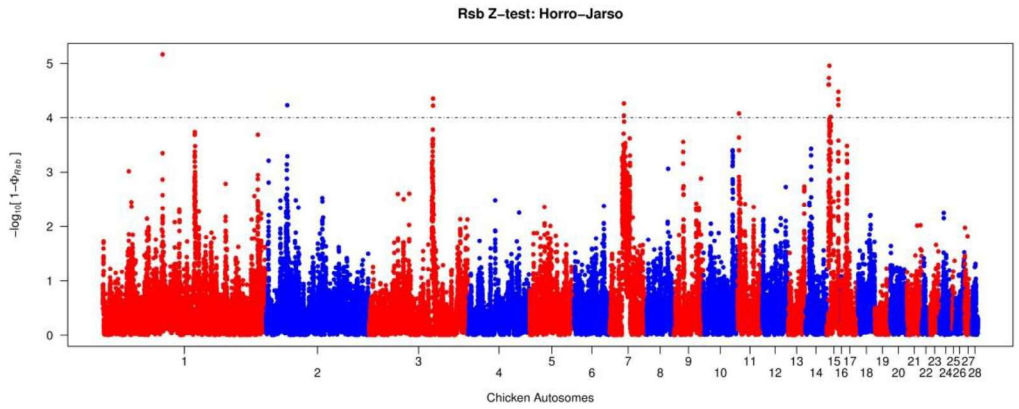


Figure 7.8 Rsb analysis between Horro and Jarso chickens using 53533 SNPs.

Table 7.6 The most significant ten SNPs showing divergent positive selection between Horro and Jarso chickens based on the smaller dataset.

Chr	Pos	Rsb score	-log$_{10}$ (P)	Associated genes
1	66043320	4.348409	5.163902221	SOX5
15	3388580	4.242694	4.956927781	PIWIL1
15	2625837	4.12438	4.730797023	ENSGALG00000002336, PUS1, ULK1
15	2228128	4.058824	4.608003735	
15	2684336	4.058692	4.607758188	ULK1
15	10417568	3.987953	4.477257422	RTN4R
3	72249742	3.919637	4.353192969	FHL5, UFL1
15	10402441	3.913596	4.342314659	RTN4R
7	16295914	3.870069	4.264383262	CHN1
15	10409842	3.854553	4.236791484	RTN4R

Genetic differentiation

A sliding window F_{ST} analysis identified thirty five outlier windows (10 SNPs each) showing F_{ST} value of ≥ 0.20 with an associated z-score of ≥ 13.9 on GGA1, 3, 4, 6, 7, 10, 12 and 27 (Figure 7.9). Forty two putative genes were mapped in genomic regions showing a considerable level of genetic divergence. Among these, twelve genes have been already characterized in chicken. Particularly, the peaks on GGA6 & 7 are clearly evident, which supports Rsb and iHS analyses. For example, among the putative genes that are located close to genomic regions showing a considerable genetic divergence, MYPN has been suggested as a genetic marker for meat quality traits in cattle (Jiao et al., 2010) and in commercial pigs (Zhai et al., 2010; Braglia et al., 2013) and it also involve in actin organization and cell motility (Goicoechea et al., 2008). No duplicates were found between lists of putative genes produced by Rsb and F_{ST} analyses.

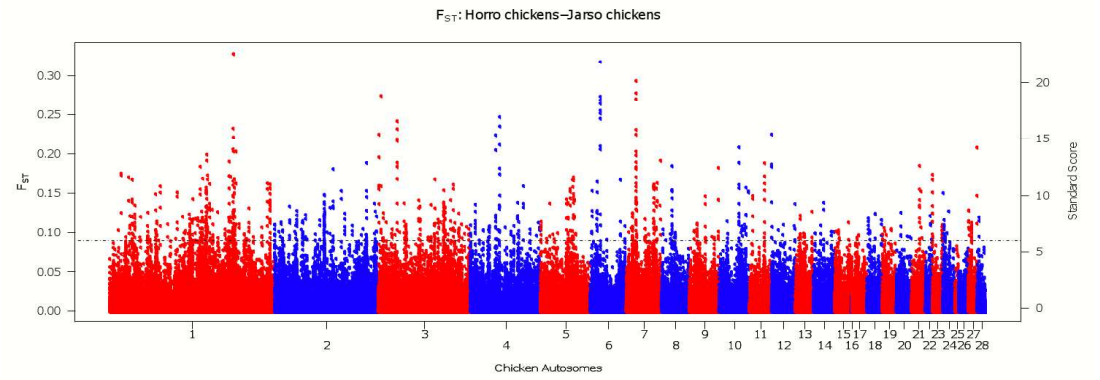


Figure 7.9 Raw and standardized ten SNPs sliding window FST result for Horro and Jarso chickens.

Variation in linkage disequilibrium

A standardized varLD score plot is presented in Figure 7.10. A clearly visible variation in LD pattern was observed between Horro and Jarso chickens for most of the autosomes. However, a high variation was observed on GGA2, 7 & 13. The varLD plots made for each autosome are displayed in Figure S7.5a–e. One thousand six hundred sixty genomic regions having a standardized varLD score of ≥ 4 were found. Two hundred twelve putative genes were mapped in regions showing a considerable variation in LD between the two chicken populations. Thirty two of

these genes have been already characterized in chicken. CDH18 is among cadherin adhesion molecules regulating cellular interactions during embryogenesis and morphogenesis and later in life it is useful to maintain the integrity of tissues (Chalmers et al., 1999). FOXI1 is among transcriptional regulators of the inner ear and pharyngeal arch derivatives development (Khatri & Groves, 2013). DOCK2 is a Rac guanine exchange factor catalysing the GDP-GTP exchange of Rac downstream chemokine receptors (Gollmer et al., 2009). FAM196B is reported to express in developing ovary of mouse and may play a role in ovarian differentiation (Chen et al., 2012). Clint1 is reported to involve in epidermal development and inflammation in zebrafish (Dodd et al., 2009). MCM3AP acetylates the replication protein MCM3 (Takei et al., 2001).

Nine functional annotation clusters were found from gene list produced from regions showing variation in LD, however, none of them were crossed a cut-off value set for enrichment score (1.3). We merged the gene list found from Rsb and varLD analyses to identify genes that have been mapped by the former due to variation LD. Four putative genes (KCNU1, NUDT21, AMFR and GNAO1) were mapped both in Rsb and varLD analyses. These genes represent both divergently selected loci and loci that locate in genomic regions showing variation in LD between the two chicken populations. The calcium insensitive but pH sensitive KCNU1 involves in potassium channel (Beisel et al., 2007). In the chicken both NUDT21 and AMFR can be referred as AMFR, which serves as autocrine motility factor receptor (Darmon & Lutz, 2012). However, no common putative genes were found from gene list produced from F_{ST} and varLD analyses.

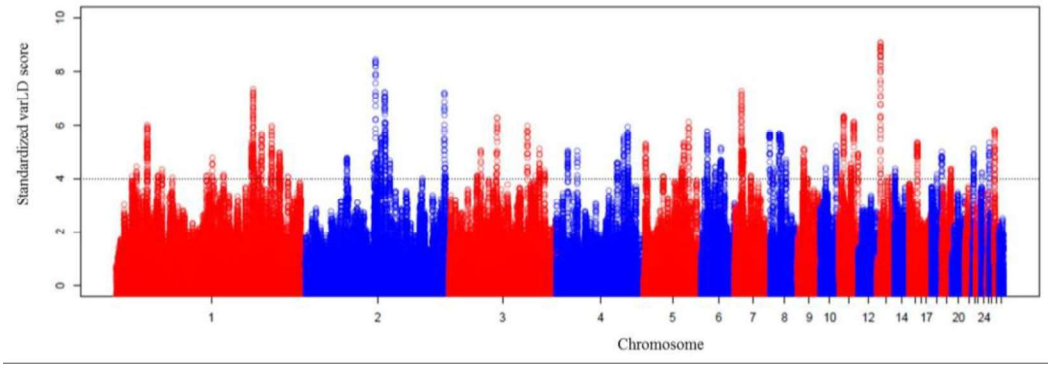


Figure 7.10 The standardized varLD score of Horro and Jarso chickens.

Functional annotation

Functional annotation was done for all the statistical tests performed to map genomic regions showing a selective sweep at intrapopulaton and interpopulation level. This report includes results of functional annotation cluster and pathway analyses performed using DAVID.

Integrated haplotype score

Using a gene list obtained from iHS test in Horro chicken DAVID identified twenty functional annotation clusters of which six have enrichment score ranging from 1.65 to 2.21. The identified clusters among others involve in morphogenesis (homobox family genes), transcription regulation, lumen, immune system development and cell activation. However, no genes from the list were found to involve in a pathway.

Using a gene list obtained from iHS test in Jarso chicken, DAVID identified three functional annotation clusters however none of them has enrichment score of 1.3. The clusters among other involve in nucleotide binding, ion binding and phosphorylation. Two genes (DDC and GOT1) from the list involve in phenylalanine metabolism pathway. This pathway produces the amino acid tyrosine that involves in the synthesis of melanin.

SweeD

The gene list produced by SweeD analysis in Horro chickens yielded two functional annotation clusters with enrichment score of 1.49 and 1.95. The clusters among others involve in anterior/posterior pattern formation, homobox, skeletal system development, DNA binding, embryonic morphogenesis and transcription regulation. However, no genes from the list were found to involve in a pathway.

The gene list obtained from SweeD analysis in Jarso chickens produced one functional annotation cluster with enrichment score of 0.57. This cluster involves in metal ion binding, cation binding and ion binding. DAVID fails to give interesting result in Jarso, because out of the thirty seven genes mapped by SweeD twenty one

are coding for uncharacterized protein. However, the heavily selected but uncharacterized loci would be good candidates for genetic adaptation studies at locus level. No pathway was found from a gene list produced by SweeD analysis in Jarso chickens.

The combined list of integrated haplotype score and SweeD

We found twenty four clusters from the combined gene list of iHS and SweeD analyses in Horro chickens. Among the functional annotation clusters identified, eight have enrichment score ranging from 1.53 to 4.11. The clusters among others involve in homobox, transcription regulation, skeletal system development, organelle lumen, DNA binding, limb development etc. No pathway was found from the combined list of genes from iHS and SweeD analyses in Horro chickens.

We found five clusters from the combined gene list of iHS and SweeD analyses in Jarso chickens, none of the functional annotation clusters identified however show enrichment score of 1.3. The clusters among others involve in phosphorylation, metal ion binding, organelle lumen, nucleotide binding, etc. From the list two genes (DDC and GOT1) involve in phenylalanine metabolism pathway. Phenylalanine is a precursor to tyrosine and oxidation of tyrosine produces melanin. This implies that DDC and GOT1 may indirectly involve in melanogenesis – a pathway detected by Rsb analysis to show a divergent selection between the two chicken populations.

We found twenty five clusters from the combined gene list of iHS and SweeD analyses at the entire population level, i.e. putative genes that are subjected to either ongoing or strong selective sweep in the two chicken populations. Among the twenty five functional annotation clusters found, five have enrichment score of 1.47 to 3.51. The clusters among others involve in homobox, organelle lumen, limb development, transcription regulation, immune system development, cell activation etc. However, no pathway was found from the combined list of genes from iHS and SweeD analysis at the entire population level.

Rsb and F_{ST}

A functional annotation cluster analysis performed using putative genes obtained from Rsb analysis produced five functional annotation clusters and of which one has enrichment score of 1.57. The annotated clusters among others involve in sexual reproduction, cation binding and regulation of cell death. One pathway was found from list of genes obtained from Rsb analysis – melanogenesis pathway. Three of the putative genes from Rsb analysis (FZD10, GNAO1 and PRKCA) involve in melanogenesis pathway. However, a single functional annotation cluster with low enrichment score of 0.23 and involving in metal ion binding, cation binding and ion binding was found from the gene list produced by F_{ST} analysis. No pathway was found from putative genes list produced by F_{ST} analysis. The combined gene list from Rsb and F_{ST} analyses yielded ten functional annotation clusters among others involving in sexual reproduction, ion binding and regulation of cell death, however only a single cluster shows enrichment score of 1.46. The combined gene list from Rsb and F_{ST} analyses also produced melanogenesis pathway as Rsb gene list does and four genes (FZD10, FZD7, GNAO1 and PRKCA) found to involve in this pathway.

varLD

Using a gene list produced by a varLD analysis we found sixteen functional annotation clusters however none of these have enrichment score of 1.3. The clusters among others involve in nucleotide binding, immunoglobulin, nucleotide regulatory activity, protein localization and cation binding. Four pathways (VEGF signalling pathway, MAPK signalling pathway, Arachidonic acid metabolism and GnRH signalling pathway) were found from gene lists produced by varLD analysis. Four genes from the list (MAPK13, Mapk14, pla2g4a and PLA2G10) involve in GnRH signalling pathway. Three genes (pla2g4a, PLA2G10 and PTGS2) from the list involve in Arachidonic acid metabolism pathway. Seven genes (RAPGEF2, BDNF, MAPK13, MAPK14, NLK, pla2g4a and PLA2G10) from the list involve in MAPK signalling pathway and five genes (MAPK13, MAPK14, pla2g4a, PLA2G10 and PTGS2) from the list involve in VEGF signalling pathway.

Rsb, F_{ST} and varLD

A combined gene list produced by Rsb, F_{ST} and varLD analyses yielded twenty eight functional annotation clusters however none of these have enrichment score of 1.3. Genes that involve in three pathways (VEGF signalling pathway, MAPK signalling pathway and Arachidonic acid metabolism) were found from putative genes mapped from interpopulation selective sweep analyses and variation in LD. Six genes (MAPK13, MAPK14, pla2g4a, PLA2G10, PTGS2 and PRKCA) from the list involve in VEGF signalling pathway. Eight genes (RAPGEF2, BDNF, MAPK13, MAPK14, NLK, pla2g4a, PLA2G10 and PRKCA) from the list involve in MAPK signalling pathway. In line with Rsb and F_{ST} combined gene list, three genes (pla2g4a, PLA2G10 and PTGS2) found to involve in Arachidonic acid metabolism pathway.

Discussion

Signature of selection may take two forms: hard or soft sweeps. Loci that contribute a higher proportion of underlying genetic variation could be subjected to a stronger selection than loci that account for a low proportion. However, the strength of a selective sweep (positive directional selection) depends on conservation status of the loci, age of the selective sweep, inbreeding, genetic admixture, local recombination rate and demographic history of the population. The signature of selection may be extensive in a population that has been subjected to intense artificial selection while it may be less evident in panmictic populations that have been mainly evolved under natural (balancing) selection and displaying a high standing genetic variation. Artificial selection for traits of special significance, natural selection for local adaptation and genetic drift may lead to a strong selective sweep. Artificially induced selective sweeps, however, may loss quickly following a random mating. Like artificial selection, natural selection is not a random event as it is dictated by the type and amount of resource required for local adaptation.

Stochastic forces like genetic drift however may have little impact on village chickens that have been commonly experiencing a random mating (Desta et al., 2013) and characterized by a high standing genetic variation (Lyimo et al., 2014).

Signature of selection in natural populations may not as extensive as populations that have been intensively selected for breed formation and show traits. Large sample size reduces the proportion of false positive selective sweeps that could otherwise occur due to local demographic effects. Increasing sample size may also counteract with the impact of ascertainment bias arising from genomic information included while developing a SNP chip. Genetic admixture and standing gene flow blur with selective sweep signal and may reduce the power of selective sweep mapping. A fine-scale genetic substructure that arises from local spatial barriers may also reduce the power of selective sweep mapping.

Loci that are invariably conserved across species may regulate vital biological functions. Most of the putative genes mapped by our selective sweep analyses are conserved in other species (data not shown). Nevertheless, although most of the loci are conserved among species, in most instances they are divergently selected in Horro and Jarso chickens. This may arise due to a considerable variation in biotic and abiotic factors between the two study sites besides a limited gene flow between the two populations attributable to their distant geographical location and absence of trading network and variation in the landscape of genetic drift and demographic history. Environmental variables like predator pressure, disease challenge, climate, feed etc. may also to some extent vary between the two sites which may make natural selection to act divergently. If there is a variation in a locally driven physiological demand, different ecologically important traits may be subjected to (divergent) selection in the two populations. Moreover, the demographic structure of the community may impose different pressure while selecting for traits of visual appeal. The putative loci mapped by the selective sweep analyses are known to have an adaptive value and a number of noble loci were also identified. The two chicken populations also may have a considerably contrasting demographic history, which may have resulted in different selective sweep landscape. Age of an allele determines the extent of selective sweep; older alleles may have short tract of selective sweep than younger alleles due to the impact of historical recombination. Rapid changes in environment and climate can also lead to a paradigm shift in allelic frequency in the selective sweep region. Loci may evolve slowly in stable than dynamic environment. A locus can be quickly fixed by selection than attributable to genetic drift. Similarly, neutral varition may take longer to fixation than adaptive

variation. Evolution of a genetic variant has spatial and temporal dimension. A genetic variant that is advantageous under one temporal and spatial setting may be detrimental under another circumstance depending on the type and amount of resources required for local adaptation.

Selective sweep can be break down by historical recombination; therefore extensive tract of positive selection may be partly the consequence of a recent adaptive process and/or anthropogenic effect. However, in recombination coldspot regions extended homozygote haplotypes may be found, which could be identified by mapping the geography of recombination rate across a genome. Nevertheless pattern of recombination coldspot may vary among populations and even among individuals within a population, for example, as it has been reported in human (Broman et al., 1998). Alleles that have been favored by adaptive evolution become more frequent, however, a selective sweep arouse from a standing genetic variation may leave weak signals (Barrett & Schluter, 2008). Beneficial mutations rapidly increase in frequency with associated loss of a local variation – a deviation in site frequency spectrum, which then creates an extended LD block (Qanbari et al., 2014).

Linkage disequlbrium is expected to be less intense in populations that have been experiencing a long history of breeding (random mating) and a limited impact of anthropogenic effect. A weak LD improves the resolution power of mapping studies and helps to locate fine-scale selective sweeps. However, a large number of markers are required to map selective sweep in panmictic populations with a fine-scale LD block. The confounding impact of anthropogenic effects with signature of selection may be less evident in panmictic populations. However, farmers prference for visual traits as commonly practiced in Ethiopia for a comb shape (Desta et al., 2013) may confound with true signals of natural selection. For example, the strong signal of selective sweep detected on GGA7 in Horro chickens likely represents the impact of both artificial and natural selections. Farmers selection in favour of a rose comb has left its own signature in this region. Moreover, Horro chickens are averagely heavier than Jarso and a number of loci that regulate growth and devolopment processes are located on GGA7 (see RESULT section). Besides an extensive discussion made for some of the GGA7 genes that are under an incomplete selective sweep somewhere in the RESULTS section, MTX2 (GGA7) involves in physical restructuring of an

embryo (epiboly) in line with gastrulation and it also dictates morphogenetic movements of endoderm and axial mesoderm (Wilkins et al., 2008).

Although village chickens have been evolved under uncontrolled breeding, they may have experienced a genetic bottleneck in the course of their breeding history, which may have also left a signature of selective sweep. Due to village chickens' long breeding history, the selective sweeps detected in the two chicken populations may be largely of an ancient origin. Selection in natural populations involves adaptation traits and these traits are largely impacted by environmental variations, hence a diverged selective sweep is not uncommon even among panmictic populations. For instance, Horro and Jarso chickens have been divergently selected among others for some of the loci involving in melanogenesis pathway. A divergent selection may have also partly made Horro chickens to exhibit vivid plumage than Jarso chickens (personal observation). A divergnet selection can be occured among populations due to imbalance between existing resources and the type and amount of resources required for local adaptation.

Selective sweep reduces effective population size and genetic variation in the selected regions of a genome. Selective sweep represents functional genomic regions that are targeted by selection. Most of the selective sweeps may be found in genic regions because commonly genes are the targets of selection, which may result in a long tract of LD in the selected genic region. Most of the footprints of selection detected by selective sweep studies may be of an adaptive origin. Ecological cline may cause a selective sweep to appear in different regions of a genome based on the demand of locally-driven biological processes. Locally-driven adaptive needs may require a specific combination of resources, which lead to a divergent selection among populations living in different environments. However, for adaptive selection that involves common vital functions, populations may have similar genomic regions that have been subjected to (could be of different extent) selective sweep.

A selective sweep mapping result is commonly varied depending on the type of a statistical test used (Utsunomiya et al., 2013). For example, F_{ST} and SweeD test provide a strong signal of selection for loci that approached to fixation. Specifically, F_{ST} analysis comfortably identifies loci that are under a considerable impact of a

divergent selection between populations and may be largely affected by the demographic history of the populations and variation in natural selection pressure. Fixation of alleles is presumably a long-term process; therefore, selective sweeps identified using these two tests likely are of an ancient origin. Rsb is a powerful test to detect loci that have been heavily selected in one population but not in the other. However, iHS maps those loci that have been selected at intermediate frequency and under incomplete selective sweep (an ongoing selection – the selected loci is still segregating). Loci that have been mapped by iHS test could be of a recent origin or may have been derived from a standing genetic variation. Therefore, although iHS and SweeD both are used to map intrapopulation selective sweep, they may not identify a similar genomic region. Unlike Qanbari et al. (2014), small proportion of overlapping regions were detected by iHS and SweeD analyses in the two chicken populations. Intrapopulation analysis of a selective sweep between two alleles of a locus as in iHS test losses power when one allele is heavily selected and completely fails to map a selective sweep when one allele is fixed (Tang et al., 2007). Haplotype based tests like iHS and Rsb are less sensitive to demographic effects however they are substantially impacted by variation in a recombination rate (Quintana-Murci & Clark, 2013). All these tests however, invariably detect genes that have been involved in a recent local adaptation (Qanbari et al., 2014).

Selective sweep may also arise due to variation in local recombination rate; however, this is proved to have less impact in the chicken population studied. Most of the putative loci that have been subjected to a divergent selection and have been mapped by Rsb and F_{ST} analyses do not overlap with loci that show variability in LD pattern between the two chicken populations. However, on the other side, variation in LD pattern could arise due to a divergent selection, demographic history, variation in recombination rate, a fine-scale genetic structure, genetic outliers, length of breeding history, variation in the impact of evolutionary forces and due to rare and population-specific haplotypes (e.g. Teo et al., 2009).

In geogrphically structured populations experiencing little or no gene flow, locally-driven selective sweeps may occur (Voight et al., 2006). Mapping selective sweep peculiar to one population but not the other is not uncommon, however, some selective sweeps may be shared among populations (Voight et al., 2006). Population

specific selective sweeps could have been partly occurred due to reproductive isolation of the two chicken populations attributable to their distant geographical location (Desta et al., 2013).

Conclusion

Selective sweep analyses performed at intrapopulation and interpopulation level identify genomic regions that have been under a recent positive selection. A considerable divergent selection observed in the two chicken populations reflects the divergent act of natural selection in response to a locally-driven adaptive variation. Our analysis mapped a large number of putative loci that have been subjected to selective sweep in outbred village chicken populations, which substantiates the importance of village chickens in selective sweep mapping.

References

ALERS, S., LÖFFLER, A. S., PAASCH, F., DIETERLE, A. M., KEPPELER, H., LAUBER, K., ... & STORK, B. Atg13 and FIP200 act independently of Ulk1 and Ulk2 in autophagy induction. Autophagy, 2011, 7(12), 1423-1433.

ANGELES RABADÁN, M., USIETO, S., LAVARINO, C., & MARTÍ, E. Identification of a putative transcriptome signature common to neuroblastoma and neural crest cells. Developmental Neurobiology, 2013, 73(11), 815-827.

ANTONICKA, H., MATTMAN, A., CARLSON, C. G., GLERUM, D. M., HOFFBUHR, K. C., LEARY, S. C., ... & SHOUBRIDGE, E. A. Mutations in COX15 produce a defect in the mitochondrial heme biosynthetic pathway, causing early-onset fatal hypertrophic cardiomyopathy. The American Journal of Human Genetics, 2003, 72(1), 101-114.

AULCHENKO, Y. S., RIPKE, S., ISAACS, A. & VAN DUIJN, C. M. GenABEL: an R package for genome-wide association analysis. Bioinformatics 2007, 23(10), 1294-1296.

AUTHIER, F., POSNER, B. I., & BERGERON, J. J. Insulin-degrading enzyme. Clinical and investigative medicine. Medecine clinique et experimentale, 1996, 19(3), 149-160.

AYERS, K. L., DAVIDSON, N. M., DEMIYAH, D., ROESZLER, K. N., GRÜTZNER, F., SINCLAIR, A. H., ... & SMITH, C. A. RNA sequencing reveals sexually dimorphic gene expression before gonadal differentiation in chicken and allows comprehensive annotation of the W-chromosome. Genome Biology, 2013, 14(3), R26.

BENNETT, D. C. & LAMOREUX, M. L. The color loci of mice–a genetic century. Pigment Cell Research, 2003, 16(4), 333-344.

BAR-PELED, L., CHANTRANUPONG, L., CHERNIACK, A. D., CHEN, W. W., OTTINA, K. A., GRABINER, B. C., ... & SABATINI, D. M. A Tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. Science, 2013, 340(6136), 1100-1106.

BARRETT, R. D., & SCHLUTER, D. Adaptation from standing genetic variation. Trends in Ecology & Evolution, 2008, 23(1), 38-44.

BEISEL, K. W., ROCHA-SANCHEZ, S. M., ZIEGENBEIN, S. J., MORRIS, K. A., KAI, C., KAWAI, J., ... & DAVIS, R. L. Diversity of Ca^{2+}-activated K^{+} channel transcripts in inner ear hair cells. Gene, 2007, 386(1), 11-23.

BRAGLIA, S., DAVOLI, R., ZAPPAVIGNA, A., ZAMBONELLI, P., BUTTAZZONI, L., GALLO, M., & RUSSO, V. SNPs of MYPN and TTN genes are associated to meat and carcass traits in Italian Large White and Italian Duroc pigs. Molecular Biology Reports, 2013, 40(12), 6927-6933.

BROMAN, K. W., MURRAY, J. C., SHEFFIELD, V. C., WHITE, R. L., & WEBER, J. L. Comprehensive human genetic maps: individual and sex-specific variation in recombination. The American Journal of Human Genetics, 1998, 63(3), 861-869.

BRUHAT, A., CHERASSE, Y., MAURIN, A. C., BREITWIESER, W., PARRY, L., DEVAL, C., ... & FAFOURNOUX, P. ATF2 is required for amino acid-regulated transcription by orchestrating specific histone acetylation. Nucleic Acids Research, 2007, 35(4), 1312-1321.

CHALMERS, I. J., HÖFLER, H., & ATKINSON, M. J. Mapping of a cadherin gene cluster to a region of chromosome 5 subject to frequent allelic loss in carcinoma. Genomics, 1999, 57(1), 160-163.

CHEN, H., PALMER, J. S., THIAGARAJAN, R. D., DINGER, M. E., LESIEUR, E., CHIU, H., ... & WILHELM, D. Identification of novel markers of mouse fetal ovary development. Plos One, 2012, 7(7), e41683.

CHEN, R., CHANG, G., DAI, A., MA, T., ZHAI, F., XIA, M., ... & CHEN, G. Cloning and expression characterization of the chicken Piwil1 gene. Molecular Biology Reports, 2013, 40(12), 7083-7091.

CORT, E. P., MORDES, J. P., LERNMARK, Å., & KWITEK, A. E. Genetic dissection reveals diabetes loci proximal to the. Physiol. Genomics, 2010, 41(1), 9-20.

CREUTZ, C. E., TOMSIG, J. L., SNYDER, S. L., GAUTIER, M. C., SKOURI, F., BEISSON, J., & COHEN, J. The copines, a novel class of C2 domain-containing, calciumdependent, phospholipid-binding proteins conserved from *Paramecium* to humans. Journal of Biological Chemistry, 1998, 273(3), 1393-1402.

DARMON, S. K., & LUTZ, C. S. mRNA 3′ end processing factors: a phylogenetic comparison. Comparative and Functional Genomics, 2012.

DELORENZO, R. J., & RUDDLE, F. H. Glutamate oxalate transaminase (GOT) genetics in *Mus musculus*: Linkage, polymorphism, and phenotypes of the Got-2 and Got-1 loci. Biochemical Genetics, 1970, 4(2), 259-273.

DESTA, T. T., DESSIE, T., BETTRIDGE, J., LYNCH, S. E., MELESE, K., COLLINS, M., CHRISTLEY, R. M., WIGLEY, P., KAISER, P., TERFA, Z., MWACHARO, J. M. & HANOTTE, O. Signature of artificial selection and ecological landscape on morphological structures of Ethiopian village chickens. Animal Genetic Resources, 2013, 52, 1-13.

DODD, M. E., HATZOLD, J., MATHIAS, J. R., WALTERS, K. B., BENNIN, D. A., RHODES, J., ... & HUTTENLOCHER, A. The ENTH domain protein Clint1 is required for epidermal homeostasis in zebrafish. Development, 2009, 136(15), 2591-2600.

DUESTER, G. Genetic dissection of retinoid dehydrogenases. Chemico-biological interactions, 2001,130, 469-480.

FISCHER, S., KOHLHASE, J., BÖHM, D., SCHWEIGER, B., HOFFMANN, D., HEITMANN, M., ... & WIECZOREK, D. Biallelic loss of function of the promyelocytic leukaemia zinc finger (PLZF) gene causes severe skeletal defects and genital hypoplasia. Journal of Medical Genetics, 2008, 45(11), 731-737.

FLINK, L. G., ALLEN, R., BARNETT, R., MALMSTRÖM, H., PETERS, J., ERIKSSON, J., ... & LARSON, G. Establishing the validity of domestication genes using DNA from ancient chickens. Proceedings of the National Academy of Sciences, 2014, 111(17), 6184-6189.

FURUKAWA, Y., SHIMADA, T., FURUTA, H., MATSUNO, S., KUSUYAMA, A., DOI, A., ... & NANJO, K. Polymorphisms in the IDE-KIF11-HHEX gene locus are reproducibly associated with type 2 diabetes in a Japanese population. The Journal of Clinical Endocrinology & Metabolism, 2008, 93(1), 310-314.

GALY, B., FERRING-APPEL, D., SAUER, S. W., KADEN, S., LYOUMI, S., PUY, H., ... & HENTZE, M. W. Iron regulatory proteins secure mitochondrial iron sufficiency and function. Cell Metabolism, 2010, 12(2), 194-201.

GAUTIER, M., & VITALIS, R. rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. Bioinformatics, 2012, 28(8), 1176-1177.

GOICOECHEA, S. M., ARNEMAN, D., & OTEY, C. A. The role of palladin in actin organization and cell motility. European Journal of Cell Biology, 2008, 87(8), 517-525.

GOLLMER, K., ASPERTI-BOURSIN, F., TANAKA, Y., OKKENHAUG, K., VANHAESEBROECK, B., PETERSON, J. R., ... & STEIN, J. V. CCL21 mediates CD4+ T-cell costimulation via a DOCK2/Rac-dependent pathway. Blood, 2009, 114(3), 580-588.

GRØNSKOV, K., DOOLEY, C. M., ØSTERGAARD, E., KELSH, R. N., HANSEN, L., LEVESQUE, M. P., ... & ROSENBERG, T. Mutations in C10orf11, a melanocyte-differentiation gene, cause autosomal-recessive albinism. The American Journal of Human Genetics, 2013, 92(3), 415-421.

GU, A., JI, G., SHI, X., LONG, Y., XIA, Y., SONG, L., ... & WANG, X. Genetic variants in Piwi-interacting RNA pathway genes confer susceptibility to spermatogenic failure in a Chinese population. Human Reproduction, 2010, deq274.

GUAN, Y. J., YANG, X., WEI, L., & CHEN, Q. MiR-365: a mechanosensitive microRNA stimulates chondrocyte differentiation through targeting histone deacetylase 4. The FASEB Journal, 2011, 25(12), 4457-4466.

GUPTA, S. D., LEE, B. T., CAMAKARIS, J., & WU, H. C. Identification of cutC and cutF (nlpE) genes involved in copper tolerance in Escherichia coli. Journal of Bacteriology, 1995, 177(15), 4207-4215.

HAESELEER, F., HUANG, J., LEBIODA, L., SAARI, J. C., & PALCZEWSKI, K. Molecular characterization of a novel short-chain dehydrogenase/reductase that reduces all-trans-retinal. Journal of Biological Chemistry, 1998, 273(34), 21790-21799.

HANAFUSA, H., ISHIKAWA, K., KEDASHIRO, S., SAIGO, T., IEMURA, S. I., NATSUME, T., ... & MATSUMOTO, K. Leucine-rich repeat kinase LRRK1 regulates endosomal trafficking of the EGF receptor. Nature Communications, 2011, 2, 158.

HSU, R., WOODROFFE, A., LAI, W. S., COOK, M. N., MUKAI, J., DUNNING, J. P., ... & GOGOS, J. A. Nogo Receptor 1 (RTN4R) as a candidate gene for schizophrenia: analysis using human and mouse genetic approaches. Plos One, 2007, 2(11), e1234.

HUANG, D. W., SHERMAN, B. T. & LEMPICKI, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research, 2009a, 37(1), 1-13.

HUANG, D. W., SHERMAN, B. T. & LEMPICKI, R. A. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protocol, 2009b, 4(1), 44-57.

HUME, A. N., COLLINSON, L. M., RAPAK, A., GOMES, A. Q., HOPKINS, C. R., & SEABRA, M. C. Rab27a regulates the peripheral distribution of melanosomes in melanocytes. The Journal of Cell Biology, 2001, 152(4), 795-808.

IMSLAND, F., FENG, C., BOIJE, H., BED'HOM, B., FILLON, V., DORSHORST, B., ... & ANDERSSON, L. The Rose-comb mutation in chickens constitutes a structural rearrangement causing both altered comb morphology and defective sperm motility. PLoS Genetics, 2012, 8(6), e1002775.

JIAO, Y., ZAN, L. S., LIU, Y. F., WANG, H. B., & GUO, B. L. A novel polymorphism of the MYPN gene and its association with meat quality traits in Bos taurus. Genetics and Molecular Research, 2010, 9(3), 1751-1758.

JONES, F. S., CHALEPAKIS, G., GRUSS, P., & EDELMAN, G. M. Activation of the cytotactin promoter by the homeobox-containing gene Evx-1. Proceedings of the National Academy of Sciences, 1992, 89(6), 2091-2095.

KALDIS, P. & PAGANO, M. Wnt signaling in mitosis. Developmental Cell, 2009, 17(6), 749-750.

KAST, H. R., GOODWIN, B., TARR, P. T., JONES, S. A., ANISFELD, A. M., STOLTZ, C. M., ... & EDWARDS, P. A. Regulation of multidrug resistance-associated protein 2 (ABCC2) by the nuclear receptors pregnane X receptor, farnesoid X-activated receptor, and constitutive androstane receptor. Journal of Biological Chemistry, 2002, 277(4), 2908-2915.

KASTHURI, S. R., WAN, Q., UMASUTHAN, N., BATHIGE, S. D. N. K., LIM, B. S., JUNG, H. B., ... & WHANG, I. Genomic characterization, expression analysis, and antimicrobial function of a glyrichin homologue from rock bream, *Oplegnathus fasciatus*. Fish & Shellfish Immunology, 2013, 35(5), 1406-1415.

KAWAKAMI, Y., WADA, N., NISHIMATSU, S. I., KOMAGUCHI, C., NOJI, S., & NOHNO, T. Identification of chick frizzled-10 expressed in the developing limb and the central nervous system. Mechanisms of Development, 2000, 91(1), 375-378.

KHATRI, S. B., & GROVES, A. K. Expression of the Foxi2 and Foxi3 transcription factors during development of chicken sensory placodes and pharyngeal arches. Gene Expression Patterns, 2013, 13(1), 38-42.

KIM, Y. & NIELSEN, R. Linkage disequilibrium as a signature of selective sweeps. Genetics, 2004, 167, 1513-1524.

KIM, Y. M., STONE, M., HWANG, T. H., KIM, Y. G., DUNLEVY, J. R., GRIFFIN, T. J., & KIM, D. H. SH3BP4 is a negative regulator of amino acid-Rag GTPase-mTORC1 signaling. Molecular Cell, 2012, 46(6), 833-846.

KOEGL, M., HOPPE, T., SCHLENKER, S., ULRICH, H. D., MAYER, T. U., & JENTSCH, S. A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. Cell, 1999, 96(5), 635-644.

KOHROKI, J., NISHIYAMA, T., NAKAMURA, T., & MASUHO, Y. ASB proteins interact with Cullin5 and Rbx2 to form E3 ubiquitin ligase complexes. FEBS letters, 2005, 579(30), 6796-6802.

KOUTSOURAKIS, M., LANGEVELD, A., PATIENT, R., BEDDINGTON, R., & GROSVELD, F. The transcription factor GATA6 is essential for early extraembryonic development. Development, 1999, 126(4), 723-732.

KRANIS, A., GHEYAS, A. A., BOSCHIERO, C., TURNER, F., YU, L., SMITH, S., ... & BURT, D. W. Development of a high density 600K SNP genotyping array for chicken. BMC Genomics, 2013, 14(1), 59.

KURAKU, S., & MEYER, A. The evolution and maintenance of Hox gene clusters in vertebrates and the teleost-specific genome duplication. International Journal of Developmental Biology, 2009, 53(5), 765.

KUWANO, R., MIYASHITA, A., ARAI, H., ASADA, T., IMAGAWA, M., SHOJI, M., ... & IHARA, Y. Dynamin-binding protein gene on chromosome 10q is associated with late-onset Alzheimer's disease. Human Molecular Genetics, 2006, 15(13), 2170-2182.

LEFEBVRE, V., BEHRINGER, R. R., & DE CROMBRUGGHE, B. L-Sox5, Sox6 and Sox9 control essential steps of the chondrocyte differentiation pathway. Osteoarthritis and Cartilage, 2001, 9, S69-S75.

LYIMO, C. M., WEIGEND, A., MSOFFE, P. L., EDING, H., SIMIANER, H. & WEIGEND, S. Global diversity and genetic contributions of chicken populations from African, Asian and European regions. Animal Genetics, 2014, 45(6), 836–848.

MARELJA, Z., STÖCKLEIN, W., NIMTZ, M., & LEIMKÜHLER, S. A novel role for human nfs1 in the cytoplasm NFS1 acts as a sulfur donor for MOCS3, a protein involved in molybdenum cofactor biosynthesis. Journal of Biological Chemistry, 2008, 283(37), 25178-25185.

MEHLEN, P., RABIZADEH, S., SNIPAS, S. J., ASSA-MUNT, N., SALVESEN, G. S., & BREDESEN, D. E. The DCC gene product induces apoptosis by a mechanism requiring receptor proteolysis. Nature, 1998, 395(6704), 801-804.

MOSHIRI, A., CLOSE, J., & REH, T. A. Retinal stem cells and regeneration. The International Journal of Developmental Biology, 2004, 48(8-9), 1003-14.

MÜLLER-RÖVER, S., ROSSITER, H., LINDNER, G., PETERS, E. M., KUPPER, T. S., & PAUS, R. Hair follicle apoptosis and Bcl-2. In Journal of Investigative Dermatology Symposium Proceedings (Vol. 4, No. 3, pp. 272-277). Nature Publishing Group, 1999, December.

NAKAMURA, Y., INLOES, J. B., KATAGIRI, T., & KOBAYASHI, T. Chondrocyte-specific microRNA-140 regulates endochondral bone development and targets Dnpep to modulate bone morphogenetic protein signaling. Molecular and Cellular Biology, 2011, 31(14), 3019-3028.

NELSON, C. E., MORGAN, B. A., BURKE, A. C., LAUFER, E., DIMAMBRO, E., MURTAUGH, L. C., ... & TABIN, C. Analysis of Hox gene expression in the chick limb bud. Development, 1996, 122(5), 1449-1466.

NYLANDER, K., BOURDON, J. C., BRAY, S. E., GIBBS, N. K., KAY, R., HART, I., & HALL, P. A. Transcriptional activation of tyrosinase and TRP-1 by p53 links UV irradiation to the protective tanning response. The Journal of Pathology, 2000, 190(1), 39-46.

ONG, R.T. & TEO, Y. Y. varLD: a program for quantifying variation in linkage disequilibrium patterns between populations. Bioinformatics, 2010, 26, 1269-1270.

PARAKATI, R., & DIMARIO, J. X. Sp1-and Sp3-mediated transcriptional regulation of the fibroblast growth factor receptor 1 gene in chicken skeletal muscle cells. Journal of Biological Chemistry, 2002, 277(11), 9278-9285.

PAVLIDIS, P., ZIVKOVIC, D., STAMATAKIS, A. & ALACHIOTIS, N. SweeD: likelihood-based detection of selective sweeps in thousands of genomes. Molecular Biology and Evolution, 2013, 30, 2224-2234.

PIEKARSKI, A., KHALDI, S., GREENE, E., LASSITER, K., MASON, J. G., ANTHONY, N., ... & DRIDI, S. Tissue Distribution, Gender-and Genotype-Dependent Expression of Autophagy-Related Genes in Avian Species. PloS One, 2014, 9(11), e112449.

QANBARI, S., PAUSCH, H., JANSEN, S., SOMEL, M., STROM, T. M., FRIES, R., ... & SIMIANER, H. Classic selective sweeps revealed by massive sequencing in cattle. PLoS Genetics, 2014, 10(2), e1004148.

QUINTANA-MURCI, L., & CLARK, A. G. Population genetic tools for dissecting innate immunity in humans. Nature Reviews Immunology, 2013, 13(4), 280-293.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, 2013. http://www.R-project.org/.

RAMEY, H. R., DECKER, J. E., MCKAY, S. D., ROLF, M. M., SCHNABEL, R. D., & TAYLOR, J. F. Detection of selective sweeps in cattle using genome-wide SNP data. BMC Genomics, 2013, 14(1), 382.

ROZAKIS-ADCOCK, M., MCGLADE, J., MBAMALU, G., PELICCI, G., DALY, R., LI, W., ... & PAWSON, T. Association of the Shc and Grb2/Sem5 SH2-containing proteins is implicated in activation of the Ras pathway by tyrosine kinases. Nature, 1992, 360(6405), 689-692.

RUBIN, C. J., ZODY, M. C., ERIKSSON, J., MEADOWS, J. R., SHERWOOD, E., WEBSTER, M. T., ... & ANDERSSON, L. Whole-genome resequencing reveals loci under selection during chicken domestication. Nature, 2010, *464*(7288), 587-591.

SCHEET, P. & STEPHENS, M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. The American Journal of Human Genetics, 2006, 78(4), 629-644.

SCHNUR, R. E., GAO, M., WICK, P. A., KELLER, M., BENKE, P. J., EDWARDS, M. J., ... & WELEBER, R. G. OA1 mutations and deletions in x-linked ocular albinism. The American Journal of Human Genetics, 1998, 62(4), 800-809.

SEO, H. R., LEE, D. H., LEE, H. J., BAEK, M., BAE, S., SOH, J. W., ... & LEE, Y. S. Cyclin G1 overcomes radiation-induced G2 arrest and increases cell death through transcriptional activation of cyclin B1. Cell Death & Differentiation, 2005, 13(9), 1475-1484.

SHI, J. D., KUKAR, T., WANG, C. Y., LI, Q. Z., CRUZ, P. E., DAVOODI-SEMIROMI, A., ... & SHE, J. X. Molecular cloning and characterization of a novel mammalian endo-apyrase (LALP1). Journal of Biological Chemistry, 2001, 276(20), 17474-17478.

SHIMOYAMA, A., WADA, M., IKEDA, F., HATA, K., MATSUBARA, T., NIFUJI, A., ... & YONEDA, T. Ihh/Gli2 signaling promotes osteoblast differentiation by regulating Runx2 expression and function. Molecular Biology of the Cell, 2007, 18(7), 2411-2418.

SITARAM, A., PICCIRILLO, R., PALMISANO, I., HARPER, D. C., DELL'ANGELICA, E. C., SCHIAFFINO, M. V., & MARKS, M. S. Localization to mature melanosomes by virtue of cytoplasmic dileucine motifs is required for human OCA2 function. Molecular Biology of the Cell, 2009, 20(5), 1464-1477.

SIVA, K., VENU, P., MAHADEVAN, A., SHANKAR, S. K., & INAMDAR, M. S. Human BCAS3 expression in embryonic stem cells and vascular precursors suggests a role in human embryogenesis and tumor angiogenesis. Plos One, 2007, 2(11), e1202.

SMITH, L. M., & BURGOYNE, L. A. Collecting, archiving and processing DNA from wildlife samples using FTA® databasing paper. BMC Ecology, 2004, 4(1), 4.

STAINES, K., YOUNG, J. R., & BUTTER, C. Expression of chicken DEC205 reflects the unique structure and function of the avian immune system. PloS One, 2013, 8(1), e51799.

SUNRYD, J. C., CHEON, B., GRAHAM, J. B., GIORDA, K. M., FISSORE, R. A., & HEBERT, D. N. TMTC1 and TMTC2 are novel endoplasmic reticulum tetratricopeptide repeat-containing adapter proteins involved in calcium homeostasis. Journal of Biological Chemistry, 2014, 289(23), 16085-16099.

TAKEI, Y., SWIETLIK, M., TANOUE, A., TSUJIMOTO, G., KOUZARIDES, T., & LASKEY, R. MCM3AP, a novel acetyltransferase that acetylates replication protein MCM3. EMBO reports, 2001, 2(2), 119-123.

TANG, K., THORNTON, K. R., & STONEKING, M. A new approach for using genome scans to detect recent positive selection in the human genome. PLoS Biology, 2007, 5(7), e171.

TATSUMI, K., SOU, Y. S., TADA, N., NAKAMURA, E., IEMURA, S. I., NATSUME, T., ... & KOMATSU, M. A novel type of E3 ligase for the Ufm1 conjugation system. Journal of Biological Chemistry, 2010, 285(8), 5417-5427.

TEO, Y. Y., SIM, X., ONG, R. T., TAN, A. K., CHEN, J., TANTOSO, E., ... & CHIA, K. S. Singapore Genome Variation Project: a haplotype map of three Southeast Asian populations. Genome Research, 2009, 19(11), 2154-2162.

UTSUNOMIYA, Y. T., O'BRIEN, A. M. P., SONSTEGARD, T. S., VAN TASSELL, C. P., DO CARMO, A. S., MESZAROS, G., ... & GARCIA, J. F. Detecting loci under recent positive selection in dairy and beef cattle by combining different genome-wide scan methods. PloS One, 2013, 8(5), e64280.

VOIGHT, B. F., KUDARAVALLI, S., WEN, X., & PRITCHARD, J. K. A map of recent positive selection in the human genome. PLoS Biology, 2006, 4(3), e72.

WANG, C. Q., AKALU, Y. T., SUAREZ-FARINAS, M., GONZALEZ, J., MITSUI, H., LOWES, M. A., ... & KRUEGER, J. G. IL-17 and TNF synergistically modulate cytokine expression while suppressing melanogenesis: potential relevance to psoriasis. Journal of Investigative Dermatology, 2013, 133(12), 2741-2752.

WANG, S. F., OH, S., SI, Y. X., WANG, Z. J., HAN, H. Y., LEE, J., & QIAN, G. Y. Computational prediction of protein-protein interactions of human tyrosinase. Enzyme Research, 2012 (2012).

WASIM, M., CARLET, M., MANSHA, M., GREIL, R., PLONER, C., TROCKENBACHER, A., ... & KOFLER, R. PLZF/ZBTB16, a glucocorticoid response gene in acute lymphoblastic leukemia, interferes with glucocorticoid-induced apoptosis. The Journal of Steroid Biochemistry and Molecular Biology, 2010, 120(4), 218-227.

WIENER, P., & WILKINSON, S. Deciphering the genetic basis of animal domestication. Proceedings of the Royal Society B: Biological Sciences, 2011, 278(1722), 3161-3170.

WILHELMSEN, K., LITJENS, S. H., KUIKMAN, I., TSHIMBALANGA, N., JANSSEN, H., VAN DEN BOUT, I., ... & SONNENBERG, A. Nesprin-3, a novel outer nuclear membrane protein, associates with the cytoskeletal linker protein plectin. The Journal of Cell Biology, 2005, 171(5), 799-810.

WILKINS, S. J., YOONG, S., VERKADE, H., MIZOGUCHI, T., PLOWMAN, S. J., HANCOCK, J. F., ... & PERKINS, A. C. Mtx2 directs zebrafish morphogenetic movements during epiboly by regulating microfilament formation. Developmental Biology, 2008, 314(1), 12-22.

WRIGHT, D., BOIJE, H., MEADOWS, J. R., BED'HOM, B., GOURICHON, D., VIEAUD, A., ... & ANDERSSON, L. Copy number variation in intron 1 of SOX5 causes the Pea-comb phenotype in chickens. PLoS Genetics, 2009, 5(6), e1000512.

YAMABE, K., MAEDA, H., KOKEGUCHI, S., SOGA, Y., MEGURO, M., NARUISHI, K., ... & TAKASHIBA, S. Antigenic group II chaperonin in Methanobrevibacter oralis may cross-react with human chaperonin CCT. Molecular Oral Microbiology, 2010, 25(2), 112-122.

YAMAMOTO, M., GOTOH, Y., TAMURA, K., TANAKA, M., KAWAKAMI, A., IDE, H., & KUROIWA, A. Coordinated expression of Hoxa-11 and Hoxa-13 during limb muscle patterning. Development, 1998, 125(7), 1325-1335.

YANG, Y., YAN, X., CAI, Y., LU, Y., SI, J., & ZHOU, T. NudC-like protein 2 regulates the LIS1/dynein pathway by stabilizing LIS1 with Hsp90. Proceedings of the National Academy of Sciences, 2010, 107(8), 3499-3504.

YATSU, A., OHBAYASHI, N., TAMURA, K., & FUKUDA, M. Syntaxin-3 is required for melanosomal localization of Tyrp1 in melanocytes. Journal of Investigative Dermatology, 2013, 133(9), 2237-2246.

ZHAI, L. W., WANG, L. X., ZHOU, W. L., & WANG, C. D. Association of the MTPN gene polymorphisms with meat quality in commercial pigs. Journal of Animal and Veterinary Advances, 2010, 9(4), 705-709.

Chapter 8

Genome-wide association study of threshold and Mendelian traits in Ethiopian village chickens

Abstract

Traits showing Mendelian mode of inheritance have been the subject of intense research. Chickens are usually used as non-laboratory model species in Mendelian genetics. We map putative loci that underlie variation in threshold traits of panmictic Ethiopian village chickens using a high density (600K) SNP array. We perform association mapping of pigmentation traits and variants of crest, comb shape and a lightly feathered shank. We remapped some loci underlying variation in Mendelian traits of the chicken. Our analysis also identifies causative variants of pigmentation traits that have been mapped in mouse along with a long list of *de novo* loci. Crest phenotype shows a significant association with Z chromosome SNPs, which then implicate the impact of sex chromosome on the trait thought to be autosomal. We found that traits like crested head and comb shape which have been thought to be simple traits may be under the control of many loci. A considerable proportion of the putative loci mapped were overlapped among the traits studied, which could then implicate a pleiotropic interaction. Moreover, many loci may control the genetic basis of each trait indicating a considerable impact of epistatic interaction. In this study, we produce a long list of putative loci underlying the variation in threshold and Mendelian traits and for the first time we map the genetic control of a buttercup comb mutation. Our analysis corroborates the appropriateness of outbred populations in fine-scale mapping.

Introduction

Village chickens show a dazzling array of morphological diversity arising from the impact of natural selection and anthropogenic effects. Mendelian traits are highly segregating in village chickens due to absence of strong artificial selection which otherwise leads to fixation of selected variants of a morph. Most of the studies conducted on morph variants of village chickens have been limited to morphological

scoring. Little effort has been made to map causative variants underlying Mendelian traits variation in village chickens using high density markers (Wragg et al., 2012; Li et al., 2014). Mendelian traits however play a significant role in socio-cultural life of subsistence farmers and in local adaptation. Farmers and consumers usually show different degree of preference to variants of a morph (e.g. plumage colour, comb shape, shank colour *etc.*). Preference of farmers however may vary based on their demographic structure and locally-driven use values of chickens.

Morphological diversity accelerates adaptive radiation of species. For example, dull plumage enables chickens to blend with their micro-habitat, which then helps them to avoid predators. Farmers in Ethiopia usually report exposure of white chickens to visually hunting predators, which has made the self-white plumage to be less preferred by some farmers (CHAPTER THREE). The junglefowl hens also display a cryptic plumage that enables them to avoid predators. Pigmentation traits could also involve in intraspecific and interspecific communication and photoprotection (see Hubbard et al., 2010 and the references therein). Comb type is involving in thermoregulation (Richards, 1971; Gerken et al., 2006) and homozygous rose comb may reduce fertility in cocks (Hindhaugh, 1932; Hutt, 1940; Crawford 1971). An increased in surface area as in a single comb facilitates evaporative cooling in warm climates, while reduced surface area as in the derived comb variants (e.g. rose, pea, walnut) may help to reduce heat loss in cold environments. Morphological appearance influences interaction among mates (Zuk et al., 1995), which then impacts their reproductive success.

Mendelian traits have been used for long in classical genetic studies. Among food animals, chickens are commonly used as model species to study the genetic control of loci showing Mendelian mode of inheritance. Mendelian traits commonly thought to be under the control of a single locus or a few loci, which make them to display a discrete class of phenotypes. There are traits that can be classified as quasi-Mendelian (threshold traits); in spite of their polygenic nature, some loci underlying these traits explain a large proportion of a genetic variation, which then make a phenotype to show a graded (qualitative) mode of expression. Plumage colour and pattern in chickens can be classified as threshold trait despite a vast array of variation among non-self-type (nondescript) plumages. Other pigmentation traits like earlobe

and shank colour also tend to show a Mendelian proportion inheritance pattern despite their polygenic nature.

Outbred populations like village chickens are good resources in fine mapping studies, because historical recombination breaks down a long-range linkage disequilibrium, which then refines further causative genomic regions. Most of the Mendelian traits mapping conducted so far have been concentrated on fancy chickens that have been intensively selected for show traits (Dorshorst et al., 2010; Wragg et al., 2012; Siwek et al, 2013). A high diversity in Mendelian traits of village chickens thought to be a consequence of natural selection besides a mild impact of anthropogenic effect (Desta et al., 2013). Mapping studies in village chickens therefore could serve as a bench-mark to study the evolutionary genetics of panmictic populations.

We map polygenic traits displaying a Mendelian proportion of variation in outbred Ethiopian village chickens sampled from two distantly located geographical regions (Desta et al., 2013) using a high density (600K) SNP array (Kranis et al., 2013). Our study involves pigmentation traits and variants of a lightly feathered shank, crested head and comb shape. Our analyses remapped causative variants previously identified in livestock and mouse and we identify a large number of *de novo* loci that underlie the Mendelian variation of the traits studied. We also substantiate that the genetic basis of the thought to be simple traits (e.g. comb, crest) is likely under the control of many loci.

Materials and methods

The study populations and data quality control

A genome-wide association study of traits showing Mendelian mode of inheritance pattern was performed in Horro (n = 383) and Jarso (n = 373) chickens of Ethiopia using a high density (600K) SNP array. Village chickens show a vast array of morphological diversity as they have not been selected or standardized for any traits of special interest, which then in some instances makes the scoring of morph variants a challenging task. In the GWAS analysis we include a subset of samples that

display less ambiguous variants of a morph to improve the power of the association mapping.

We perform a quality control (QC) on 760 chickens and 572762 SNPs assayed on twenty eight autosomes (GGA1–28) and the Z chromosome. We used the R (R Core Team, 2013) package GenABEL (Aulchenko et al., 2007) both for QC and association mapping. The QC criteria adopted are minor allele frequency (MAF = 0.01), SNP and individual calling rate (0.9), ibs.mrk (All), ibs.threshold (0.95) and ibs.exclude (lower). Following the QC, 29950 and 1578 SNPs were excluded due to low MAF and low calling rate respectively. Four chickens were excluded due to high IBS. The QC retains 541236 SNPs and 756 chickens for a downstream analysis.

Data analysis

We analyse variants of plumage, earlobe and shank colour and a lightly feathered shank, comb shape and crest. Two GenABEL functions – qtscore (fast score test for association) and ccfast (allelic chi-square test) were used for GWAS analysis. False discovery rate (FDR) corrected P-value (Benjamini & Hochberg, 1995) was set to 10% using the qvaluebh95 function of GenABEL and is displayed on Manhattan plot along the Bonferroni corrected P-value (9.23811E-08) that accounts for a multiple testing. We used a p-value corrected for inflation factor (Pc1df) in ccfast analysis to account for stratification effects.

A number of pairwise comparisons were made between contrasting variants of each morphological trait to check how often a causative variant can be mapped by different analyses. The traits analysed are crested head (n = 102) versus plain (n = 379) hens as crest was commonly scored in hens. Shank colour was contrasted between wild (dark shanks vis-v-vis black (n = 25), slate blue (n = 25) and green (n = 20) versus derived variants (brightly-coloured – white (n = 126) and yellow (n = 403), among five commonly observed variants of shank colour (yellow, white, black, slate blue and green), between wild variants and yellow shank and between white and yellow shank. A GWAS analysis also made on a lightly feathered (n = 6) versus non-feathered shank (n = 746). Earlobe colour was analysed for spotted (n = 316) versus plain earlobe (n = 406), among four commonly observed variants of earlobe

(purple, white spotted purple, red, white spotted red) and red versus white earlobe. Plumage colour and pattern was mapped for ten pairs of analyses involving self-white cocks and hens (n = 20), self-black hens (n = 25), red cocks (n = 91), silver birchen cocks (n = 5) and red-pyle cocks (n = 28). Population (Horro and Jarso) and/or sex (cock and hen) used as covariate(s) in the qtscore analysis. As in the case of a lightly feathered shank, we tried to balance the number of cases and controls but this has reduced the resolution power of mapping studies. Therefore, according to our finding the use of highly contrasting and clearly defined phenotypes is the one matters the quality of GWAS than balancing the number of cases and controls.

A genomic region bracketing 50Kb upstream and 50Kb downstream of a SNP showing significant association was scanned to identify putative loci underlying the causative variant. We used BioMart portal of the Ensembl genome browser (Ensembl genes 77 and *Gallus gallus* genes (Galgal4)) to map putative genes. Functional annotation of putative genes was performed using DAVID (http://david.abcc.ncifcrf.gov/, Huang et al., 2009a & b). List of pigmentation genes that have been mapped in mouse was obtained from http://www.espcr.org/micemut/ and this has been used as a reference to validate putative loci that have been found to associate with pigmentation traits. Moreover, a summary presented at http://www.informatics.jax.org/searches/GOannot_report.cgi?id=GO:0043473 was used in some instances as a reference to search literatures dealing with pigmentation genes besides genecards (http://www.genecards.org/) and google scholar.

Results

A genome-wide association study was performed between and among variants of the morphological traits analysed. For a considerable proportion of the association mapping performed we used a moderate number of samples (see MATERIALS and METHODS section), which may be adequate enough to map traits showing Mendelian proportion of inheritance pattern.

Earlobe colour

A genome-wide association mapping of earlobe colour was performed for spotted versus plain earlobe, and among four commonly observed variants (purple, white spotted purple, red, white spotted red) and red versus white earlobe. A considerable number of overlapping loci were mapped among the three GWAS analyses, which could then validate our mapping approach.

Spotted versus plain earlobe

A ccfast analysis performed on plain versus spotted earlobe identifies seven SNPs crossing the FDR corrected P-value and four of these also passed the Bonferroni corrected P-value. We found significantly associated SNPs both on the autosomes and the Z chromosome with lack of evidence for sexual dimorphism of earlobe spotting. Seven SNPs that show a significant association are presented in Table 8.1. Five putative genes which are closely located with the SNPs were mapped and among these CTNND2 (GGA2) may play antagonistic role to tyrosine phosphorylation (Martinez et al., 2003). Moreover, a gene with Ensembl ID ENSGALG00000014545 coding for uncharacterized protein on GGA1 (161024899–161059435) was mapped and this novel loci may also contribute to earlobe spotting as a clearly visible peak is located close to this gene (Figure 8.1).

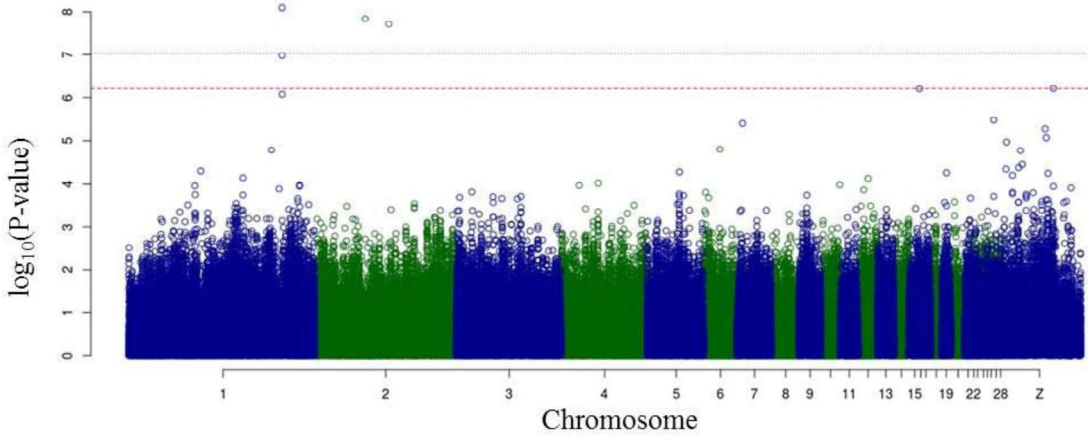


Figure 8.1 A single marker genome-wide association study of spotted versus plain earlobe colour. The x-axis represents the chromosomal position of each SNP and the y-axis indicates –log10 P-value. The dotted red line represents the FDR corrected P-value, while the black dotted line represents the Bonferonni P-value corrected for multiple testing.

Table 8.1 Seven SNPs that passed FDR corrected P-value and that associate with earlobe spotting.

Chr	Pos	P-value (Pc1df)	qvalue	Associated genes
1	160965706	7.96E-09	0.003485	
2	53094118	1.44E-08	0.003485	
2	77917805	1.93E-08	0.003485	CTNND2
1	160965926	1.05E-07	0.014194	
Z	52929780	6.19E-07	0.056535	
15	11448619	6.27E-07	0.056535	C15H12ORF49, RNFT2
1	161040375	8.46E-07	0.065387	ENSGALG00000014545

Earlobe colour variants

Ten SNPs crossed the FDR corrected P-value from a qtscore analysis of four earlobe colour variants (purple, white spotted purple, red, white spotted red), however none of these passed the Bonferonni corrected P-value. List of the ten SNPs is presented in Table 8.2. A strong signal of association was observed on GGA1 & 3 (Figure 8.2). Among the thirty four putative genes closely associated with the SNPs, PTPRT may involve in tyrosine phosphorylation (Wang et al., 2004) and PLCB1 among other functions regulates cell growth (Peruzzi et al., 2002). Zeb2 involves in the regulation of melanocytes development and differentiation (Denecker et al., 2014) and Szt2 may involve in pigmentation (Frankel et al., 2009).

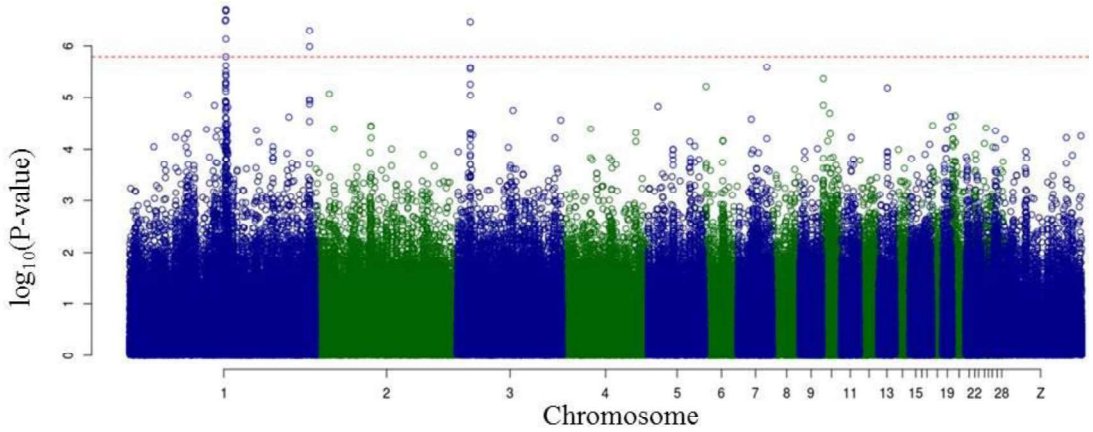


Figure 8.2 A single marker genome-wide association study of four classes of earlobe colours. The x-axis represents the chromosomal position of each SNP and the y-axis indicates –log10 P-value. The dotted red line represents FDR corrected P-value, while the black dotted line represents Bonferonni P-value corrected for multiple testing.

Table 8.2 Description of ten SNPs that passed the FDR corrected P-value and found to associate with most commonly observed earlobe colour variants in Ethiopian village chickens.

Chr	Pos	P-value (P1df)	qvalue	Associated genes
1	100286171	3.23E-07	0.030489	NCAM2
1	101015707	1.60E-06	0.086488	
1	101021944	2.11E-07	0.030489	
1	101025002	2.07E-07	0.030489	
1	101030605	1.98E-07	0.030489	
1	101034579	3.09E-07	0.030489	
1	101038793	7.41E-07	0.050128	
1	188959995	5.01E-07	0.038718	DLG2
1	188966676	1.02E-06	0.061604	DLG2
3	13827915	3.38E-07	0.030489	PLCB1

Red versus white earlobe

Three thousand and ninety seven SNPs crossed the FDR corrected P-value (Figure 8.3) and seven hundred seven putative genes are closely located. Among the putative loci mapped, TYRP1 and TYR have a well-established role in avian melanogenesis (Domyan et al., 2014). Among genes that have a systemic effect on melanogenesis we map HEPH, RXRA, RAB27A, RAB14, RABGAP1, RABEPK, RAB5A, RAB42 and RAB23 (http://www.espcr.org/micemut/) and Rab1 may involve in melanosomes trafficking (Hume et al., 2001). MAPK regulates the function of MITF at post-transcriptional level (Ebanks et al., 2009) and CREB binds and activates MITF promoter through cyclic adenosine monophosphate (cAMP) (Saha et al., 2006). DDC may involve in inhibiting of tyrosine activity (Satoh & Mishima, 1969) and BCL2 may inhibit apoptosis of the melanocytes (Müller-Röver et al., 1999). GRB2 activates the Ras guanine nucleotide-binding protein by tyrosine kinases (Rozakis-Adcock et al., 1992). F-KER and a number of uncharacterized loci (ENSGALG00000025900, ENSGALG00000024141, ENSGALG00000027640, ENSGALG00000028192, ENSGALG00000028211, ENSGALG00000026987, ENSGALG00000028843, ENSGALG00000009189 and ENSGALG00000018882) may involve in keratinocytes synthesis as it has been found from the description provided by BioMart portal of the Ensembl genome browser. Keratinocytes are structurally closely associated with melanocytes (Yu et al., 2004). Among the SNPs showing a significant association, the top ten are presented in Table 8.3.

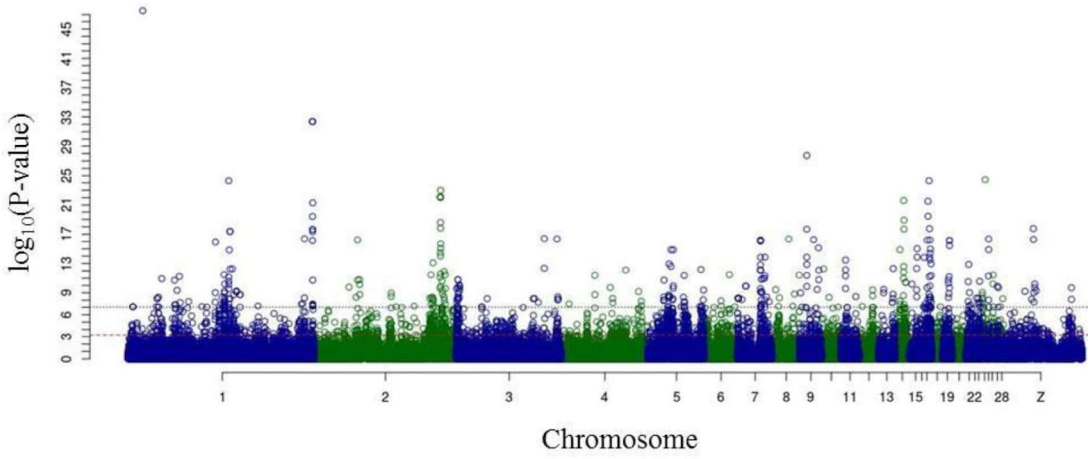


Figure 8.3 A single marker genome-wide association study of white versus red earlobe in Ethiopian village chickens. The x-axis represents the chromosomal position of each SNP and the y-axis indicates –log10 P-value. The dotted red line represents the FDR corrected P-value, while the black dotted line represents the Bonferonni corrected P-value.

Table 8.3 The most significantly associated ten SNPs with red and white earlobe colour variants in Ethiopian village chickens.

Chr	Pos	P-value (P1df)	qvalue	Associated genes
1	15396635	3.00E-48	1.62E-42	CPNE8
1	193425774	4.33E-33	5.86E-28	NUMA1, LCMT2
1	193432162	4.33E-33	5.86E-28	LCMT2, DCHS1
1	193435256	4.33E-33	5.86E-28	LCMT2, DCHS1
9	7681609	1.84E-28	1.99E-23	SCG2
24	3729134	3.53E-25	3.19E-20	GRIK4, ARHGEF12, TMEM136
17	7417865	5.25E-25	3.56E-20	COL5A1
1	105552906	5.26E-25	3.56E-20	RUNX1
2	132419051	1.03E-23	6.17E-19	
2	132243686	8.13E-23	3.67E-18	

Plumage colour and pattern

Red pyle versus red plumage

Five hundred and thirty four SNPs were found to show a significant association (Figure 8.4). Eight hundered one putative genes are closely linked with these SNPs. Among the putative genes ADAM17 is known to result in an irregular pigment formation (http://www.espcr.org/micemut/), which may be accomplished by reducing the synthesis of insoluble Pmel17 fragments (Kummer et al., 2009). ASIP is known to involve in melanogenesis and it favours the synthesis of pheomelanin (Kanetsky et al., 2002) and MED1 involves in melanogensis (http://www.espcr.org/micemut/). EDAR may involve in skin appendage development (Cui et al., 2007). EDNRB results in white spotting in megacolon and other neural crest defects (http://www.espcr.org/micemut/; Hosoda et al., 1994). EED results in diluted coat, FGFR2 makes lighter skin, FOXN1 results in travelling waves of dark/light and GATA3 results in an irregular pigment deposition (http://www.espcr.org/micemut/). MYO7A inovolves in melanosome transport in etinal pigment epithelium (RPE) (Klomp et al., 2007), PTS leads to coat dilution with low biopterin and high phenylalanine, TBX15 alters dorsoventral color pattern, UNC119 results in mottling of RPE and WNT3A makes defects of neural crest including the melanoblasts (http://www.espcr.org/micemut/). ERBB4 may associate with melanogenesis (Zhao et al., 1998; Prickett et al., 2009) and GSK3B regulates MITF at post-transcriptional level (Ebanks et al., 2009). PAX7 controls the early development of neural crest from which melanocytes are derived (Maczkowiak et al., 2010) and MAPK regulates the function of MITF at post-transcriptional level (Ebanks et al., 2009). DBH involves in hydroxylation of dopamine (Slominski et al., 2004). Rab29 (Wang et al., 2014); Rab11b (Tarafder et al., 2013) and Rab1 (Hume et al., 2001) may involve in melanosomes trafficking. MYC may involve in localization of melanosomes (Yatsu et al., 2013) and BCL2 may inhibit apoptosis of melanocytes (Müller-Röver et al., 1999). Ten SNPs that show the most significant association are presented in Table 8.4.

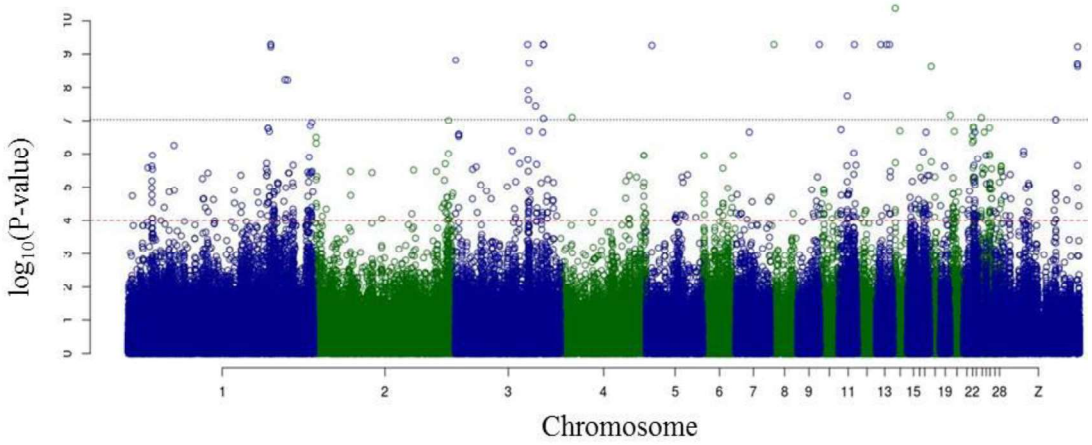


Figure 8.4 A single marker genome-wide association study of red pyle versus red cocks in Ethiopian village chickens. The x-axis represents the chromosomal position of each SNP and the y-axis indicates –log10 P-value. The dotted red line represents FDR corrected P-value, while the black dotted line represents Bonferonni P-value corrected for multiple testing.

Table 8.4 The most significant ten SNPs associate with switch between red and red pyle (red saddled (smoky) white) plumage in Ethiopian village chickens.

Chr	Pos	P-value (P1df)	qvalue	Associated genes
14	2300177	4.34E-11	2.23E-05	ENSGALG00000021238, ENSGALG00000004050, C7orf50, GPR146, ENSGALG00000021226
1	149954907	5.29E-10	2.23E-05	
1	149966517	5.29E-10	2.23E-05	
3	76168978	5.29E-10	2.23E-05	
3	92678994	5.29E-10	2.23E-05	MYT1L
3	92686630	5.29E-10	2.23E-05	MYT1L
8	3785708	5.29E-10	2.23E-05	ATF6, OLFML2B, HSD17B7, UAP1, UHMK1, ENSGALG00000002749
9	22854332	5.29E-10	2.23E-05	7., MBNL1
11	16296449	5.29E-10	2.23E-05	COTL1, KLHL36, USP10, ENSGALG00000020995
13	4771481	5.29E-10	2.23E-05	TENM2

Black versus red plumage

Eight hundred and fifty four SNPs crossed the FDR corrected P-value (Figure 8.5). One thousand two hundered thirty putative genes are colsely located with these SNPs. Among the putative genes, FZD4 results in light or silver coat (http://www.espcr.org/micemut/) and serves as a receptor for Wnt signalling pathway and is among the most common markers of melanocyte lineages (Yamada et al., 2013). MAP2K1 may inhibits the action of MITF and results in pigmentation disorder besides its role in melanosomes trafficking (Baxter & Pavan, 2013). ASIP and MC1R involve in eumelanin/pheomelanin switch, OCA2 involves in melanosome biogenesis and RS1 results in tiny patches of depigmentation in RPE (http://www.espcr.org/micemut/). TRPM7 results in pale colour and is associated with melanin synthesis (McNeill et al., 2007). ERBB4 may associate with melanogenesis (Zhao et al., 1998; Prickett et al., 2009) and Wnt7B is reported as pigmentation gene (Trantow et al., 2010). MED1 involves in melanogenesis (http://www.espcr.org/micemut/) and CREB binds and activates the MITF promoter through the cyclic adenosine monophosphate (cAMP) (Saha et al., 2006). TP53 activates tyrosinase and tyrosinase-related protein-1 (TRP-1) biosynthesis (Nylander, et al., 2000) and DBH involves in hydroxylation of dopamine (Slominski et al., 2004). DRD2 is known to involve in melanin pathway (Lao et al., 2007) and it darkenes Agouti color (http://www.espcr.org/micemut/). TRPM1 regulates differentiation and proliferation of melanocyte (Lu et al., 2010). Melanophores are sensitive to reduction in TRPM7 level, implicating the role of this gene in melanogenesis (McNeill et al., 2007). The cappuccino (CNO) may involve in organelle biogenesis associated with melanosomes (Huang et al., 2012) and Zeb2 involves in regulation of melanocytes development and differentiation (Denecker et al., 2014). Rab11b (Tarafder et al., 2013), Rab11a (Lapierre et al., 2001) and Rab1 (Hume et al., 2001) may involve in melanosomes trafficking. Oca2 involves in pigmentation (Sitaram et al., 2009) and Eda may involve in skin appendage development (Cui et al., 2007). CDH3 may involve in hair follicle morphogenesis (Jamora et al., 2003) and BCL2 may inhibit apoptosis of the melanocytes (Müller-Röver et al., 1999). The list of ten SNPs that show the most significant association is presented in Table 8.5.

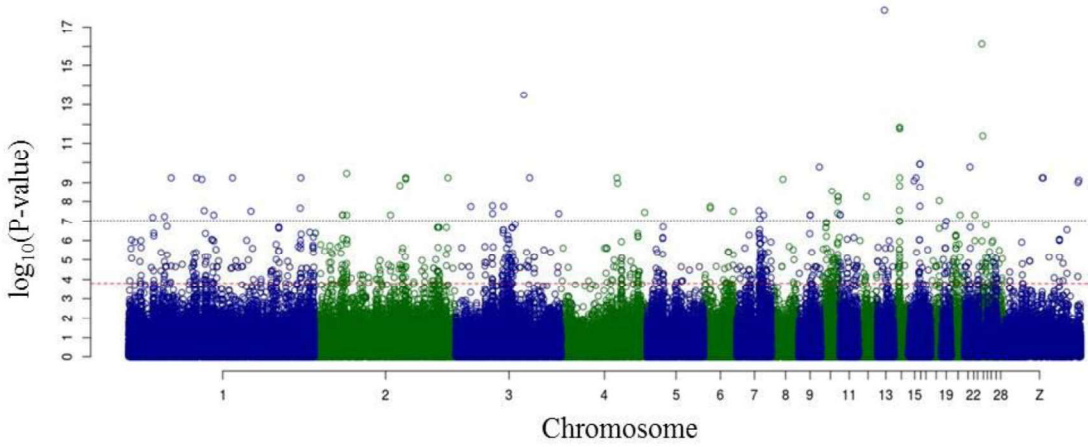


Figure 8.5 A single marker genome-wide association study of self-black hens versus red cocks in Ethiopian village chickens. The x-axis represents the chromosomal position of each SNP and the y-axis indicates –log10 P-value. The dotted red line represents FDR corrected P-value, while the black dotted line represents Bonferonni P-value corrected for multiple testing.

Table 8.5 The most significant ten SNPs associated with switch between red and black plumage in Ethiopian village chickens.

Chr	Pos	P-value (P1df)	qvalue	Associated genes
13	7497516	1.38E-18	7.48E-13	IL12B, FBXO38
24	1427085	7.43E-17	2.01E-11	TMEM45B, APLP2, ENSGALG00000024296
3	71206117	3.20E-14	5.76E-09	ENSGALG00000022967
14	5683794	1.41E-12	1.43E-07	
14	5726227	1.41E-12	1.43E-07	SOX8, LMF1
14	5676470	1.78E-12	1.43E-07	
14	5839933	1.86E-12	1.43E-07	LMF1
24	2481874	4.24E-12	2.87E-07	IGSF9B, JAM3, NCAPD3
15	11873941	1.21E-10	6.54E-06	
15	11875354	1.21E-10	6.54E-06	

White versus red plumage

One hundred and ninety eight SNPs exceed the FDR corrected P-value (Figure 8.6). Particularly, GGA2 & 6 show sharp association peaks indicating a mutation that underie a siwtch between red and white plumage. The commonly known TYR gene was mapped within ~ 8.6Kb downstream of an oultier SNP and SLC24A5 was also mapped by this analysis. Among others MYO5A involves in distribution and trafficking of melanosomes (Rodriguez & Cheney, 2002) and TP53 activates tyrosinase and tyrosinase-related protein-1 (TRP-1) biosynthesis (Nylander, et al., 2000). MAP2K1 may inhibits the action of MITF and results in pigmentation disorder besides its role in melanosomes trafficking (Baxter & Pavan, 2013). RAB27A involve in trnasportation of melanocytes and Notch1 may involve in graying of a plumage (http://www.espcr.org/micemut/) due to its role in cell growth, surivial and differentiation (Liu et al., 2006). POLG, POLH and TRPM7 have a systemic effect in melanogenesis (McNeill et al., 2007). CREB binds and activates MITF promoter through the cyclic adenosine monophosphate (cAMP) (Saha et al., 2006) and FGF7 mediates the transfer of melanocytes to keratinocytes (Cardinali et al., 2007). FGF18 is expressed in hair follicles (Kawano et al., 2005) and ADAMTSL3, ADAMTS17, ADAMTS7, ADAMTSL2 may involve in pigmentation as ADAMTS 20 does (Bennett and Lamoreux, 2003). GPC3 may cause a spotted pigmentation and ITGB1 may make a patchy hypopigmentation (http://www.espcr.org/micemut/). MAPK regulates the function of MITF at post-transcriptional level (Ebanks et al., 2009) and DBH involves in hydroxylation of dopamine (Slominski et al., 2004). TRPM1 regulates differentiation and proliferation of melanocyte (Lu et al., 2010). Melanophores are sensitive to reduction in TRPM7 level, implicating the role of this gene in melanogenesis (McNeill et al., 2007). Vps33b may involve in clustering and fusion of pigment granules (Gissen et al., 2005) and Rab11a (Lapierre et al., 2001) and Rab1 (Hume et al., 2001) may involve in melanosomes trafficking. BCL2 may inhibit apoptosis of the melanocytes (Müller-Röver et al., 1999) and GRB2 activates the Ras guanine nucleotide-binding protein by tyrosine kinases (Rozakis-Adcock et al., 1992). Ten SNPs that show the most significant association are presented in Table 8.6.

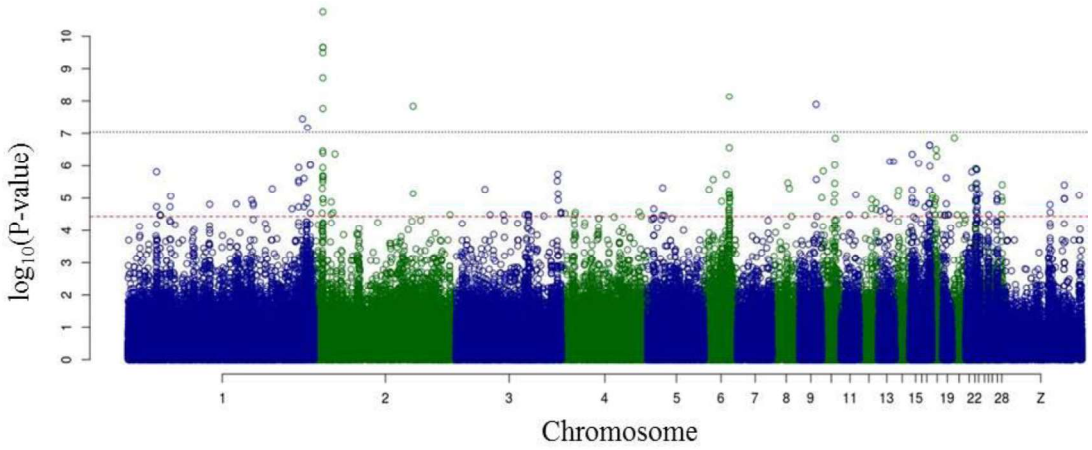


Figure 8.6 A single marker genome-wide association study of red cocks versus self-white hens and cocks in Ethiopian village chickens. The x-axis represents the chromosomal position of each SNP and the y-axis indicates –log10 P-value. The dotted red line represents FDR corrected P-value, while the black dotted line represents Bonferonni P-value corrected for multiple testing.

Table 8.6 The most significant ten SNPs associated with switch between red and white plumage in Ethiopian village chickens.

Chr	Pos	P-value (P1df)	qvalue	Associated genes
2	9018069	1.71E-11	9.24E-06	PTPRN2
2	9037808	2.14E-10	3.86E-05	PTPRN2
2	9038922	2.14E-10	3.86E-05	PTPRN2
2	9014404	3.09E-10	4.18E-05	PTPRN2
2	9003298	1.88E-09	0.000203	PTPRN2
6	26364260	7.05E-09	0.000636	VTI1A
9	17553016	1.25E-08	0.000967	TBL1XR1
2	103694921	1.44E-08	0.000974	
2	9068691	1.69E-08	0.001017	PTPRN2
1	182953790	3.44E-08	0.001862	FAR-2

Shank colour

We perform a GWAS on dark (black, green and slate blue) versus brightly coloured shanks (white and yellow), darkly coloured shanks versus yellow and yellow versus white.

Dark versus vividly coloured shanks

We analysed wild (dark coloured: black, slaty blue and green) shank colour versus derived variants (white and yellow) and a number of association peaks were found on the autosomes and the Z chromosome (Figure 8.7). However, sharp peaks were observed on GGA2, 3, 5, 10, 11, 13 & 14. Among three hundred thirty six putative genes mapped, CDKN2A and CDKN2B involve in melanoma (Soto et al., 2005). MED1, RB1, TYR, MC1R, DPH1, GPC3 and NF1 are pigmentation genes and Notch1 may involve in greying of plumage (http://www.espcr.org/micemut/) attributable to its effect on cell growth, survival and differentiation (Liu et al., 2006). Repressive cross-regulatory interactions between Sox2 and MITF involve in melanocytes development (Adameyko et al., 2012) and Wnt7B is reported as a pigmentation gene (Trantow et al., 2010). The most significant ten SNPs are presented in Table 8.7. Three hundred seven genes were commonly mapped by GWAS analyses performed on wild versus derived variants of shank colour and between wild and yellow shank colour. However, a GWAS performed on white versus yellow shank couldn't produce common genes with the two former analyses.

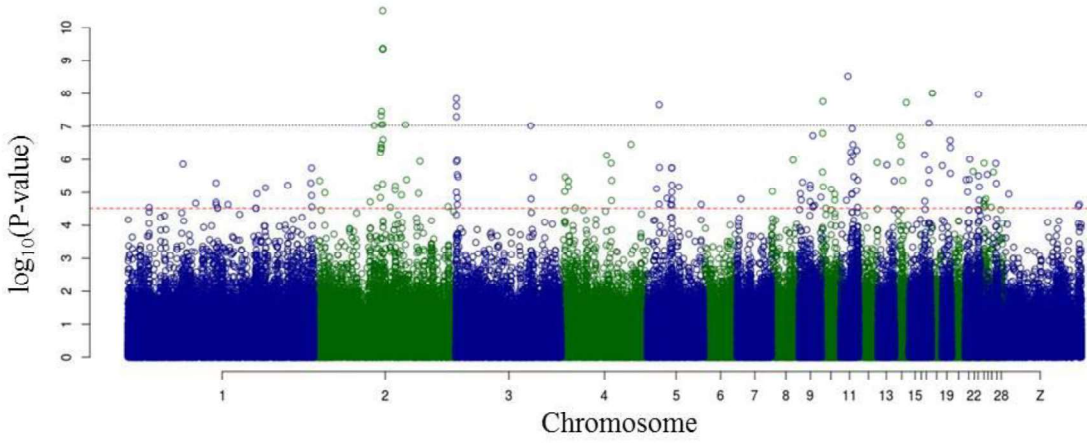


Figure 8.7 A single marker genome-wide association study of wild versus derived shank colour variants in Ethiopian village chickens. The x-axis represents the chromosomal position of each SNP and the y-axis indicates –log10 P-value. The dotted red line represents FDR corrected P-value, while the black dotted line represents Bonferonni P-value corrected for multiple testing.

Table 8.7 The most significant ten SNPs associated with a switch from wild to derived shank colour variants in Ethiopian village chickens.

Chr	Pos	P-value (P1df)	qvalue	Associated genes
2	72006740	3.05E-11	1.65E-05	
2	72037382	4.36E-10	4.72E-05	
2	72047880	4.36E-10	4.72E-05	
2	72084096	4.36E-10	4.72E-05	
2	72123591	4.36E-10	4.72E-05	
11	7960385	2.98E-09	0.000269	ENSGALG00000004489, CCNE1, URI1
18	958775	9.69E-09	0.000632	
18	960088	9.69E-09	0.000632	
23	2985528	1.05E-08	0.000632	
3	603436	1.42E-08	0.000771	

Dark versus yellow shank

Yellow shank is the most commonly observed variant in Horro and Jarso chickens (Desta et al., 2013). Two hundred and thirty two SNPs crossed the FDR corrected P-value (Figure 8.8). A genome-wide association study performed for yellow shank versus wild variants identifies seven hundred ninety six putative genes that may underlie the variation observed. Among the putative loci mapped, CDKN2A and CDKN2B involve in melanoma (Soto et al., 2005) and MED1 is reported as a pigmentation gene (http://www.espcr.org/micemut/). NF1 involves in melanosomes localization (Arun et al., 2013) and Notch1 may involve in greying of a plumage (http://www.espcr.org/micemut/) following its effect on cell growth, survival and differentiation (Liu et al., 2006). GPC3, FOXN1, UNC119 and MC1R are colour genes and PTS has a systemic effect in melanogenesis (http://www.espcr.org/micemut/). Wnt7B has been reported as a pigmentation gene (Trantow et al., 2010) and Spns2 may also involve in pigmentation (Chen et al., 2014). The most significant ten SNPs from this analysis are presented in Table 8.8.

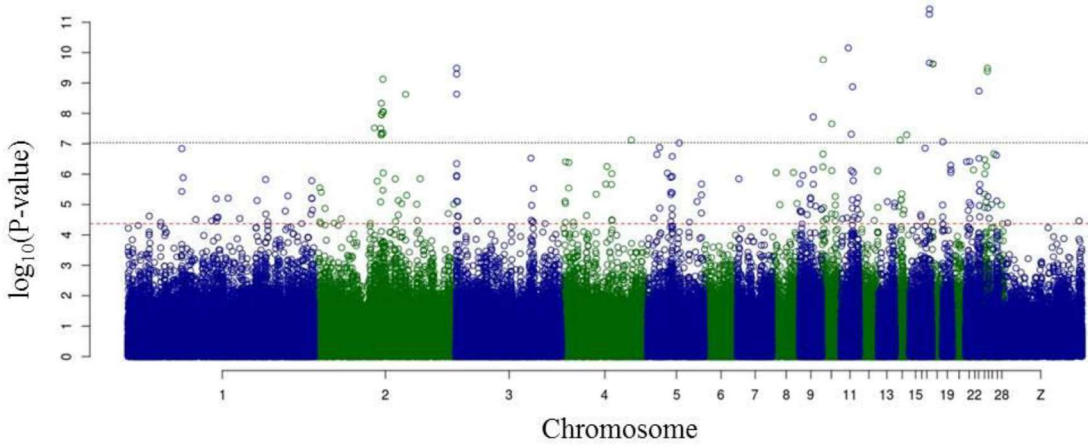


Figure 8.8 A single marker genome-wide association study of wild versus yellow shank colour variants in Ethiopian village chickens. The x-axis represents the chromosomal position of each SNP and the y-axis indicates –log10 P-value. The dotted red line represents FDR corrected P-value, while the black dotted line represents Bonferonni P-value corrected for multiple testing.

Table 8.8 The most significant ten SNPs associated with a switch from dark to yellow shank in Ethiopian village chickens.

Chr	Pos	P-value (P1df)	qvalue	Associated genes
17	7929540	3.76E-12	1.49E-06	INPP5E, PMPCA, SDCCAG3, SNAPC4, CARD9, DNLZ, GPSM1
17	7745104	5.49E-12	1.49E-06	PPP1R26, MRPS2, SNORA17, FAM69B, AGPAT2, EGFL7, gga-mir-126
11	7960385	7.08E-11	1.28E-05	ENSGALG00000004489, CCNE1, URI1
10	1586459	1.73E-10	1.82E-05	HCN4
17	7784420	2.19E-10	1.82E-05	AGPAT2, EGFL7, gga-mir-126, NOTCH1
18	958775	2.36E-10	1.82E-05	
18	960088	2.36E-10	1.82E-05	
3	603436	3.29E-10	1.99E-05	
24	6143938	3.30E-10	1.99E-05	ENSGALG00000007846, PTS, BCO2, IL18, SDHD, C11orf57, PIH1D2, DLAT, DIXDC1
24	6136244	4.14E-10	2.24E-05	ENSGALG00000007846, PTS, BCO2, IL18, SDHD, C11orf57, PIH1D2, DLAT

Yellow versus white shank

We perform association mapping for yellow versus white shank and we found a sharp peak on GGA24. Twenty one SNPs showing a significant association passed the FDR corrected p-value on GGA1–3, 7, 9, 10, 17, 23 and 24 (Figure 8.9). Twenty four putative genes that may underlie a switch between white and yellow shank are located close to the SNPs showing a significant association. The putative genes among others include the BCO2 (beta-carotene oxygenase 2) on GGA24. This locus was mapped by Rubin et al. (2010) and they described it as a domestication gene.

TCF is known to be a colour gene and PTS has systemic effect in melanogenesis (http://www.espcr.org/micemut/). Moreover, sever deficiency in PTS is known to cause hyperphenylalaninemia (elevated level of phenylalanine) and monoamine neurotransmitter deficiency (Blau et al., 2000). Phenylalanine is the precursor of tyrosine; a molecule that is used in pigment synthesis. IL-18 may involve in prevention of apoptosis in keratinocytes (Schwarz et al., 2006). Ten SNPs that show the most significant association are presented in Table 8.9.

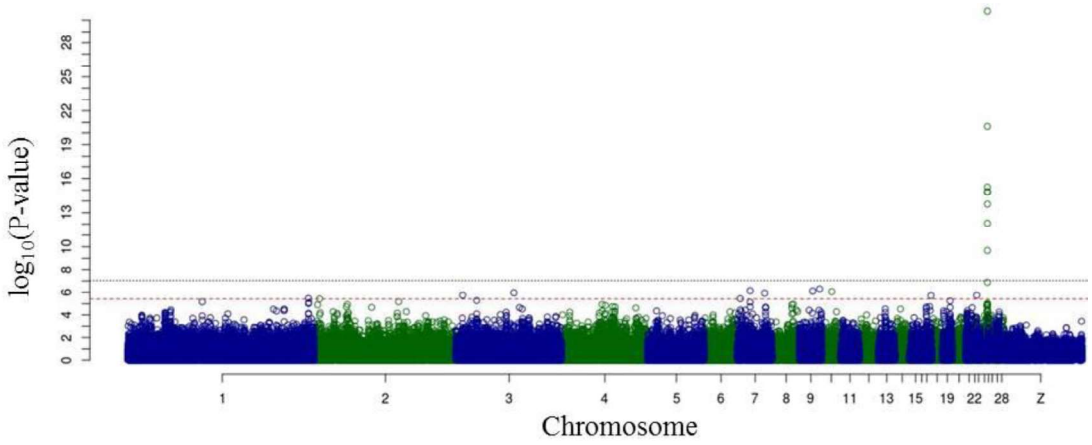


Figure 8.9 Manhattan plot of genome-wide -log10 (P-value) for white and yellow shank colour in Horro and Jarso chickens. The horizontal dashed black line represents the Bonferroni significance threshold while the dotted red line indicates the FDR corrected p-value. The x-axis is the position of each SNP on chicken autosomes and the y-axis is the –log10 P-value.

Table 8.9 The most significant ten SNPs associated with a switch from white to yellow shank in Ethiopian village chickens.

Chr	Pos	P-value (P1df)	qvalue	Associated genes
24	6143938	1.68E-31	9.09E-26	ENSGALG00000007846, PTS, BCO2, IL18, SDHD, C11orf57, PIH1D2, DLAT, DIXDC1
24	6136244	2.20E-21	5.95E-16	ENSGALG00000007846, PTS, BCO2, IL18, SDHD, C11orf57, PIH1D2, DLAT
24	6159205	5.43E-16	9.80E-11	ENSGALG00000007846, PTS, BCO2, IL18, SDHD, C11orf57, PIH1D2, DLAT, DIXDC1
24	6160905	1.34E-15	1.65E-10	ENSGALG00000007846, PTS, BCO2, IL18, SDHD, C11orf57, PIH1D2, DLAT, DIXDC1
24	6166964	1.52E-15	1.65E-10	ENSGALG00000007846, PTS, BCO2, IL18, SDHD, C11orf57, PIH1D2, DLAT, DIXDC1
24	6165853	1.83E-14	1.65E-09	ENSGALG00000007846, PTS, BCO2, IL18, SDHD, C11orf57, PIH1D2, DLAT, DIXDC1
24	6163838	8.47E-13	6.55E-08	ENSGALG00000007846, PTS, BCO2, IL18, SDHD, C11orf57, PIH1D2, DLAT, DIXDC1
24	6162692	2.07E-10	1.40E-05	ENSGALG00000007846, PTS, BCO2, IL18, SDHD, C11orf57, PIH1D2, DLAT, DIXDC1
24	6182044	1.32E-07	0.007951	PTS, BCO2, IL18, SDHD, C11orf57, PIH1D2, DLAT, DIXDC1, HSPB2
9	21220767	5.04E-07	0.027272	

Crested versus plain head

A GWAS performed using the qtscore function of GenABEL identifies a significant association for a switch between crested (Cr) and plain head in hens in five hundred and ninety SNPs located both on the autosomes and the Z chromosome (Figure 8.10). Among these, ten SNPs that show the most significant association with crest mutation are presented in Table 8.10. Crest has been known to show incomplete dominant autosomal mutation (Somes, 1990), however, we found a large number of Z chromosome SNPs associating with this mutation and crest is also reported to show sexual dimorphism (Wang et al., 2012; Desta et al., 2013), which may implicate the character of a sex-influenced trait. Eight hundred thirty seven putative genes are closely linked with SNPs showing a significant association. Among the putative genes, LAPTM4B and MATN2 are known to involve in transmembrane development (Liu et al., 2009) and extra cellular matrix assembly (Mátés et al., 2002 and the references therein) respectively. CSMD1 may involve in tumour suppression (Toomes et al., 2003) and SCML2 may maintain transcriptionally repressive state of homeotic genes (Montini et al., 1999), i.e., genes regulating the development of anatomical structures (Hirth et al., 1998). REPS2 may involve in growth factor signalling by its impact on the Ral signalling pathway (Ikeda et al., 1998). NRG1 may induce growth and differentiation in epithelial, glial, neuronal, and skeletal muscle cells (Zhao et al., 1998) and ZBTB16 involves in limb and skeletal patterning (Wasim et al., 2010). HpS4 may involve in alteration of cytoskeletal elements (Wei, 2006) and Gli3 may regulate craniofacial development (Vortkamp et al., 1991). The homobox containing gene en1 may involve in morphogenesis (Logan et al., 1989) and Brwd may regulate cytoskeletal organization and cell morphology (Bai et al., 2011). MED1 may involve in tissue development and differentiation by bridging transcription activators with RNA polymerase II (Jiang et al., 2010). Our GWAS analysis indicates that the genetic control of crest in chicken may be under the control of a polygenic effect.

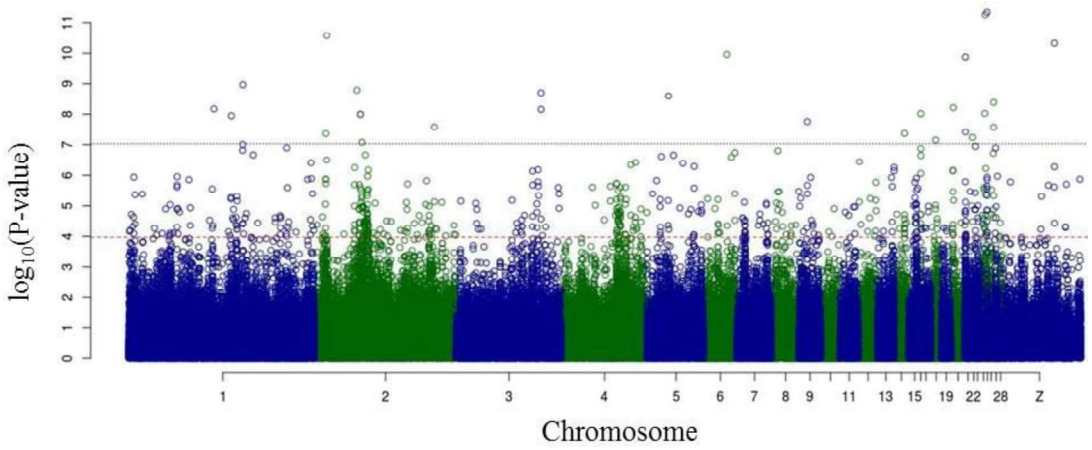


Figure 8.10 A single marker genome-wide association study of crested versus plain head in hens. The x-axis represents the chromosomal position of each SNP and the y-axis indicates –log10 P-value. The dotted red line represents FDR corrected P-value, while the dotted black line represents Bonferonni corrected P-value.

Table 8.10 The most significant ten SNPs associating with variation underlying crest phenotype.

Chr	Pos	P-value (P1df)	qvalue	Associated genes
25	460105	4.65E-12	1.52E-06	PEAR1, NTRK1, INSRR, ENSGALG00000026070, SH2D2A, PRCC, ENSGALG00000022570, MRPL24, RRNAD1, CRABP2, ENSGALG00000017589, BCAN, NES, ENSGALG00000029025
24	4938085	5.60E-12	1.52E-06	ENSGALG00000007032
2	12790364	2.69E-11	4.84E-06	FZD8, GJD4, CCNY
Z	53966263	4.75E-11	6.43E-06	NRG1
6	24409812	1.09E-10	1.18E-05	ENSGALG00000027136
21	791177	1.31E-10	1.19E-05	AJAP1
1	120087458	1.07E-09	8.28E-05	SCML2
2	44506894	1.62E-09	0.000109	
3	89227837	1.98E-09	0.000119	CSMD1
5	22574044	2.54E-09	0.000137	C11orf49, LRP4

Comb types

We perform a GWAS among eight comb shape variants (buttercup, cushion, duplex, pea, rose, single, strawberry and walnut) and single versus rose comb.

Comb shape variants

Mapping of comb variants was performed in cocks only due to a reduced comb size in hen. Forty nine SNPs crossed the FDR corrected P-value (Figure 8.11). Ninety nine putative genes are located close to SNPs showing a significant association. A GWAS performed using the qtscore function shows a clearly evident association peaks on GGA5, 7, 19 and the Z chromosome. Functional annotation analysis made using DAVID yielded nine clusters of which one has enrichment score of 1.62. The clusters among others involve in extracellular structure organization, muscle tissue/ organ development, striated muscle differentiation, organelle lumen etc. Three pathways: ErbB signalling pathway, beta-Alanine metabolism and tight junction were identified by DAVID. Among the putative genes Shc4, NRG1 and Pik3r1 involve in ErbB signalling pathway – a pathway that involves in embryogenesis (Alroy & Yarden, 1997). Gad1 and HADHA involve in beta-Alanine metabolism and ASH1L, CTNNA3 and ZAK involve in tight junction pathway. Ten SNPs showing the most significant association are presented in Table 8.11.

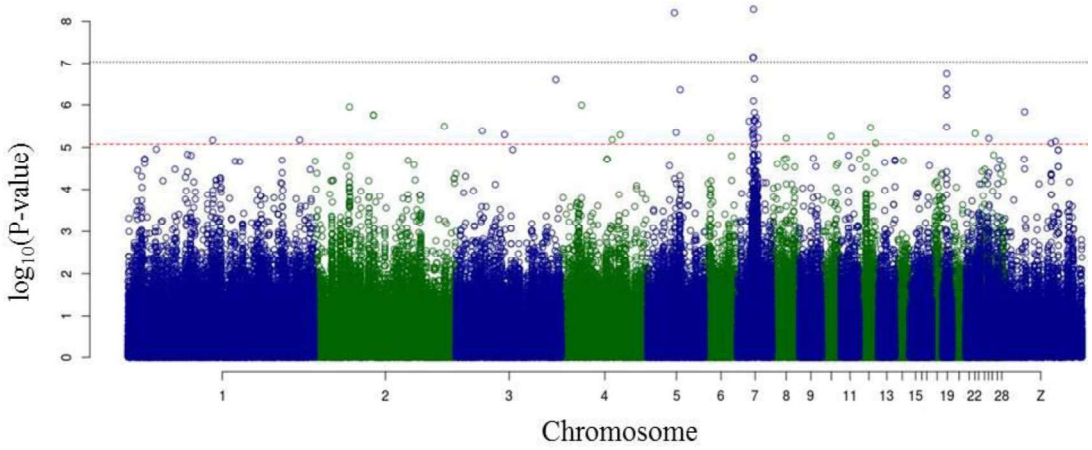


Figure 8.11 A single marker genome-wide association study of comb shape variants in cocks. The x-axis represents the chromosomal position of each SNP and the y-axis indicates –log10 P-value. The dotted red line represents FDR corrected P-value, while the dotted black line represents Bonferonni corrected P-value.

Table 8.11 The most significant ten SNPs associating with variation underlying comb variants commonly observed in Ethiopian village chickens.

Chr	Pos	P1df	qvalue	Associated genes
7	16799579	5.31E-09	0.001733	CDCA7, ENSGALG00000009344
5	28404929	6.40E-09	0.001733	GPHN
7	16803667	7.22E-08	0.009855	CDCA7, ENSGALG00000009344
7	16311522	7.28E-08	0.009855	CHN1, gga-mir-1570, CHRNA1, WIPF1
19	4313443	1.79E-07	0.019057	SRRM3, MDH2, STYXL1, TMEM120A, POR, TAF15
7	17824041	2.36E-07	0.019057	MYO3B
3	104555890	2.46E-07	0.019057	KIF3C, RAB10, GAREML, HADHA, ENSGALG00000028395
19	4225530	4.03E-07	0.025198	ENSGALG00000001876, UPK3B, ENSGALG00000001885, DTX2, SSC4D, YWHAG, HSPB1, SRRM3
5	34367364	4.19E-07	0.025198	NPAS3
19	4180338	5.90E-07	0.031913	PRKRIP1, ORAI2, LRWD1, ALKBH4, ENSGALG00000001861, ENSGALG00000001876, DTX2, ENSGALG00000001885, UPK3B, SSC4D

Single versus rose comb

A sharp association peak made by ~ 2.6k SNPs that associate with rose comb mutation and passed both the FDR and Bonferonni corrected P-value were found on GGA7 alone (Figure 8.12). Six hundred and nine putative genes were mapped in this region. However, it is not clear whether structural variation alone (Imsland et al., 2012) or polymorphism in the putative genes account for a rose comb mutation. A structural variant underlying the rose comb phenotype was mapped to this region (Imsland et al., 2012) and this region was also noted for its association with rose comb phenotype from GWAS (Wragg et al., 2012). Among the putative loci mapped, RNF20 of the Z chromosome involves in transcriptional activation of hox genes (Zhu et al., 2005) and hox genes are known to regulate patterns of anatomical development (Gellon & McGinnis, 1998). The sex chromosome gene RNF20 may contribute to dimorphism in comb size between hens and cocks through the action of hox genes. DLX3 as family of homeobox genes may involve in craniofacial patterning and morphogenesis (Merlo et al., 2000). LFNG may involve in mediating somite segmentation and patterning (Aulehla & Herrmann, 2004) and TGFB1 regulates the function other growth factors (Ignotz & Massague, 1986). IHH is known to regulate tissue patterning, skeletogenesis and cellular proliferation (see Shimoyama et al., 2007 and the references therein) and GRB2 activates the Ras guanine nucleotide-binding protein by tyrosine kinases (Rozakis-Adcock et al., 1992). Ten SNPs that show a highly significant association are presented in Table 8.12. There are a number of loci in 50K up and down stream of the outlier SNPs (see Table 8.12); therefore the sharp peak region mapped likely represent a gene-rich region.

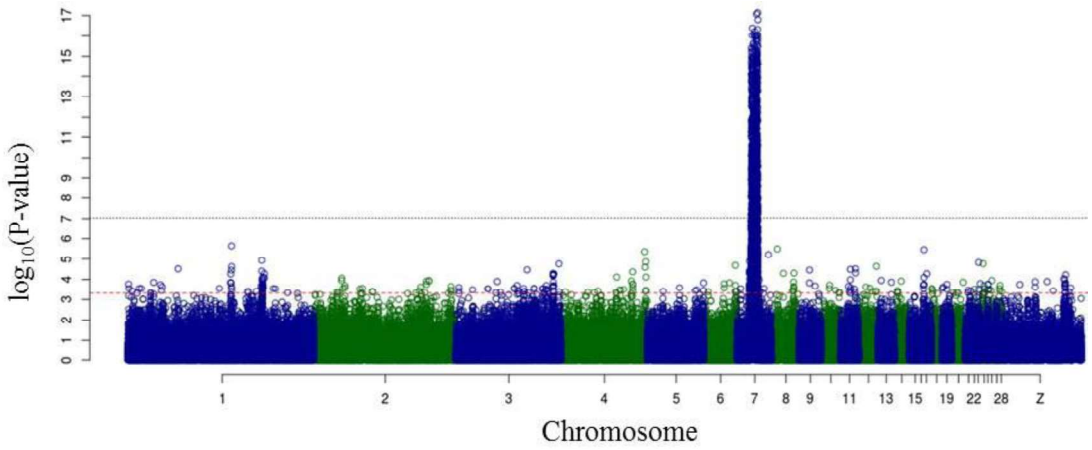


Figure 8.12 A single marker genome-wide association study of single versus rose comb in cocks. The x-axis represents the chromosomal position of each SNP and the y-axis indicates –log10 P-value. The dotted red line represents FDR corrected P-value, while the dotted black line represents Bonferonni corrected P-value.

Table 8.12 The most significant ten SNPs associated with switch from single to rose comb in Ethiopian village chickens.

Chr	Pos	P-value (P1df)	qvalue	Associated genes
7	21583378	7.27E-18	2.33E-12	STK11IP, OBSL1, INHA, CHPF, ENSGALG00000011233, TMEM198, ASIC4, GMPPA, ENSGALG00000011252, ENSGALG00000026721
7	20413623	8.60E-18	2.33E-12	KCNH7
7	21754266	1.61E-17	2.90E-12	GLB1L, STK16, ANKZF1, ATG9A, ABCB6, ZFAND2B, ENSGALG00000000433, IHH, FAM134A, SLC23A3, NHEJ1
7	16110177	4.10E-17	4.49E-12	ATF2, ATP5G3
7	21383967	4.71E-17	4.49E-12	PLA2R1, LY75, ENSGALG00000027744, ENSGALG00000011172
7	18512158	5.97E-17	4.49E-12	CERS6
7	17730249	8.42E-17	4.49E-12	ENSGALG00000021856, GAD67, SP5, MYO3B
7	21464697	9.44E-17	4.49E-12	LY75, ENSGALG00000011172, SLC4A3
7	15743377	9.46E-17	4.49E-12	MTX2, HOXD4, HOXD8, ENSGALG00000023420, gga-mir-6624, gga-mir-1713, gga-mir-10b
7	21260451	9.46E-17	4.49E-12	RBMS1, ITGB6

Lightly feathered shank

A ccfast analysis identifies association between three hundred and forty nine SNPs and causative genomic variants underlying the mutation of a lightly feathered shank

on most of the autosomes and the Z chromosome (Figure 8.13). One hundred sixty eight putative genes are located along the SNPs. Among the putative loci mapped, RBM19 may involve in embryo pre-implantation development and EDIL3 may involve in the regulation of vascular morphogenesis (Feng et al., 2014 and the references therein). MED1 may involve in tissue development and differentiation by bridging transcription activators with RNA polymerase II (Jiang et al., 2010). Among four loci that are located on GGA15 and showing a sharp association peak HIC2 up regulates Myc/Max pathway and down regulates Oxidative Stress/Notch signalling (regulates cell-fate determination during development) (Lv et al., 2014). ANAPC7 is required for vertebrates' mitotic cell cycle progression (Ho et al., 2013) and MED13L involves in Rb/E2F control of cell growth through complete suppression and cell cycle inhibition of target genes (Angus & Nevins, 2012). RBM19 is involved in preimplantation development – a critical developmental stage (Zhang et al., 2008). The most significant ten SNPs that show association with a lightly feathered shank mutation are presented in Table 8.13. Twelve functionally annotated clusters were identified by DAVID, two of these show enrichment score of 1.33 and 1.36. The clusters among others involve in egf-like domain, regulation of kinase and transferase activity, nucleotide binding, disulphide bond and localization of cells. However, no pathway has been detected.

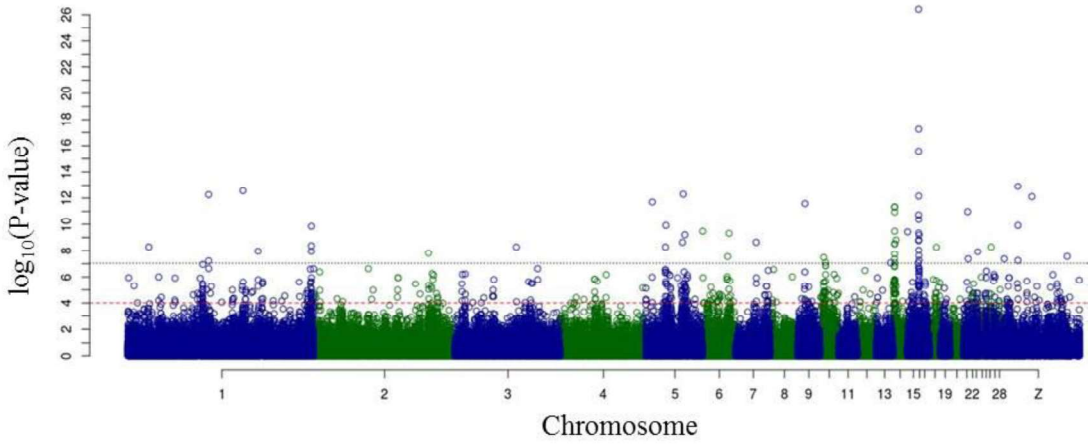


Figure 8.13 A single marker genome-wide association study of lightly feathered shank mutation. The x-axis represents the chromosomal position of each SNP and the y-axis indicates –log10 P-value. The dotted red line represents FDR corrected P-value, while the dotted black line represents Bonferonni corrected P-value.

Table 8.13 The most significant ten SNPs associated with a switch from plain to lightly feathered shank in Ethiopian village chickens.

Chr	Pos	P-value (Pc1df)	qvalue	Associated genes
15	11907142	3.15E-13	1.22E-07	
15	11763058	4.49E-13	1.22E-07	MED13L
15	11859242	1.39E-12	2.51E-07	
1	193034865	2.73E-12	3.69E-07	WNT11
5	40399721	3.01E-11	3.25E-06	
Z	16962607	3.77E-11	3.40E-06	gp130, ANKRD55
6	27818538	4.74E-11	3.66E-06	ATRNL1
Z	2495193	7.93E-11	5.36E-06	
2	126256102	2.11E-09	0.000127	NDUFAF6, PLEKHF2, C8orf37
15	11814356	2.80E-09	0.000152	

Discussion

Genome wide association study investigates the strength of relationship between genetic markers and causative variants that underlie a phenotypic variation. GWAS therefore links causative variants with phenotype of interest via closely linked genetic markers. GWAS identifies genetic variant that explain differences between contrasting phenotypes for example as in case and control studies. GWAS in outbred population has a special advantage over the classical quantitative genetics approach as a fine scale genetic map can be performed even in non-pedigreed populations using genetic markers information only (Emara and Kim, 2003). The use of high density genetic markers that are tightly linked with causal genomic variants improves the resolution power of a fine mapping and can be used to localize the underlying loci. GWAS combines linkage analysis with association test (Estus et al., 2013), which has made it a powerful approach.

A GWAS analysis that involves a large number of samples and that corrects for population stratification effect reduces the proportion of false positives. The use of

highly contrasting phenotypes and accuracy in phenotypic scoring improves the resolution power of an association mapping. True signal of GWAS is identified either directly when the genetic markers themselves making the causative variants or indirectly when the genetic markers are in sufficient linkage disequilibrium (LD) with causal genetic polymorphisms. Extensive LD (as in recombination coldspot regions), familial relationship among cohorts, small sample size and inaccurate scoring of phenotypes may lead to spurious association (false positives). Moreover, all associated genomic variants may not necessarily underlie a genetic variation (Keller et al., 2010). Therefore, a subsequent validation study is required to functionally annotate the putative loci identified. When a large number of loci make the genetic basis of a trait effect size (the proportion of the variance explained by a locus) will reduce proportionately, this makes GWAS to lose its power of detecting underlying causative variants. Genetic heterogeneity as in genetically highly diverse village chicken populations reduces the power of association mapping, therefore to counteract with this problem, a large sample size is required in GWAS of such populations.

To date large number of loci that form the genetic basis of threshold traits (Dorshorst et al., 2010; Wragg et al., 2012; Siwek et al., 2013) have been mapped in chicken using a GWAS approach. However, most of the GWAS performed in chicken were concentrated on disease (Sironi et al., 2011; Connell et al., 2013; Li et al., 2013; Luo et al., 2013; Wolc et al., 2013; Luo et al., 2014) and production traits (Ankra-Badu et al., 2010; Gao et al., 2011; Gu et al., 2011; Liu et al., 2011; Xie et al., 2012; Xu et al., 2013). Polygenic traits are considerably influenced by environmental effect, which reduces the resolution of a genetic mapping. Mapping of polygenetic traits is further complicated when gene by environment interaction is considered. GWAS is basically a molecular survey as it involves in course mapping of a large region of a genome that harbours a causative variant. However, as the number of markers included in GWAS increases, a fine mapping that localizes genomic regions underlying a causative variant becomes more feasible.

Scoring of morphological traits in outbred village chickens is a challenging task (Wragg et al., 2012); despite this we map the genetic control of a number of traits displaying less ambiguous variants. However, due to unique expression pattern of

threshold traits, graded classification of some phenotypes can be performed with little doubt. Due to absence of strong anthropogenic effect, useful LD extends to short region; which indicates the importance of village chickens in fine mapping. However a large number of SNPs is required for fine mapping in outbred populations. Our GWAS also benefited from a large number of samples included in the analysis. GWAS in outbred populations also has an added advantage because it helps to identify the impact of natural selection as human driven selection is thought to be less intense in panmictic populations. Our analysis shows that although the traits analysed tend to show Mendelian mode of inheritance, large number of loci may form the genetic basis of the traits. For example, crested head thought to be a simple binary trait; however it may assume continuous distribution for its size. Besides polygenic nature of the traits analysed, epistasis and pleiotropic interaction may confound with additive and dominance effect of each locus. A considerable proportion of the traits analysed show sexual dimorphism, which reflects the feature of sex influenced and sex linked traits. For example, crested head is mostly observed in hen and self-black plumage is entirely limited to hen in the two chicken populations. Traits of these kind may considerably influenced by sex linked loci (Dorshorst & Ashwell, 2009) and hormones (Yu et al., 2004). Variation in comb size between cocks and hens may show the character of a sex influenced trait. Chickens are known to have a number of traits displaying sexual dimorphism.

A high diversity in plumage colour and pattern of domestic fowl has attracted a number of studies (Werret et al., 1959; Cole and Jeffers, 1963; Silversides and Crawford, 1990; Klungland & Vage, 2000; Kerje et al., 2003; Kerje et al., 2004; Gunnarsson et al., 2007; Liu et al., 2010; Hoque et al., 2013; Park et al., 2013). Despite a long list of studies few loci (DCT, EDN3, KIT, ASIP, MC1R, PEML17 and SLC45A2) have been known to underlie the genetic control of plumage pigmentation. Moreover, a simple mode of inheritance pattern was reported from experimental crossing involving different comb shape (Bateson, 1902; Punnett, 1923; Warren, 1949; Somes, 1991); however, this trait could be under the control of many loci than what has been thought. For example, the derived comb shape variants (e.g. rose and pea comb) each represented by different forms displaying a subtle variation. The genetic basis of threshold traits needs to be revisited using a high density genetic markers and whole genome resequencing data to advance our

knowledge on the genetic control of these traits. Epistasis and pleiotropic effects need to be considered while dealing with the genetic control of threshold traits, because these traits could be impacted by complex biological processes (Cole and Jeffers, 1963; Bitgood, 1999; Wright et al., 2010).

Moreover, it is also important to understand the impact of threshold traits on adaptive radiation of the domestic fowl (e.g. comb shape, earlobe colour, plumage colour) and production traits. Threshold traits are important resources to study the impact of domestication events and dispersion pattern in domestic fowl. For example, loci that underlie derived comb shape variants can be considered as domestication genes because extant junglefowl species are known to display single comb variant. However, human driven selection may to some extent confound with adaptive radiation, though artificial selection on morph traits is less intense (Desta et al., 2013). Threshold traits displaying a graded expression pattern are good resources in QTL mapping than metric polygenic traits showing a continuous variation. In the former a few genes contribute to a large proportion of the underlying genetic variation, which forms the basis of a QTL mapping study. Threshold traits may show a swift shift between morph variants when a gene product making a variant to express excels or below the required limit (Roff, 1998).

From traits that have been included in our GWAS analysis, loci that control morphological variation have been reported for plumage colour (Kerje et al., 2003; Dorshorst et al., 2010; Liu et al., 2010, Park et al., 2013), shank colour (Dorshorst et al., 2010; Siwek et al., 2013; Li et al., 2014), earlobe colour (Wragg et al., 2012), rose comb (Dorshorst et al., 2010; Wragg et al., 2012; Imsland et al., 2012), pea comb (Wright et al., 2009); crested head (Wang et al., 2012) and feathered shank (Dorshorst et al., 2010). However, we mapped a large number of putative loci that closely linked with SNPs showing a significant association with variants of the traits studied. We also further refined the genetic control of a lightly feathered shank mutation. We found a large number of genomic regions associating with crested head phenotype than what has been known (e.g. Wang et al., 2012). Our findings to some extent disagree with previous reports on the genetic control of a feathered shank (Dorshorst et al., 2010) and crested head (Wang et al., 2012), which need to be addressed by future studies. Our finding substantiates a sex-influenced character of

crested head, which also disagrees with previous finding that describe crest as an autosomal trait (Hurst, 1905; Davenport, 1906). Polygenic nature of the traits analysed may reduce the resolution power of our mapping, becuase these traits to some extent tend to show a continuous variation. Moreover, even in threshold traits, large number of loci but each having little additive effect may contribute to a considerable proportion of the total genetic variation, therefore their cumulative effect may have a substantial impact on a phenotype. Furthermore, due to a high genetic diversity of village chickens (e.g. Lyimo et al., 2014) even individual loci may take different forms, which then reduce the resolution power of a genetic mapping.

Our GWAS shows that there is a considerable overlap among loci underlying variation in pigmentation in the traits analysed for colour polymorphism, which implicate an extensive pleiotropic effect of pigmentation loci. Pleiotropic effect was even detected among pigmentation and other traits particularly for those loci involving in morphogenesis and tissue development (e.g. GRB2, ZBTB16). Moreover, a large numbers of loci may control genetic variation of the traits studied, which shows a wide range of epistatic interaction. Genomic analyses of this kind therefore need to be supported by a gene expression analysis to gain insight on the impact of the underlying loci. Targeted sequencing and candidate gene approach will provide further evidence down to the identification of a causative point mutation.

Conclusion

Our GWAS maps a large number of putative loci that underlie the genetic control of a range of Mendelian and threshold traits in outbred village chickens. This analysis shows the importance of village chickens in genetic mapping studies. We produce a long list of putative loci that form the genetic basis of pigmentation traits and we also map putative loci contributing for buttercup comb mutation. In line with village chickens breeding history, the traits mapped may have been largely shaped by natural selection, which then enables us to uncover the impact of natural selection on traits showing Mendelian pattern of inheritance.

References

ADAMEYKO, I., LALLEMEND, F., FURLAN, A., ZININ, N., ARANDA, S., KITAMBI, S. S., ... & ERNFORS, P. Sox2 and Mitf cross-regulatory interactions consolidate progenitor and melanocyte lineages in the cranial neural crest. Development, 2012, 139(2), 397-410.

ALROY, I., & YARDEN, Y. The ErbB signaling network in embryogenesis and oncogenesis: signal diversification through combinatorial ligand-receptor interactions. FEBS letters, 1997, 410(1), 83-86.

ANGUS, S. P., & NEVINS, J. R. A role for Mediator complex subunit MED13L in Rb/E2F-induced growth arrest. Oncogene, 2012, 31(44), 4709-4717.

ANKRA-BADU, G. A., BIHAN-DUVAL, E. L., MIGNON-GRASTEAU, S., PITEL, F., BEAUMONT, C., DUCLOS, M. J., ... & AGGREY, S. E. Mapping QTL for growth and shank traits in chickens divergently selected for high or low body weight. Animal Genetics, 2010, 41(4), 400-405.

ARUN, V., WORRELL, L., WILEY, J. C., KAPLAN, D. R., & GUHA, A. Neurofibromin interacts with the cytoplasmic Dynein Heavy Chain 1 in melanosomes of human melanocytes. FEBS Letters, 2013, 587(10), 1466-1473.

AULCHENKO, Y. S., RIPKE, S., ISAACS, A. & VAN DUIJN, C. M. GenABEL: an R package for genome-wide association analysis. Bioinformatics, 2007, 23(10), 1294–1296.

AULEHLA, A., & HERRMANN, B. G. Segmentation in vertebrates: clock and gradient finally joined. Genes & Development, 2004, 18(17), 2060-2067.

BAI, S. W., HERRERA-ABREU, M. T., ROHN, J. L., RACINE, V., TAJADURA, V., SURYAVANSHI, N., ... & RIDLEY, A. J. Identification and characterization of a set of conserved and new regulators of cytoskeletal organization, cell morphology and migration. BMC Biology, 2011, 9(1), 54.

BATESON, W. Experiments with poultry. Rep Evol Comm Roy Soc, 1902, 1, 87-124.

BAXTER, L. L. & PAVAN, W. J. The etiology and molecular genetics of human pigmentation disorders. Wiley Interdiscip Rev Dev Biol. 2013, 2, 379-392.

BENJAMINI, Y. & HOCHBERG, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological), 1995, 289-300.

BENNETT, D. C., & LAMOREUX, M. L. The color loci of mice–a genetic century. Pigment Cell Research, 2003, 16(4), 333-344.

BITGOOD, J. J. Linkage relationships of the Z-linked silver, slow feathering, and pop-eye loci. Poultry Science, 1999, 78, 1100-1101.

BLAU, N., SCHERER-OPPLIGER, T., BAUMER, A., RIEGEL, M., MATASOVIC, A., SCHINZEL, A., ... & THÖNY, B. Isolated central form of tetrahydrobiopterin deficiency associated with hemizygosity on chromosome 11q and a mutant allele of PTPS. Human Mutation, 2000, 16(1), 54-60.

CARDINALI, G., BOLASCO, G., ASPITE, N., LUCANIA, G., LOTTI, L. V., TORRISI, M. R., & PICARDO, M. Melanosome transfer promoted by keratinocyte growth factor in light and dark skin-derived keratinocytes. Journal of Investigative Dermatology, 2007, 128(3), 558-567.

CHEN, J., INGHAM, N., KELLY, J., JADEJA, S., GOULDING, D., PASS, J., ... & STEEL, K. P. Spinster Homolog 2 (Spns2) Deficiency Causes Early Onset Progressive Hearing Loss. PLoS Genetics, 2014, 10(10), e1004688.

COLE, R. K. & JEFFERS, T. K. Allelism of silver, gold, and imperfect albinism in the fowl. Nature, 1963, 200, 1238-1239.

CONNELL, S., MEADE, K. G., ALLAN, B., LLOYD, A. T., DOWNING, T., O'FARRELLY, C., & BRADLEY, D. G. Genome-wide association analysis of avian resistance to *Campylobacter jejuni* colonization identifies risk locus spanning the CDH13 gene. G3: Genes| Genomes| Genetics, 2013, 3(5), 881-890.

CRAWFORD, R. D. Rose comb and fertility in Silver Spangled Hamburgs. Poultry Science, 1971, 47, 867-869.

CUI, C. Y., KUNISADA, M., ESIBIZIONE, D., GRIVENNIKOV, S. I., PIAO, Y., NEDOSPASOV, S. A., & SCHLESSINGER, D. Lymphotoxin-β regulates periderm differentiation during embryonic skin development. Human Molecular Genetics, 2007, 16(21), 2583-2590.

DAVENPORT, C. B. Inheritance in poultry. Washington, D.C., Carnegie Institution of Washington. Publication, 1906, 52, 1-136.

DENECKER, G., VANDAMME, N., AKAY, Ö., KOLUDROVIC, D., TAMINAU, J., LEMEIRE, K., ... & BERX, G. Identification of a ZEB2-MITF-ZEB1 transcriptional network that controls melanogenesis and melanoma progression. Cell Death & Differentiation, 2014, 21(8), 1250-1261.

DESTA, T. T., DESSIE, T., BETTRIDGE, J., LYNCH, S. E., MELESE, K., COLLINS, M., CHRISTLEY, R. M., WIGLEY, P., KAISER, P., TERFA, Z., MWACHARO, J. M. & HANOTTE, O. Signature of artificial selection and ecological landscape on morphological structures of Ethiopian village chickens. Animal Genetic Resources, 2013, 52, 1-13.

DOMYAN, E. T., GUERNSEY, M. W., KRONENBERG, Z., KRISHNAN, S., BOISSY, R. E., VICKREY, A. I., ... & SHAPIRO, M. D. Epistatic and combinatorial effects of pigmentary gene mutations in the domestic pigeon. Current Biology, 2014, 24(4), 459-464.

DORSHORST, B. J. & ASHWELL, C. M. Genetic mapping of the sex-linked barring gene in the chicken. Poultry Science, 2009, 88(9), 1811-1817.

DORSHORST, B., OKIMOTO, R., & ASHWELL, C. Genomic regions associated with dermal hyperpigmentation, polydactyly and other morphological traits in the Silkie chicken. Journal of Heredity, 2010, 101(3), 339-350.

EBANKS, J. P., WICKETT, R. R., & BOISSY, R. E. Mechanisms regulating skin pigmentation: the rise and fall of complexion coloration. International Journal of Molecular Sciences, 2009, 10(9), 4066-4087.

EMARA, M. G., & KIM, H. Genetic markers and their application in poultry breeding. Poultry Science, 2003, 82(6), 952-957.

ESTUS, J. L., OF NEPHROPATHY, F. I., FARDO, D. W., & DIABETES RESEARCH GROUP. Combining genetic association study designs: a GWAS case study. Frontiers in Genetics, 2013, 4.

FENG, M. X., MA, M. Z., FU, Y., LI, J., WANG, T., XUE, F., ... & XIA, Q. Elevated autocrine EDIL3 protects hepatocellular carcinoma from anoikis through RGD-mediated integrin activation. Molecular Cancer, 2014, 13(1), 226.

FRANKEL, W. N., YANG, Y., MAHAFFEY, C. L., BEYER, B. J., & O'BRIEN, T. P. Szt2, a novel gene for seizure threshold in mice. Genes, Brain and Behavior, 2009, 8(5), 568-576.

GAO, Y., FENG, C. G., SONG, C., DU, Z. Q., DENG, X. M., LI, N., & HU, X. X. Mapping quantitative trait loci affecting chicken body size traits via genome scanning. Animal Genetics, 2011, 42(6), 670-674.

GELLON, G., & MCGINNIS, W. Shaping animal body plans in development and evolution by modulation of Hox expression patterns. BioEssays, 1998, 20(2), 116-125.

GERKEN, M., AFNAN, R., & DÖRL, J. Adaptive behaviour in chickens in relation to thermoregulation. Archiv fur Geflugelkunde, 2006, 70(5), 199-207.

GISSEN, P., JOHNSON, C. A., GENTLE, D., HURST, L. D., DOHERTY, A. J., O'KANE, C. J., ... & MAHER, E. R. Comparative evolutionary analysis of VPS33 homologues: genetic and functional insights. Human Molecular Genetics, 2005, 14(10), 1261-1270.

GU, X., FENG, C., MA, L., SONG, C., WANG, Y., DA, Y., ... & LI, N. Genome-wide association study of body weight in chicken F2 resource population. Plos One, 2011, 6(7), e21872.

GUNNARSSON, U., HELLSTRÖM, A. R., TIXIER-BOICHARD, M., MINVIELLE, F., BED'HOM, B., ITO, S. I., ... & ANDERSSON, L. Mutations in SLC45A2 cause plumage color variation in chicken and Japanese quail. Genetics, 2007, 175(2), 867-877.

HINDHAUGH, W. L. S. Some observations on fertility in White Wyandottes. Poultry J, 1932, 17, 555–560.

HIRTH, F., HARTMANN, B., & REICHERT, H. Homeotic gene action in embryonic brain development of Drosophila. Development, 1998, 125(9), 1579-1589.

HO, A. W., GARG, A. V., MONIN, L., SIMPSON-ABELSON, M. R., KINNER, L., & GAFFEN, S. L. The anaphase-promoting complex protein 5 (AnapC5) associates with A20 and inhibits IL-17-mediated signal transduction. Plos One, 2013, 8(7), e70168.

HOQUE, M. R., JIN, S., HEO, K. N., KANG, B. S., JO, C., & LEE, J. H. Investigation of MC1R SNPs and their relationships with plumage colors in Korean native chicken. Asian-Australasian Journal of Animal Sciences, 2013, 26(5), 625-629.

HOSODA, K., HAMMER, R. E., RICHARDSON, J. A., BAYNASH, A. G., CHEUNG, J. C., GIAID, A., & YANAGISAWA, M. Targeted and natural (piebald-lethal) mutations of endothelin-B receptor gene produce megacolon associated with spotted coat color in mice. Cell, 1994, 79(7), 1267-1276.

HUANG, D. W., SHERMAN, B. T. & LEMPICKI, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research, 2009a, 37(1), 1-13.

HUANG, D. W., SHERMAN, B. T. & LEMPICKI, R. A. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protocol, 2009b, 4(1), 44-57.

HUANG, Z. M., CHINEN, M., CHANG, P. J., XIE, T., ZHONG, L., DEMETRIOU, S., ... & WEI, M. L. Targeting protein-trafficking pathways alters melanoma treatment sensitivity. Proceedings of the National Academy of Sciences, 2012, 109(2), 553-558.

HUBBARD, J. K., UY, J. A. C., HAUBER, M. E., HOEKSTRA, H. E., & SAFRAN, R. J. Vertebrate pigmentation: from underlying genes to adaptive function. Trends in Genetics, 2010, 26(5), 231-239.

HUME, A. N., COLLINSON, L. M., RAPAK, A., GOMES, A. Q., HOPKINS, C. R., & SEABRA, M. C. Rab27a regulates the peripheral distribution of melanosomes in melanocytes. The Journal of Cell Biology, 2001, 152(4), 795-808.

HURST, C. C. Experiments with poultry. Reports to the Evolution Committee of the Royal Society II, 1905,131-154.

HUTT, F. B. A relationship between breed characteristics and poor reproduction in white Wyandotte fowls. American Naturalist, 1940, 74, 148-156.

IGNOTZ, R. A., & MASSAGUE, J. Transforming growth factor-beta stimulates the expression of fibronectin and collagen and their incorporation into the extracellular matrix. Journal of Biological Chemistry, 1986, 261(9), 4337-4345.

IKEDA, M., ISHIDA, O., HINOI, T., KISHIDA, S., & KIKUCHI, A. Identification and characterization of a novel protein interacting with Ral-binding protein 1, a putative effector protein of Ral. Journal of Biological Chemistry, 1998, 273(2), 814-821.

IMSLAND, F., FENG, C., BOIJE, H., BED'HOM, B., FILLON, V., DORSHORST, B., ... & ANDERSSON, L. The Rose-comb mutation in chickens constitutes a structural rearrangement causing both altered comb morphology and defective sperm motility. PLoS Genetics, 2012, 8(6), e1002775.

JAMORA, C., DASGUPTA, R., KOCIENIEWSKI, P., & FUCHS, E. Links between signal transduction, transcription and adhesion in epithelial bud development. Nature, 2003, 422(6929), 317-322.

JIANG, P., HU, Q., ITO, M., MEYER, S., WALTZ, S., KHAN, S., ... & ZHANG, X. Key roles for MED1 LxxLL motifs in pubertal mammary gland development and luminal-cell differentiation. Proceedings of the National Academy of Sciences, 2010, 107(15), 6765-6770.

KANETSKY, P. A., SWOYER, J., PANOSSIAN, S., HOLMES, R., GUERRY, D., & REBBECK, T. R. A polymorphism in the agouti signaling protein gene is associated with human pigmentation. The American Journal of Human Genetics, 2002, 70(3), 770-775.

KAWANO, M., KOMI-KURAMOCHI, A., ASADA, M., SUZUKI, M., OKI, J., JIANG, J., & IMAMURA, T. Comprehensive analysis of FGF and FGFR expression in skin: FGF18 is highly expressed in hair follicles and capable of inducing anagen from telogen stage hair follicles. Journal of Investigative Dermatology, 2005, 124(5), 877-885.

KELLER, B., MARTINI, S., SEDOR, J., & KRETZLER, M. Linking variants from genome-wide association analysis to function via transcriptional network analysis. In Seminars in nephrology (Vol. 30, No. 2, pp. 177-184). WB Saunders, 2010, March.

KERJE, S., LIND, J., SCHÜTZ, K., JENSEN, P., & ANDERSSON, L. Melanocortin 1-receptor (MC1R) mutations are associated with plumage colour in chicken. Animal Genetics, 2003, 34(4), 241-248.

KERJE, S., P. SHARMA, U. GUNNARSSON, H. KIM, S. BAGCHI et al. The Dominant white, Dun and Smoky color variants in chicken are associated with insertion/deletion polymorphisms in the PMEL17 gene. Genetics, 2004, 168, 1507–1518.

KLOMP, A. E., TEOFILO, K., LEGACKI, E., & WILLIAMS, D. S. Analysis of the linkage of MYRIP and MYO7A to melanosomes by RAB27A in retinal pigment epithelial cells. Cell Motility and the Cytoskeleton, 2007, 64(6), 474-487.

KLUNGLAND, H., & VAGE, D. I. Molecular genetics of pigmentation in domestic animals. Current Genomics, 2000, 1(3), 223-242.

KRANIS, A., GHEYAS, A. A., BOSCHIERO, C., TURNER, F., YU, L., SMITH, S., ... & BURT, D. W. Development of a high density 600K SNP genotyping array for chicken. BMC Genomics, 2013, 14(1), 59.

KUMMER, M. P., MARUYAMA, H., HUELSMANN, C., BACHES, S., WEGGEN, S., & KOO, E. H. Formation of Pmel17 amyloid is regulated by juxtamembrane metalloproteinase cleavage, and the resulting C-terminal fragment is a substrate for γ-secretase. Journal of Biological Chemistry, 2009, 284(4), 2296-2306.

LAO, O., DE GRUIJTER, J. M., VAN DUIJN, K., NAVARRO, A., & KAYSER, M. Signatures of positive selection in genes associated with human skin pigmentation as revealed from analyses of single nucleotide polymorphisms. Annals of Human Genetics, 2007, 71(3), 354-369.

LAPIERRE, L. A., KUMAR, R., HALES, C. M., NAVARRE, J., BHARTUR, S. G., BURNETTE, J. O., ... & GOLDENRING, J. R. Myosin vb is associated with plasma membrane recycling systems. Molecular Biology of the Cell, 2001, 12(6), 1843-1857.

LI, D. F., LIAN, L., QU, L. J., CHEN, Y. M., LIU, W. B., CHEN, S. R., ... & YANG, N. A genome-wide SNP scan reveals two loci associated with the chicken resistance to Marek's disease. Animal Genetics, 2013, 44(2), 217-222.

LI, G., LI, D., YANG, N., QU, L., HOU, Z., ZHENG, J., ... & CHEN, S. A genome-wide association study identifies novel single nucleotide polymorphisms associated with dermal shank pigmentation in chickens. Poultry Science, 2014, 93(12), 2983-2987.

LIU, W. B., CHEN, S. R., ZHENG, J. X., QU, L. J., XU, G. Y., & YANG, N. Developmental phenotypic-genotypic associations of tyrosinase and melanocortin 1 receptor genes with changing profiles in chicken plumage pigmentation. Poultry Science, 2010, 89(6), 1110-1114.

LIU, W., LI, D., LIU, J., CHEN, S., QU, L., ZHENG, J., ... & YANG, N. A genome-wide SNP scan reveals novel loci for egg production and quality traits in white leghorn and brown-egg dwarf layers. Plos One, 2011, 6(12), e28600.

LIU, X., XIONG, F., WEI, X., YANG, H., & ZHOU, R. LAPTM4B-35, a novel tetratransmembrane protein and its PPRP motif play critical roles in proliferation and metastatic potential of hepatocellular carcinoma cells. Cancer Science, 2009, 100(12), 2335-2340.

LIU, Z. J., XIAO, M., BALINT, K., SMALLEY, K. S., BRAFFORD, P., QIU, R., ... & HERLYN, M. Notch1 signaling promotes primary melanoma progression by activating mitogen-activated protein kinase/phosphatidylinositol 3-kinase-Akt pathways and up-regulating N-cadherin expression. Cancer Research, 2006, 66(8), 4182-4190.

LOGAN, C., WILLARD H. F., ROMMENS J. M. AND JOYNER, A. L. Chromosomal localization of the human homeobox-containing genes, EN1 and EN2. Genomics, 1989, 4, 206-209.

LU, S., SLOMINSKI, A., YANG, S. E., SHEEHAN, C., ROSS, J., & CARLSON, J. A. The correlation of TRPM1 (Melastatin) mRNA expression with microphthalmia-associated transcription factor (MITF) and other melanogenesis-related proteins in normal and pathological skin, hair follicles and melanocytic nevi. Journal of Cutaneous Pathology, 2010, 37(s1), 26-40.

LUO, C., QU, H., MA, J., WANG, J., HU, X., LI, N., & SHU, D. A genome-wide association study identifies major loci affecting the immune response against infectious bronchitis virus in chicken. Infection, Genetics and Evolution, 2014, 21, 351-358.

LUO, C., QU, H., MA, J., WANG, J., LI, C., YANG, C., ... & SHU, D. Genome-wide association study of antibody response to Newcastle disease virus in chicken. BMC Genetics, 2013, 14(1), 42.

LV, L., DENG, H., LI, Y., ZHANG, C., LIU, X., LIU, Q., ... & ZHU, J. The DNA methylation-regulated miR-193a-3p dictates the multi-chemoresistance of bladder cancer via repression of SRSF2/PLAU/HIC2 expression. Cell Death & Disease, 2014, 5(9), e1402.

LYIMO, C. M., WEIGEND, A., MSOFFE, P. L., EDING, H., SIMIANER, H. & WEIGEND, S. Global diversity and genetic contributions of chicken populations from African, Asian and European regions. Animal Genetics, 2014, 45(6), 836-848.

MACZKOWIAK, F., MATÉOS, S., WANG, E., ROCHE, D., HARLAND, R., & MONSORO-BURQ, A. H. The Pax3 and Pax7 paralogs cooperate in neural and neural crest patterning using distinct molecular mechanisms, in Xenopus laevis embryos. Developmental Biology, 2010, 340(2), 381-396.

MARTINEZ, M. C., OCHIISHI, T., MAJEWSKI, M., & KOSIK, K. S. Dual regulation of neuronal morphogenesis by a δ-catenin–cortactin complex and Rho. The Journal of Cell Biology, 2003, 162(1), 99-111.

MÁTÉS, L., KORPOS, É., DEÁK, F., LIU, Z., R BEIER, D., ASZÓDI, A., & KISS, I. Comparative analysis of the mouse and human genes (Matn2 and MATN2) for matrilin-2, a filament-forming protein widely distributed in extracellular matrices. Matrix Biology, 2002, 21(2), 163-174.

MCNEILL, M. S., PAULSEN, J., BONDE, G., BURNIGHT, E., HSU, M. Y., & CORNELL, R. A. Cell death of melanophores in zebrafish trpm7 mutant embryos depends on melanin synthesis. Journal of Investigative Dermatology, 2007, 127(8), 2020-2030.

MERLO, G. R., ZEREGA, B., PALEARI, L, TROMBINO, S., MANTERO, S., & LEVI, G. Multiple functions of Dlx genes. International Journal of Developmental Biology, 2000, 44(6), 619-626.

MONTINI, E., BUCHNER, G., SPALLUTO, C., ANDOLFI, G., CARUSO, A., DEN DUNNEN, J. T., ... & FRANCO, B. Identification of SCML2, a second human gene homologous to the Drosophila sex comb on midleg (Scm): A new gene cluster on Xp22. Genomics, 1999, 58(1), 65-72.

MÜLLER-RÖVER, S., ROSSITER, H., LINDNER, G., PETERS, E. M., KUPPER, T. S., & PAUS, R. Hair follicle apoptosis and Bcl-2. In Journal of Investigative Dermatology Symposium Proceedings (Vol. 4, No. 3, pp. 272-277). Nature Publishing Group, December 1999.

NYLANDER, K., BOURDON, J. C., BRAY, S. E., GIBBS, N. K., KAY, R., HART, I., & HALL, P. A. Transcriptional activation of tyrosinase and TRP-1 by p53 links UV irradiation to the protective tanning response. The Journal of Pathology, 2000, 190(1), 39-46.

PARK, M. N., CHOI, J., LEE, K. T., LEE, H. J., CHOI, B. H., KIM, H., ... & LEE, T. Genome-wide association study of chicken plumage pigmentation. Asian-Australasian Journal of Animal Sciences, 2013, 26(11), 1523-1528.

PERUZZI, D., ALUIGI, M., MANZOLI, L., BILLI, A. M., DI GIORGIO, F. P., MORLEO, M., MARTELLI, A. M. & COCCO, L. Molecular characterization of the human PLC beta1 gene. Biochim Biophys Acta, 2002, 1584(1), 46-54.

PRICKETT, T. D., AGRAWAL, N. S., WEI, X., YATES, K. E., LIN, J. C., WUNDERLICH, J. R., ... & SAMUELS, Y. Analysis of the tyrosine kinome in melanoma reveals recurrent mutations in ERBB4. Nature Genetics, 2009, 41(10), 1127-1132.

PUNNETT, R.C. Heredity in poultry. Macmillan and Co., London, 1923, 8vo.pp.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013, ISBN 3-900051-07-0. http://www.R-project.org/.

RICHARDS, S. A. The significance of changes in the temperature of the skin and body core of the chicken in the regulation of heat loss. The Journal of Physiology, 1971, 216(1), 1-10.

RODRIGUEZ, O. C., & CHENEY, R. E. Human myosin-Vc is a novel class V myosin expressed in epithelial cells. Journal of Cell Science, 2002, 115(5), 991-1004.

ROFF, D. A. Evolution of threshold traits: the balance between directional selection, drift and mutation. Heredity, 1998, 80(1), 25-32.

ROZAKIS-ADCOCK, M., MCGLADE, J., MBAMALU, G., PELICCI, G., DALY, R., LI, W., ... & PAWSON, T. Association of the Shc and Grb2/Sem5 SH2-containing proteins is implicated in activation of the Ras pathway by tyrosine kinases. Nature, 1992, 360(6405), 689-692.

RUBIN, C. J., ZODY, M. C., ERIKSSON, J., MEADOWS, J. R., SHERWOOD, E., WEBSTER, M. T., ... & ANDERSSON, L. Whole-genome resequencing reveals loci under selection during chicken domestication. Nature, 2010, 464(7288), 587-591.

SAHA, B., SINGH, S. K., SARKAR, C., BERA, R., RATHA, J., TOBIN, D. J., & BHADRA, R. Activation of the Mitf promoter by lipid-stimulated activation of p38-stress signalling to CREB. Pigment Cell Research, 2006, 19(6), 595-605.

SATOH, G. J. Z., & MISHIMA, Y. Tyrosine release and tyrosinase activity in amelanotic and melanotic malignant melanoma1. Journal of Investigative Dermatology, 1969, 52(1), 107-110.

SCHWARZ, A., MAEDA, A., STÄNDER, S., VAN STEEG, H., & SCHWARZ, T. IL-18 reduces ultraviolet radiation-induced DNA damage and thereby affects photoimmunosuppression. The Journal of Immunology, 2006, 176(5), 2896-2901.

SHIMOYAMA, A., WADA, M., IKEDA, F., HATA, K., MATSUBARA, T., NIFUJI, A., ... & YONEDA, T. Ihh/Gli2 signaling promotes osteoblast differentiation by regulating Runx2 expression and function. Molecular Biology of the Cell, 2007, 18(7), 2411-2418.

SILVERSIDES, F. G. & CRAWFORD, R. D. Genetic aspects of a new mutation (sal-s) to sex-linked imperfect albinism in chickens. Genetics Selection Evolution, 1990, 22, 447-455.

SIRONI, L., WILLIAMS, J. L., STELLA, A., MINOZZI, G., MORENO, A., RAMELLI, P., ... & MARIANI, P. Genomic study of the response of chicken to highly pathogenic avian influenza virus. In BMC proceedings (Vol. 5, No. Suppl 4, p. S25). BioMed Central Ltd, 2011, June.

SITARAM, A., PICCIRILLO, R., PALMISANO, I., HARPER, D. C., DELL'ANGELICA, E. C., SCHIAFFINO, M. V., & MARKS, M. S. Localization to mature melanosomes by virtue of cytoplasmic dileucine motifs is required for human OCA2 function. Molecular Biology of the Cell, 2009, 20(5), 1464-1477.

SIWEK, M., WRAGG, D., SŁAWIŃSKA, A., MALEK, M., HANOTTE, O., & MWACHARO, J. M. Insights into the genetic history of Green-legged Partridgelike fowl: mtDNA and genome-wide SNP analysis. Animal genetics, 2013, 44(5), 522-532.

SLOMINSKI, A., TOBIN, D. J., SHIBAHARA, S., & WORTSMAN, J. Melanin pigmentation in mammalian skin and its hormonal regulation. Physiological Reviews, 2004, 84(4), 1155-1228.

SOMES JR, R. G. Duplex comb in the chicken: a multi-allelic trait. Journal of Heredity, 1991, 82(2), 169-172.

SOMES JR, R. G. Mutations and major variants of plumage and skin in chickens. Developments in Animal and Veterinary Sciences (Netherlands), 1990.

SOTO, J. L., CABRERA, C. M., SERRANO, S., & LÓPEZ-NEVOT, M. Á. Mutation analysis of genes that control the G1/S cell cycle in melanoma: TP53, CDKN1A, CDKN2A, and CDKN2B. BMC Cancer, 2005, 5(1), 36.

TARAFDER, A. K., BOLASCO, G., CORREIA, M. S., PEREIRA, F. J., IANNONE, L., HUME, A. N., ... & SEABRA, M. C. Rab11b mediates melanin transfer between donor melanocytes and acceptor keratinocytes via coupled exo/endocytosis. Journal of Investigative Dermatology, 2013, 134(4), 1056-1066.

TOOMES, C., JACKSON, A., MAGUIRE, K., WOOD, J., GOLLIN, S., ISHWAD, C., ... & THAKKER, N. The presence of multiple regions of homozygous deletion at the CSMD1 locus in oral squamous cell carcinoma question the role of CSMD1 in head and neck carcinogenesis. Genes, Chromosomes and Cancer, 2003, 37(2), 132-140.

TRANTOW, C. M., CUFFY, T. L., FINGERT, J., KUEHN, M. H., & ANDERSON, M. Microarray analysis of iris gene expression in mice with mutations influencing pigmentation. Investigative Ophthalmology & Visual Science, 2010, iovs-10.

VORTKAMP, A., GESSLER, M. & GRZESCHIK K.H. GLI3 zincfinger gene interrupted by translocations in Greig syndrome families. Nature, 1991, 352, 539-540.

WANG, S., MA, Z., XU, X., WANG, Z., SUN, L., ZHOU, Y., ... & WANG, T. A role of Rab29 in the Integrity of the Trans-Golgi network and retrograde trafficking of Mannose-6-Phosphate receptor. Plos One, 2014, 9(5), e96242.

WANG, Y., GAO, Y., IMSLAND, F., GU, X., FENG, C., LIU, R., SONG, C., TIXIER-BOICHARD, M., GOURICHON, D., LI, Q., CHEN, K., LI, H., ANDERSSON, L., HU, X. & LI, N. The crest phenotype in chicken is associated with ectopic expression of HOXC8 in cranial skin. Plos One, 2012, 7(4), e34012.

WANG, Z., SHEN, D., PARSONS, D. W., BARDELLI, A., SAGER, J., SZABO, S., ... & VELCULESCU, V. E. Mutational analysis of the tyrosine phosphatome in colorectal cancers. Science, 2004, 304(5674), 1164-1166.

WARREN, D. C. Linkage relations of autosomal factors in the fowl. Genetics, 1949, 34(3), 333.

WASIM, M., CARLET, M., MANSHA, M., GREIL, R., PLONER, C., TROCKENBACHER, A., ... & KOFLER, R. PLZF/ZBTB16, a glucocorticoid response gene in acute lymphoblastic leukemia, interferes with glucocorticoid-induced apoptosis. The Journal of Steroid Biochemistry and Molecular Biology, 2010, 120(4), 218-227.

WEI, M. L. Hermansky–Pudlak syndrome: a disease of protein trafficking and organelle function. Pigment Cell Research, 2006, 19(1), 19-42.

WERRET, W. F., CANDY A. J., KING J. O. & SHEPPARD, P. M. Semialbino: a third sex-linked allelomorph of silver and gold in the fowl. Nature, 1959, 184, (Suppl. 7), 480.

WOLC, A., ARANGO, J., JANKOWSKI, T., SETTAR, P., FULTON, J. E., O'SULLIVAN, N. P., ... & DEKKERS, J. C. Genome-wide association study for Marek's disease mortality in layer chickens. Avian Diseases, 2013, 57(2s1), 395-400.

WRAGG, D., MWACHARO, J. M., ALCALDE, J. A., HOCKING, P. M. & HANOTTE, O. Analysis of genome-wide structure, diversity and fine mapping of Mendelian traits in traditional and village chickens. Heredity, 2012, 109, 6-18.

WRIGHT, D., BOIJE, H., MEADOWS, J. R., BED'HOM, B., GOURICHON, D., VIEAUD, A., ... & ANDERSSON, L. Copy number variation in intron 1 of SOX5 causes the Pea-comb phenotype in chickens. PLoS Genetics, 2009, 5(6), e1000512.

WRIGHT, D., RUBIN, C. J., MARTINEZ BARRIO, A., SCHÜTZ, K., KERJE, S., BRÄNDSTRÖM, H., ... & ANDERSSON, L. The genetic architecture of domestication in the chicken: effects of pleiotropy and linkage. Molecular Ecology, 2010, 19(23), 5140-5156.

XIE, L., LUO, C., ZHANG, C., ZHANG, R., TANG, J., NIE, Q., ... & ZHANG, X. Genome-wide association study identified a narrow chromosome 1 region associated with chicken growth traits. Plos One, 2012, 7(2), e30910.

XU, Z., NIE, Q. & ZHANG, X. Overview of genomic insights into chicken growth traits based on genome-wide association study and microRNA Regulation. Current Genomics, 2013, 14(2), 137.

YAMADA, T., HASEGAWA, S., INOUE, Y., DATE, Y., YAMAMOTO, N., MIZUTANI, H., ... & AKAMATSU, H. Wnt/β-catenin and Kit signaling sequentially regulate melanocyte stem cell differentiation in UVB-induced epidermal pigmentation. Journal of Investigative Dermatology, 2013, 133(12), 2753-2762.

YATSU, A., OHBAYASHI, N., TAMURA, K., & FUKUDA, M. Syntaxin-3 is required for melanosomal localization of Tyrp1 in melanocytes. Journal of Investigative Dermatology, 2013, 133(9), 2237-2246.

YU, M., YUE, Z., WU, P., WU, D. Y., MAYER, J. A., MEDINA, M., ... & CHUONG, C. M. The developmental biology of feather follicles. International Journal of Developmental Biology, 2004, 48, 181-192.

ZHANG, J., TOMASINI, A. J., & MAYER, A. N. RBM19 is essential for preimplantation development in the mouse. BMC Developmental Biology, 2008, 8(1), 115.

ZHAO, Y. Y., SAWYER, D. R., BALIGA, R. R., OPEL, D. J., HAN, X., MARCHIONNI, M. A., & KELLY, R. A. Neuregulins promote survival and growth of cardiac myocytes Persistence of ErbB2 and ErbB4 expression in neonatal and adult ventricular myocytes. Journal of Biological Chemistry, 1998, 273(17), 10261-10269.

ZHU, B., ZHENG, Y., PHAM, A. D., MANDAL, S. S., ERDJUMENT-BROMAGE, H., TEMPST, P., & REINBERG, D. Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. Molecular Cell, 2005, 20(4), 601-611.

ZUK, M., POPMA, S. L., & JOHNSEN, T. S. Male courtship displays, ornaments and female mate choice in captive red jungle fowl. Behaviour, 1995, 821-836.

Chapter 9

General Discussion and Concluding Remarks

Among food animals, chicken is the most widely used model species in evolutionary genetic studies. Chicken has a number of conspicuous morphological traits (comb, crest, earlobe, plumage, shank etc.) that show a Mendelian proportion of inheritance pattern, which has made chicken important species in Mendelian genetics. Moreover, the survival of domestic chicken progenitor in the Asian jungle has made chicken ideal species for comparative genomics and evolutionary genetic studies. Furthermore, chicken's high rate of reproduction and short generation interval make it appropriate species for genetic studies involving experimental crossing. Its long history of use in genetic studies makes the chicken to get a priority among livestock species to have its genome sequenced (Hillier et al., 2004).

Village chickens harbour a considerable level of extant genetic diversity found in domestic fowl. Studying the genetic basis of chicken diversity is a stepping stone to uncover extant intrapopulation genetic variation and interpopulations' genetic divergence, to construct their demographic history and to trace their origin. A genetic variation forms the basis of selective breeding and genetic improvement involving adaptation and economically important traits. A standing (functional) genetic variation can be used to infer status of a population and to make an informed decision to prioritize populations for conservation. Prior knowledge on population structure is inevitable to correct for population stratification effect in genome-wide association studies, otherwise this leads to a spurious association. Genetic diversity is indispensable to deal with unforeseen changes in consumers demands and to mitigate with climate change. Village chickens can be also used as a biological marker to reconstruct a prehistoric time dispersion pattern of human and an ancient social and trading network (Mwacharo et al., 2013).

Most of the studies conducted on chicken of the least developed world have been concentrated on scoring of morphological traits and description of their production environment. Cognizant of this fact, our study combines phenomic studies with genomic analysis using morphological markers and a high density SNP array.

However, the variation observed between the two chicken populations using morphological markers (Chapter 2 & 3) hasn't been as evident as the genetic divergence observed using genome-wide high density SNPs (Chapter 4). We found a considerably contrasting demographic history of the two village chicken populations and we validate the suitability of a high density SNP chip that developed using genomic information of commercial chickens (Kranis et al., 2013) to study the genetic structure of outbred village chickens (Chapter 4 & 5). Our finding supports the historical evidence that suggests two entry points of chickens in the east Africa region. A clear genetic divergence observed between Horro and Jarso chickens may partly associate with these putative entry points. Unlike other studies that reported absence of population substructure in village chickens (*e.g.* Muchadeyi et al., 2007; Osei-Amponsah et al., 2010; Dana, 2011; Lyimo et al., 2014), regardless of our analysis involving only two populations that are distantly located, we found a clearly defined substructure which could partly reflect the impact of a large number of markers.

Congruent with previous findings (Dana, 2011; Desta et al.., 2013; Lyimo et al., 2014); both morphological and genomic analyses confirm a high intrapopulation diversity of the two chicken populations. Standing gene flow, a mild impact of human driven selection, long history of breeding and uncontrolled breeding could have resulted in a high diversity of village chickens. Our study corroborates that village chickens are good resource in evolutionary genetics study and to uncover the genetic basis of threshold and Mendelian traits. However, phenotye scoring is a challenging task owing to nondescript nature of village chickens and a fine-scale population structure. High genetic diversity of village chickens also blurs signals of selective sweep and reduces association of genetic markers with phenotypes in genetic mapping studies. Traits that commonly form the basis of Mendelian mapping studies have been highly segregating in village chickens and they display a vast array of variation. Village chickens are also good resource to study the impact of natural selection and adaptive radiation on domesticates' genome.

Natural selection most likely excels artificial selection in village chickens; as a result village chickens' genome may be largely populated by fine scale footprints of natural selection, which then provides insight as to which genomic regions are contributing

to a local adaptation. Natural selection forms the basis of adaptive radiation studies and it is important to identify traits that confer local adaptation and to consider them in genetic improvement programs. We map a large number of putative loci that involve in local adaptation, divergent selection and adaptive radiation of threshold traits and to some extent traits that carry the legacy of human driven selection (e.g., the rose comb mutation). A selective sweep analysis identifies a considerable degree of divergent selection between the two chicken populations, which implicates the impact of a locally-driven demand in adaptive radiation of the two chicken populations and the impact of a localized genetic drift and their demographic history. Our analysis remapped genomic regions that have been subjected to recurrent selection besides mapping novel genomic regions underlying the genetic basis of village chickens adaptation. A strong signal of selective sweep detected for rose comb mutation on GGA7 may be the consequence of the combined effect of natural and artificial selection. A genome-wide association study also identifies a long list of traits underlying the genetic control of threshold and Mendelian traits. Our study is the first of its kind to perform an extensive mapping on pigmentation traits because most of the loci that have been mapped in mouse were remapped by our analysis. This validates our mapping approach and the importance of village chickens to study the genetic control of pigmentation (other threshold and Mendelian) traits.

Extending the classical evolutionary and population genetics studies to translational genomics would assist subsistence farmers to improve performance and welfare of their flocks. Improving the performance of indigenous chicken would help to conserve the standing genetic variation and functional diversity of the village chickens. Extensive exposure of village chickens to natural environment may have enabled them to accumulate a number of advantageous genomic variants that confer a robust character. However, the agricultural extension system of the least developed world is actively distributing commercial stocks with the intention of improving the production performance of local chickens that are thought to be inherently low producing. However, village chickens have been maintained in production systems that are constrained by scarcity of basic production inputs (feed, veterinary service, reliable market outlet etc.), which have partly made them to produce less. With little positive impact (if any) on the livelihood of subsistence farmers, extensive unidirectional gene flow from commercial populations without doubt imposes a

serious threat to genetic integrity of indigenous stock, which then gradually reduces the standing genetic variation accumulated for millennia. Loss of genetic diversity limits the responding capacity of subsistence farmers to climate change.

Our study substantiates that village chickens have to be a part of the mainstream omic studies and translational genomics as they display intermediate genetic structure that connects commercial chickens with the junglefowls. Village chickens could also serve as a model to study the impact of natural selection on domesticates genome and they also carry the legacy of subsistence farmers' maintaining them ever since a prehistoric time. A comparative study that involves a large number of representative village chicken populations across the least developed world, the junglefowl species and the commercial chickens however will provide indisputable genetic evidence on the demographic history of domestic chickens and the junglefowls.

References

DANA, N. Breeding programs for indigenous chicken in Ethiopia: analysis of diversity in production systems and chicken populations. PhD thesis, Wageningen University, 2011.

DESTA, T. T., DESSIE, T., BETTRIDGE, J., LYNCH, S. E., MELESE, K., COLLINS, M., CHRISTLEY, R. M., WIGLEY, P., KAISER, P., TERFA, Z., MWACHARO, J. M. & HANOTTE, O. Signature of artificial selection and ecological landscape on morphological structures of Ethiopian village chickens. Animal Genetic Resources, 2013, 52, 1-13.

HILLIER, L. W., MILLER, W., BIRNEY, E., WARREN, W., HARDISON, R. C., PONTING, C. P., ... & DODGSON, J. B. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature, 2004, 432(7018), 695-716.

KRANIS, A., GHEYAS, A. A., BOSCHIERO, C., TURNER, F., YU, L., SMITH, S., ... & BURT, D. W. Development of a high density 600K SNP genotyping array for chicken. BMC Genomics, 2013, 14(1), 59.

LYIMO, C. M., WEIGEND, A., MSOFFE, P. L., EDING, H., SIMIANER, H. & WEIGEND, S. Global diversity and genetic contributions of chicken populations from African, Asian and European regions. Animal Genetics, 2014, 45(6), 836-848.

MUCHADEYI, F. C., EDING, H., WOLLNY, C. B. A., GROENEVELD, E., MAKUZA, S. M., SHAMSELDIN, R., SIMIANER, H. & WEIGEND, S. Absence of population substructuring in Zimbabwe chicken ecotypes inferred using microsatellite analysis. Animal Genetics, 2007, 38, 332-339.

MWACHARO, J. M., NOMURA, K., HANADA, H., HAN, J. L., AMANO, T. & HANOTTE, O. Reconstructing the origin and dispersal patterns of village chickens across East Africa: insights from autosomal markers. Molecular Ecology, 2013, 22(10), 2683-2697.

OSEI-AMPONSAH, R., KAYANG, B. B., NAAZIE, A., OSEI, Y. D., YOUSSAO, I. A., YAPI-GNAORE, V. C., TIXIER-BOICHARD, M. & ROGNON, X. Genetic diversity of forest and savannah chicken populations of Ghana as estimated by microsatellite markers. Anim Sci Journal, 2010, 81, 297-303.

APPENDICES

Appendix 4

Distribution of SNPs across the autosomes

The mean (sd) physical distance between contiguous syntenic SNPs in the twenty eight autosomes is presented in Table S4.1.

Table S4.1. The mean (sd) physical distance (bp) between contiguous syntenic SNPs in the twenty eight autosomes.

Chromosome	**Size (Mbp)[a]**	**SNPs (n)**	**Mean(SD)**
1	195.277	72750	2683.75 (3485.458)
2	148.810	44819	3319.857 (5612.684)
3	110.448	40400	2733.859 (3367.873)
4	90.217	30853	2922.606 (3521.483)
5	59.580	20879	2852.052 (7093.337)
6	34.952	14918	2339.754 (2636.037)
7	36.245	15267	2371.001 (2573.834)
8	28.767	12424	2312.18 (2129.58)
9	23.442	12871	1820.285 (1789.048)
10	19.911	12507	1590.093 (1946.518)
11	19.401	9767	1983.061 (1841.378)
12	19.897	10452	1899.112 (2173.646)
13	17.760	7699	2295.014 (3084.617)
14	15.162	8522	1777.42 (2107.762)
15	12.657	6965	1811.061 (2149.022)
16	0.535	232	2141.835 (8440.811)
17	10.454	6113	1677.479 (2573.501)
18	11.220	5983	1871.091 (6230.673)
19	9.983	6046	1651.286 (2495.187)
20	14.303	6345	2244.424 (6042.476)
21	6.803	5832	1164.226 (1638.268)
22	4.081	2320	1754.818 (3249.774)
23	5.723	4035	1414.917 (1950.158)

Table S4.1
cont'd

24	6.323	4919	1283.623 (1539.777)
25	2.191	1518	1440.408 (8999.872)
26	5.330	3703	1437.151 (4056.433)
27	5.209	3676	1413.244 (5243.669)
28	4.743	3398	1394.111 (2926.691)

[a] Chromosome size was obtained from http://www.ncbi.nlm.nih.gov/nuccore/CM000119.3

Village level descriptive summary

Heatmap plot produced using gplots package (Warnes et al. 2013) of the R for a subset of data that includes two birds from each household is displayed at village level (Figure S4.1a–h).

From the entire dataset (747 chickens), 49 chickens are included in our analysis as single from each household whereas 698 birds as pair. In Horro 20 chickens analysed as single and 360 chickens as pair from each household. In Jarso 29 chickens analysed as single while 338 birds as pair. At the entire dataset level inbreeding coefficient calculated for households represented by a single chicken versus a pair show non-significant difference ($t_{810.414} = 0.1182$, $P = 0.9059$). Similarly, a dataset containing a pair of chickens ($n = 360$) and a single chicken ($n = 200$) at household level in Horro chickens show non-significant difference ($t_{426.429} = 0.0784$, $P = 0.9376$). Moreover, analysis of two chickens ($n = 338$) versus one ($n = 198$) from each household in Jarso show non-significant difference ($t_{388.05} = 0.1174$, $P = 0.9066$). Absence of significant difference in inbreeding level between the two datasets is in agreement with heatmap plots displayed for each village in Figure S4.1a–h. These figures show mean *IBS* value between pairs of birds sampled from each household. Therefore, there is no much harm from inbreeding point of view at household level; however this may not reflect the real case as our sampling strategy deliberately excluded sampling of birds with known history of familial relationship.

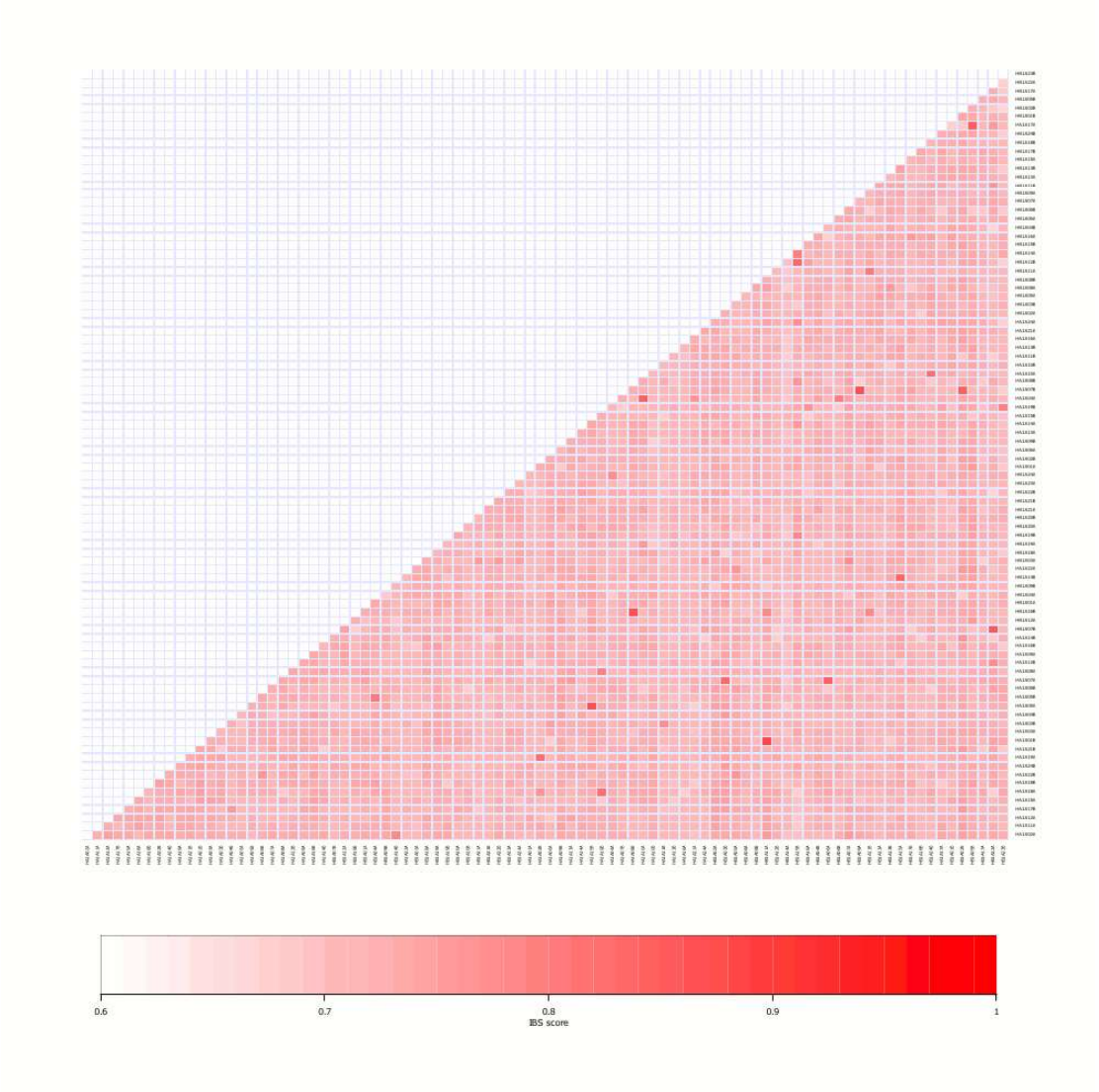


Figure S4.1a. Heatmap graph for mean IBS values in Didibe Chistana chickens.

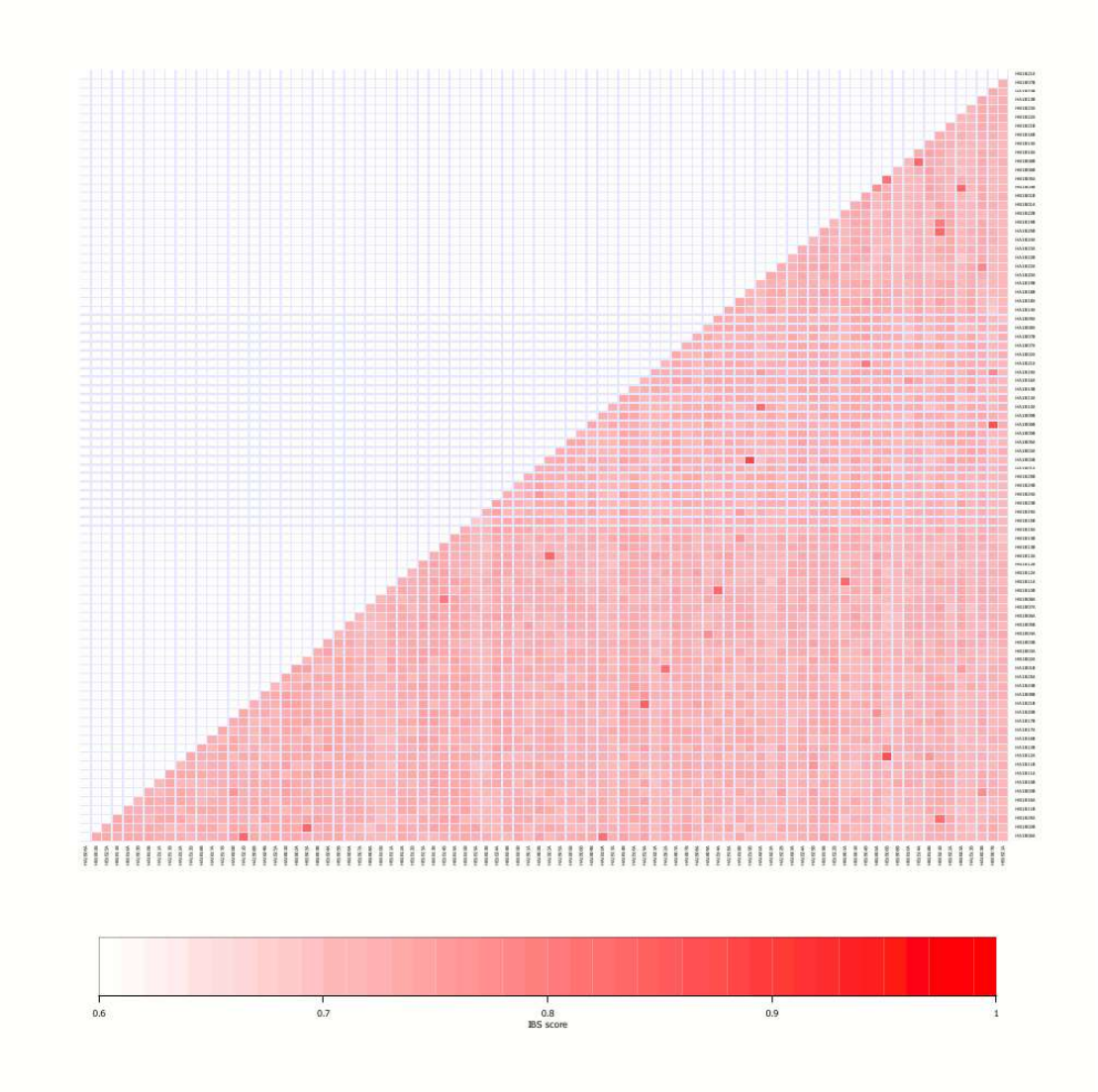


Figure S4.1b. Heatmap graph for mean IBS values in Doyo Beriso chickens.

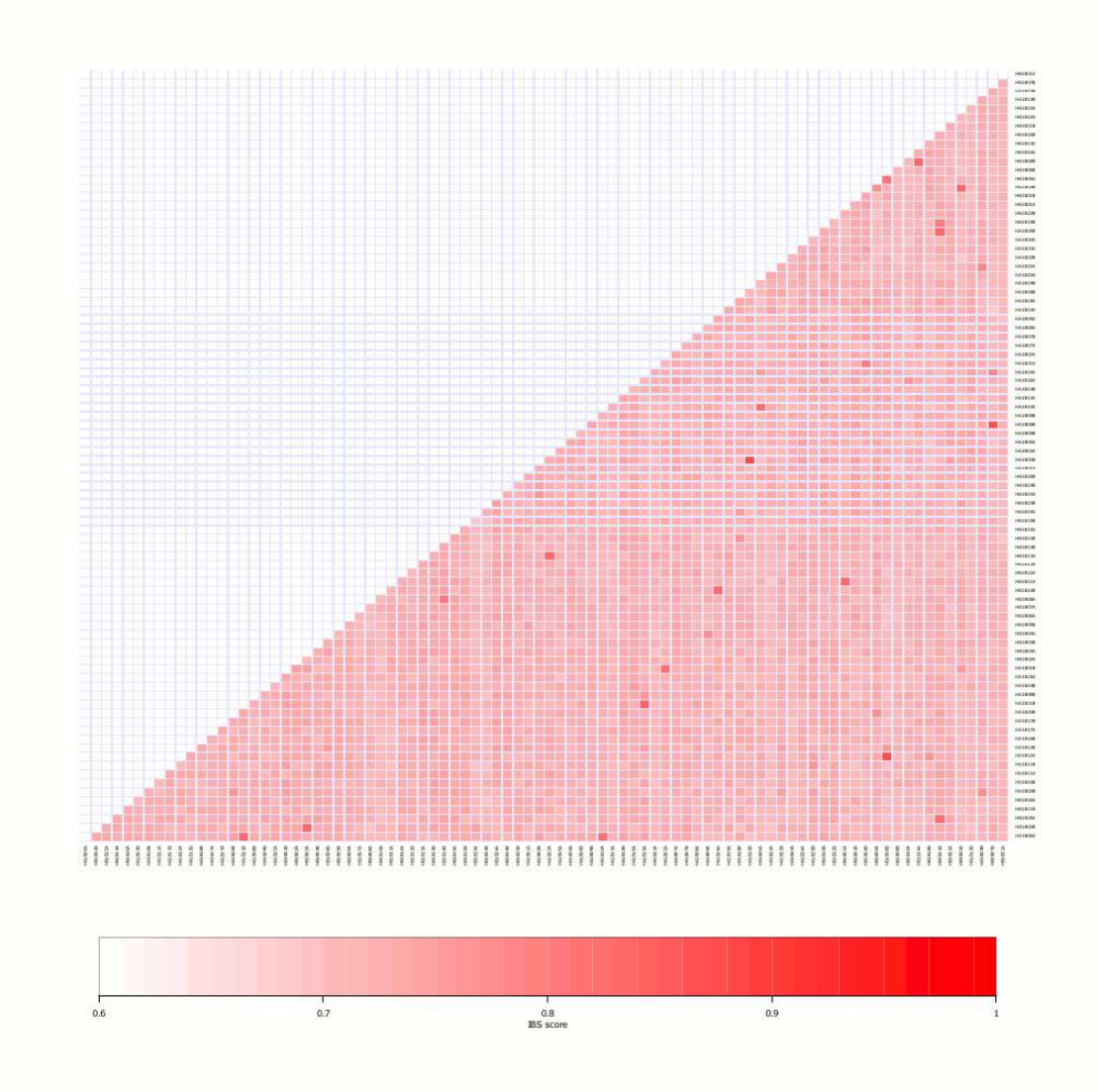


Figure S4.1c. Heatmap graph for mean IBS values in Harro Aga chickens.

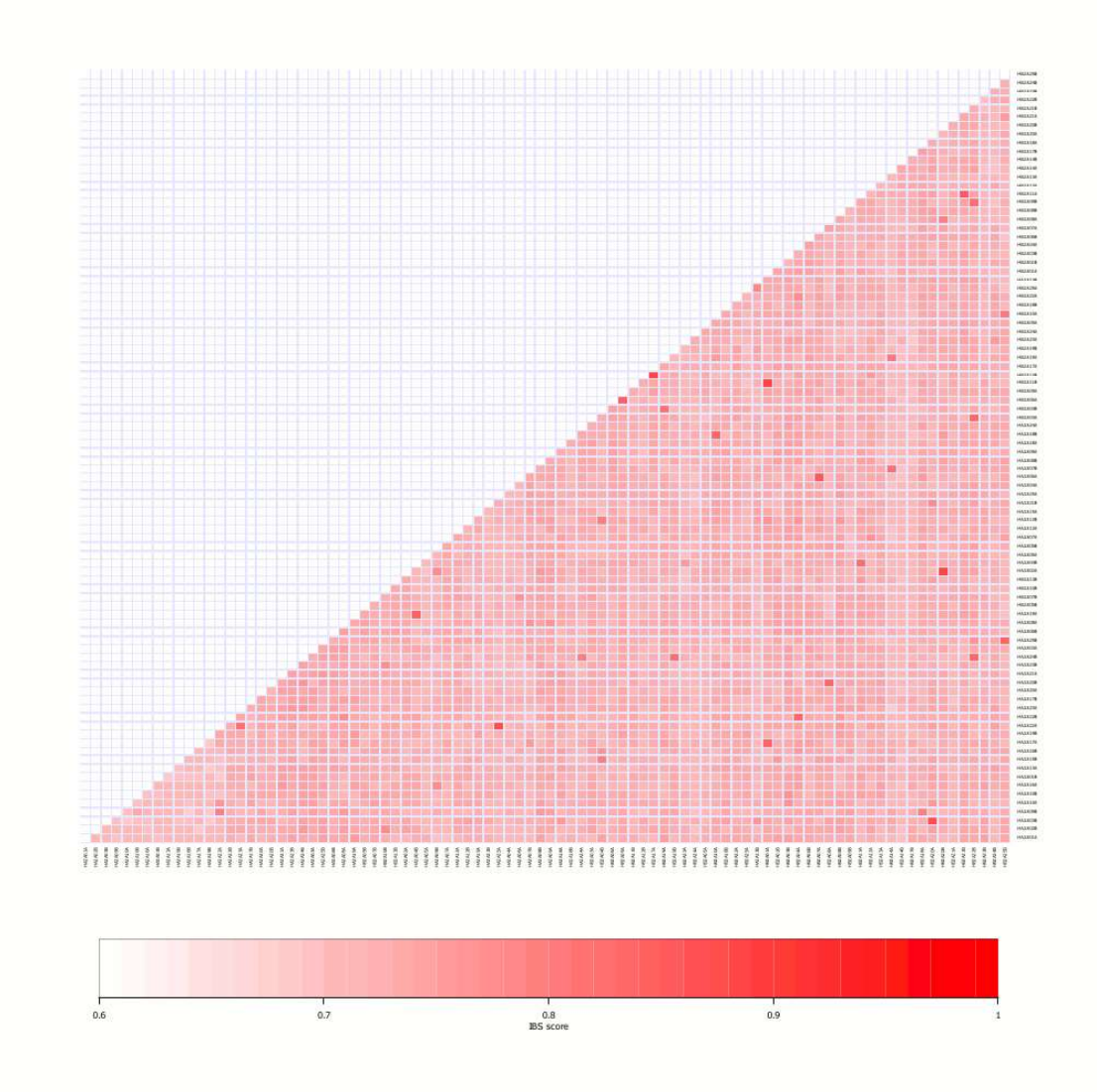


Figure S4.1d. Heatmap graph display for mean IBS values in Bonne Abunna chickens.

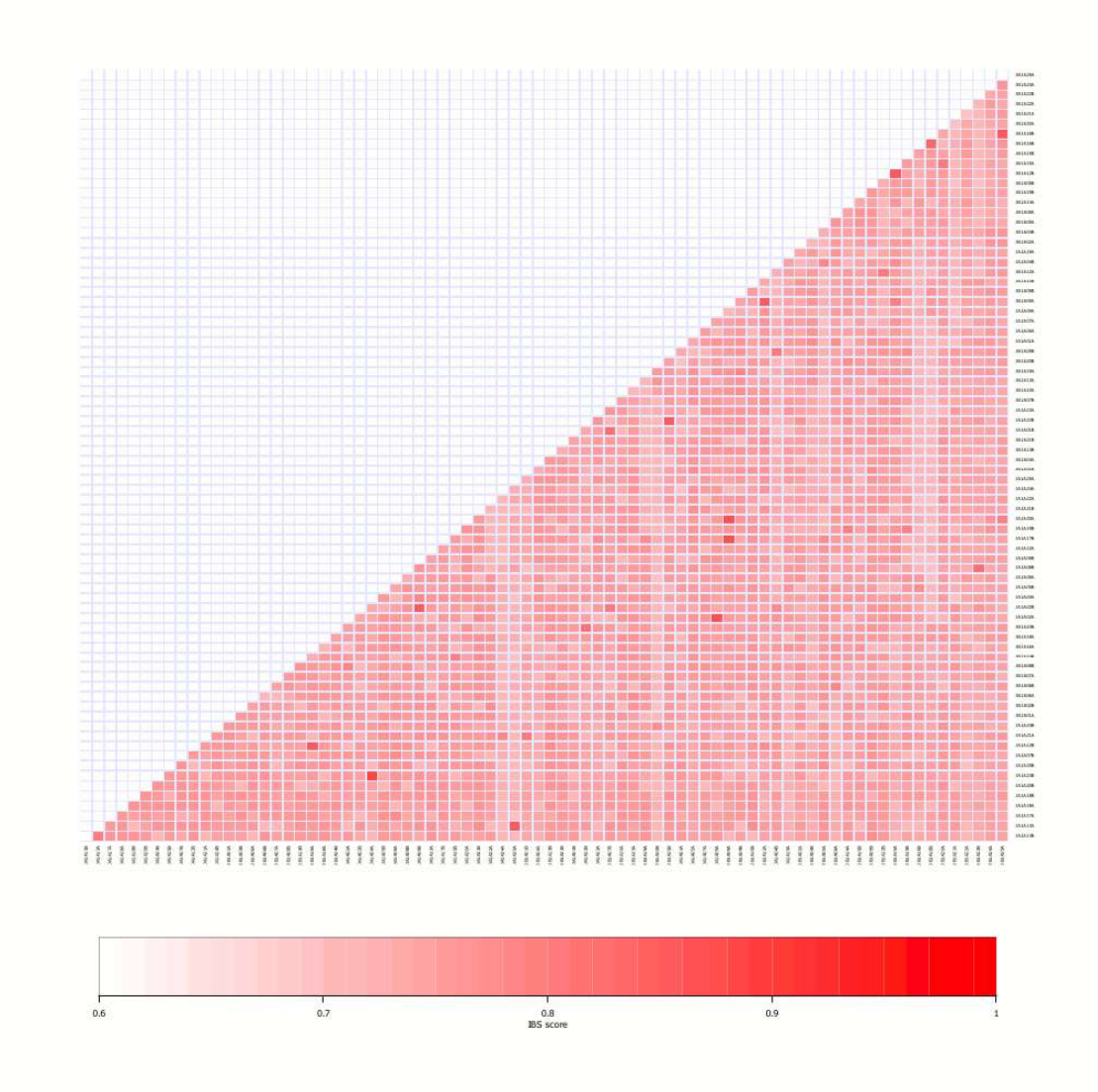


Figure S4.1e. Heatmap graph display for mean IBS values in Afgug chickens.

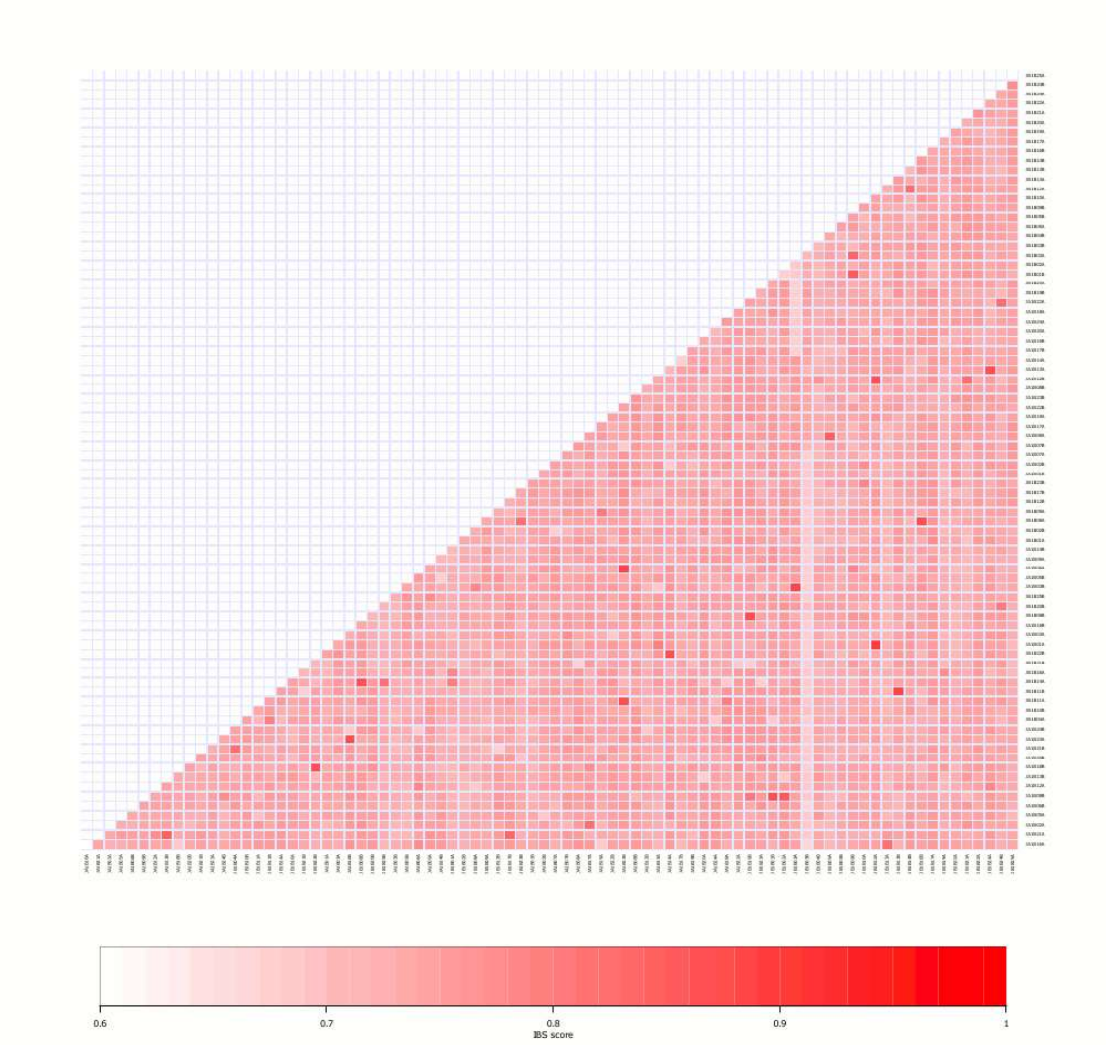


Figure S4.1f. Heatmap graph display for mean IBS values in Bedhasa chickens.

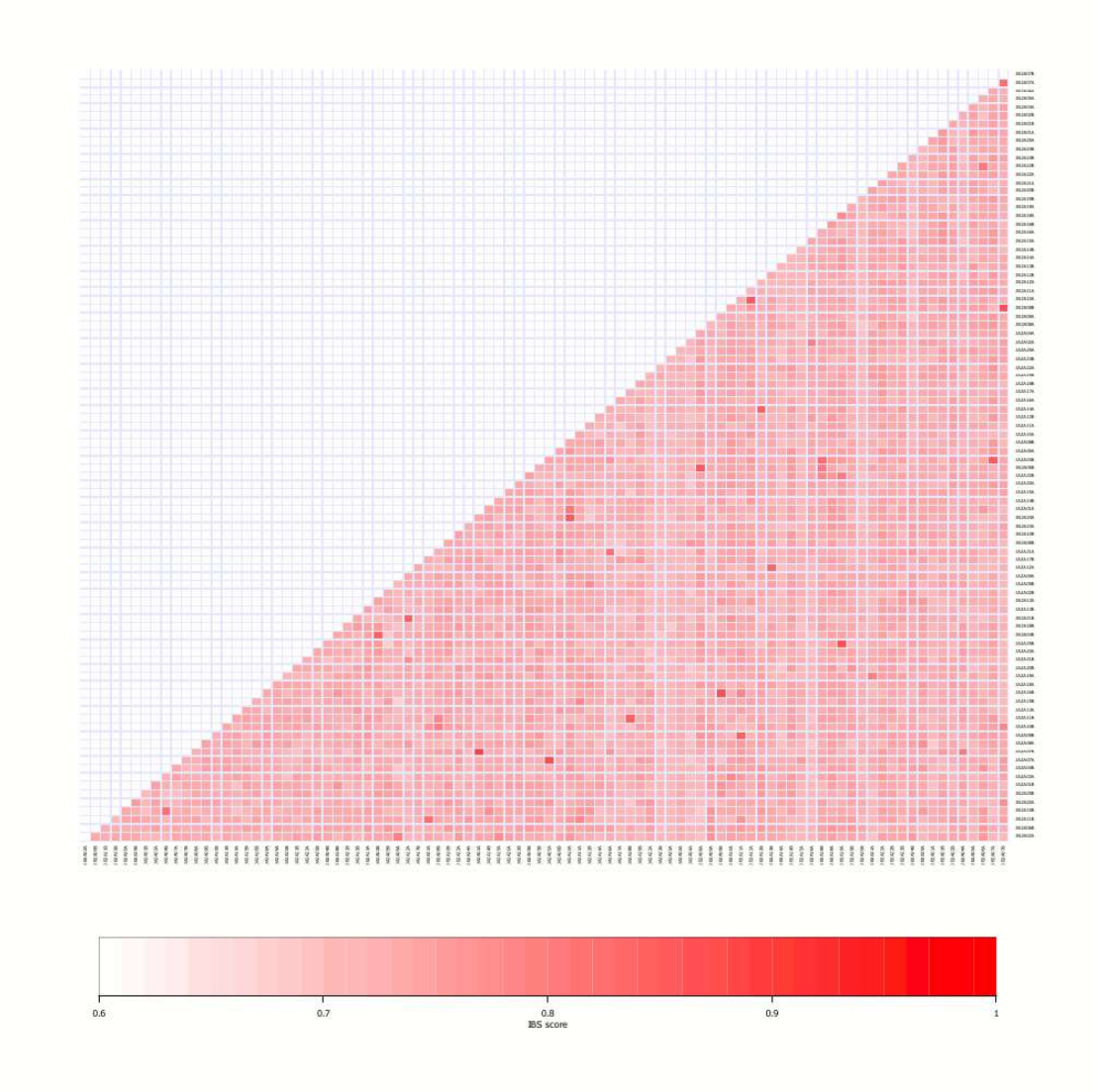


Figure S4.1g. Heatmap graph display for mean IBS values in Lafin Fedho chickens.

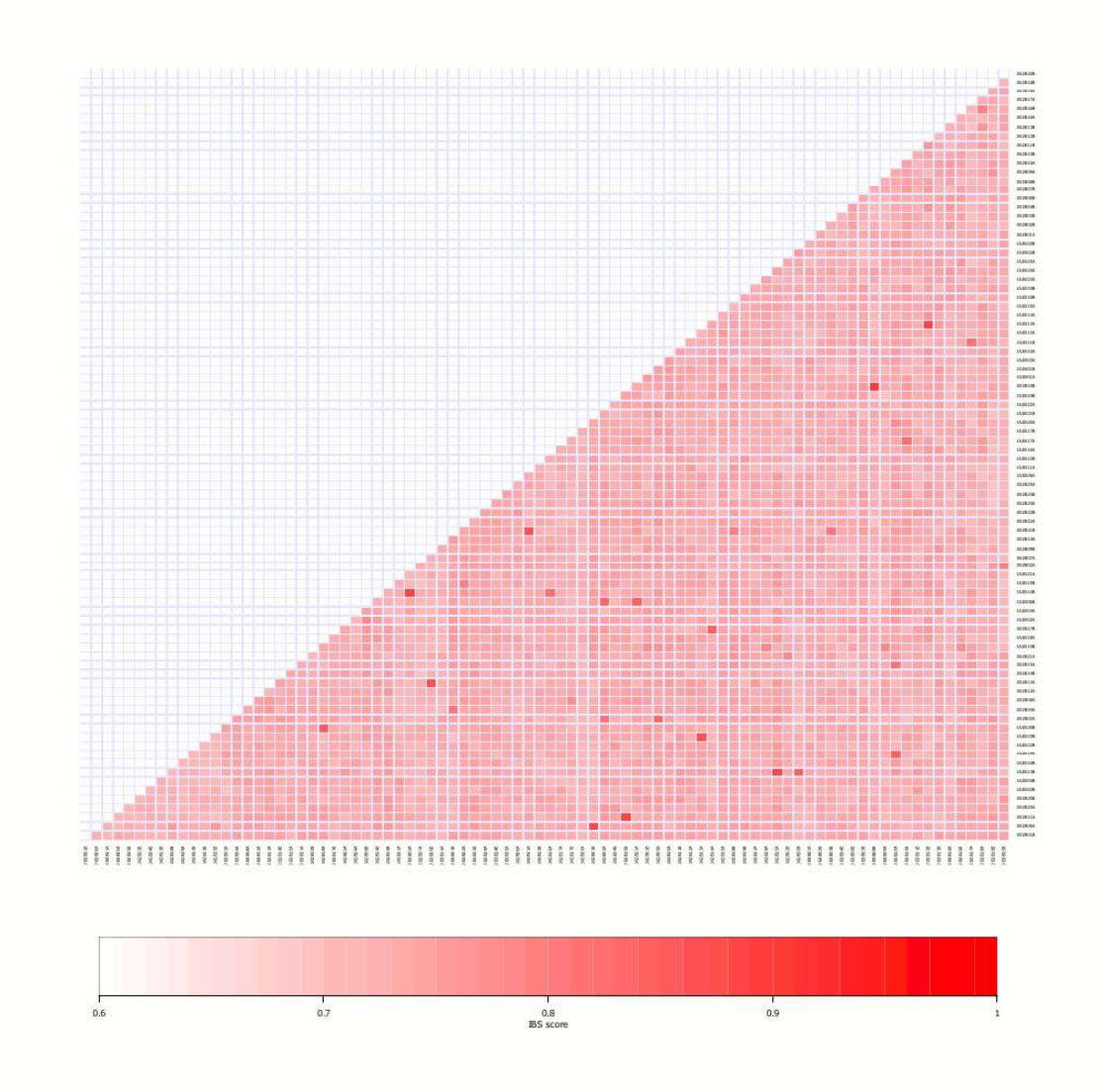


Figure S4.1h. Heatmap graph display for mean IBS values in Aman chickens.

Genetic divergence

Pairwise F_{ST} values among eight villages and four marketsheds sampled from Horro and Jarso were calculated using a custom R script is presented in Table S4.2 and S4.3 respectively. A genetic divergence at intrapopulation level is much lower than interpopulation level.

Table S4.2. Pairwise F_{ST} among chicken populations sampled across eight villages in Horro and Jarso.

Village	DC	DB	HA	BA	AF	BD	LF
DB	0.004						
HA	0.006	0.006					
BA	0.007	0.006	0.005				
AF	0.047	0.047	0.048	0.048			
BD	0.047	0.047	0.048	0.049	0.005		
LF	0.042	0.042	0.044	0.044	0.006	0.006	
AM	0.047	0.046	0.048	0.048	0.006	0.006	0.006

Table S4.3. Pairwise F_{ST} among chicken populations sampled across four marketsheds in Horro and Jarso.

Market shed	H1	H2	J1
H2	0.004		
J1	0.045	0.046	
J2	0.042	0.044	0.003

H1 and H2 refer to marketshed 1 & 2 in Horro and J1 and J2 refer to marketshed 1 & 2 in Jarso respectively.

Principal component analysis

We used the a.score function of discriminant analysis of principal components (DAPC) in adegenet package (Jombart 2008) of the R to identify the optimal number of principal components required to differentiate Horro and Jarso chickens. Figure S4.2 shows that the first PC axis is sufficient to differentiate the two chicken populations.

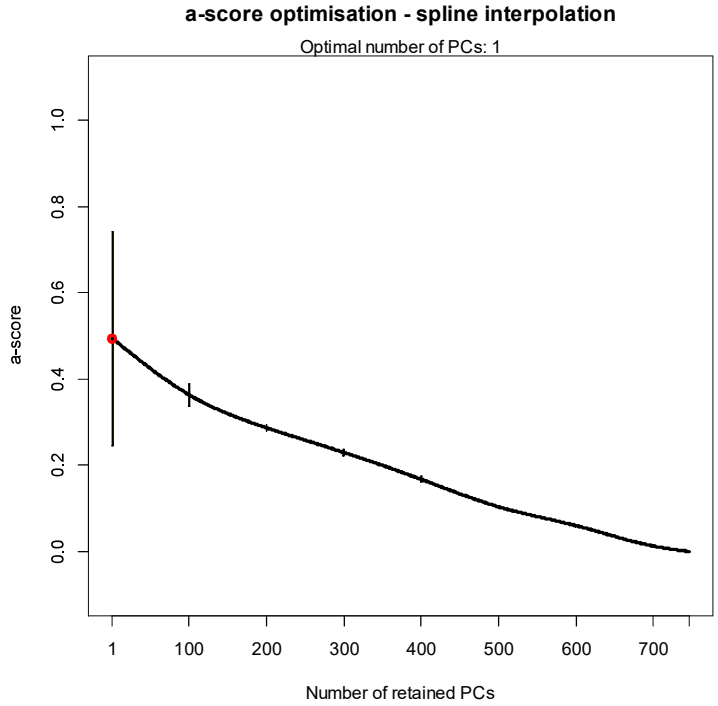


Figure S4.2. The optimum number of principal components required to run the PCA.

The optimal number of genetic clusters was inferred using the Bayesian Information Criterion (BIC) function of adegenet package. Figure S4.3 illustrates that the optimal number of genetic cluster is found to be two.

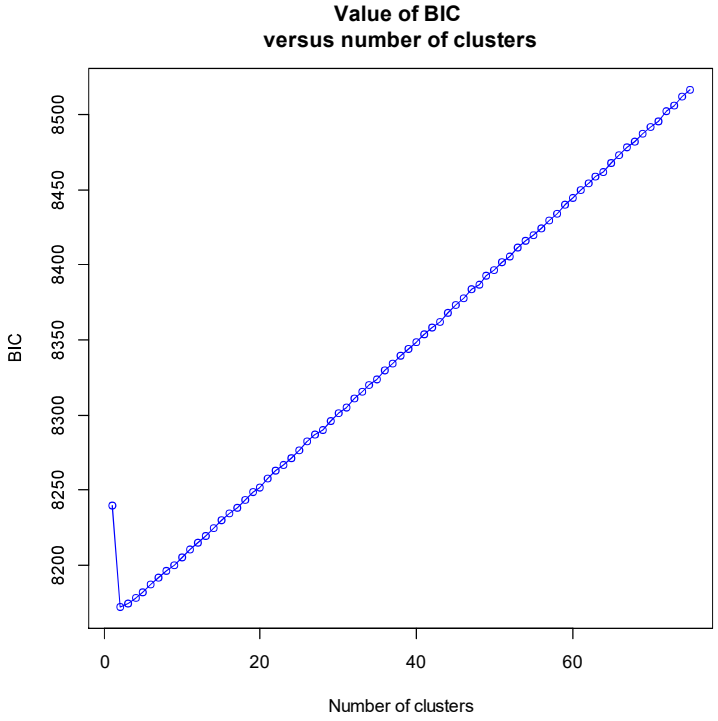


Figure S4.3. The number of optimal genetic clusters identified using BIC.

PCA analysis was performed using ade4 package (Dray & Dufour, 2007) for R. Figure S4.4a & b show absence of genetic substructure at population level. Figure S4.4c shows a PCA plot for random subsets of 25 chickens from each Horro and Jarso.

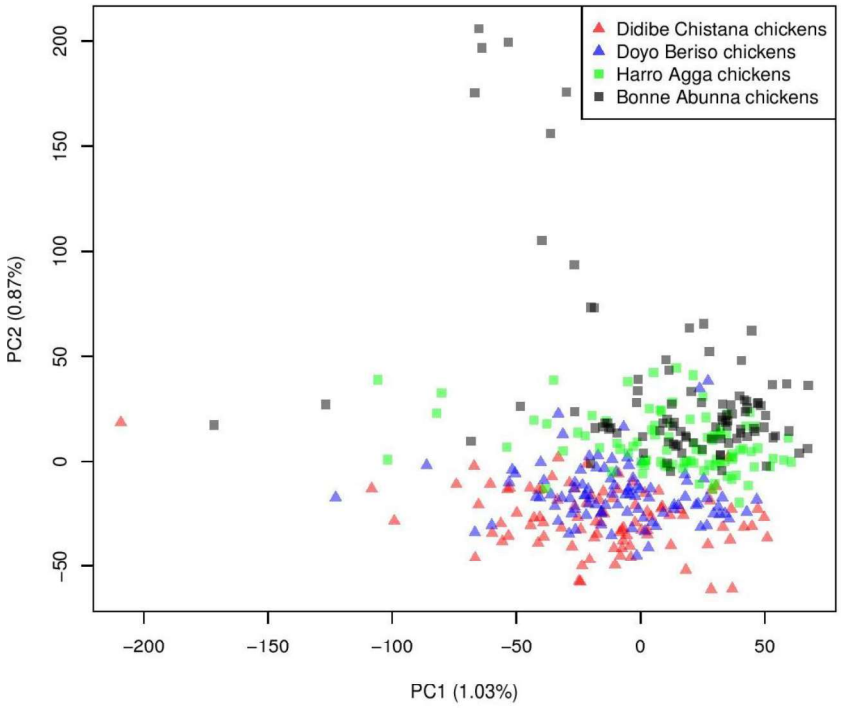


Figure S4.4a. PCA plot for chickens sampled across four villages in Horro.

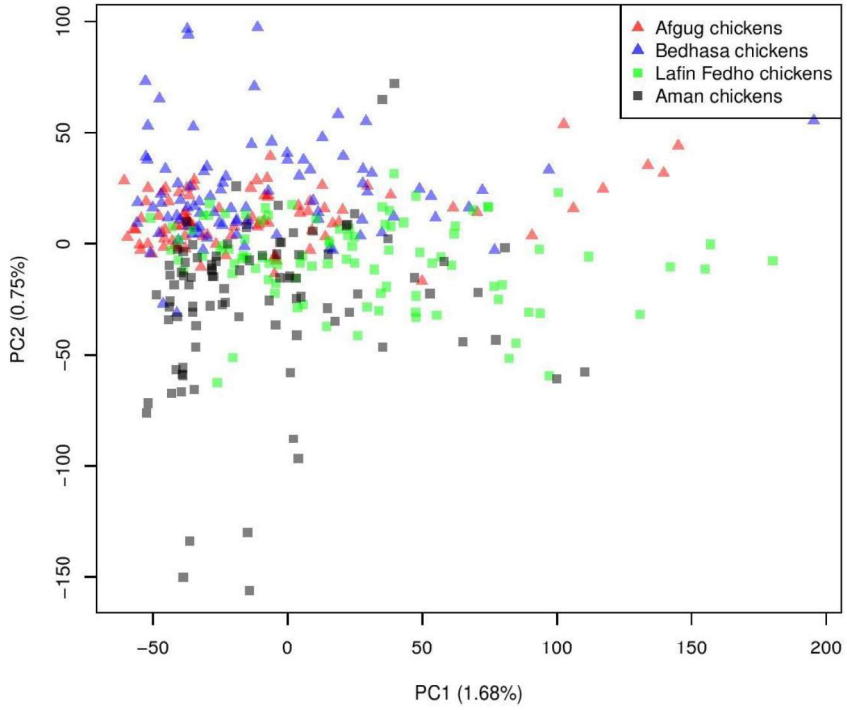


Figure S4.4b. PCA plot for chickens sampled across four villages in Jarso.

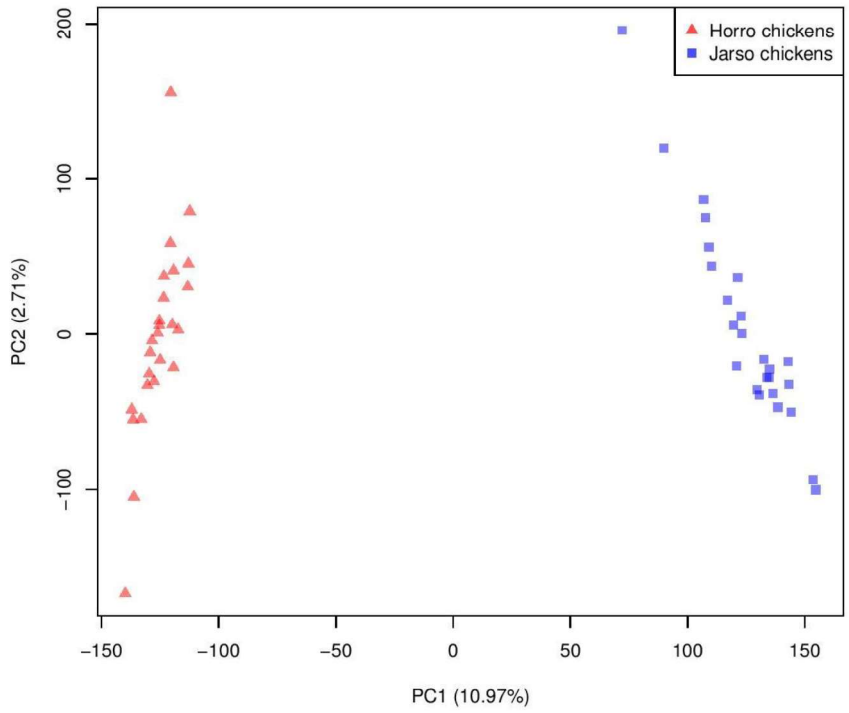


Figure S4.4c. PCA plot for random subsets of 25 chickens from each Horro and Jarso.

Genetic admixture analysis

A genetic admixture analysis performed using the dapc function of adegenet package for Horro and Jarso chicken populations is presented in Figure S4.5a & b respectively. The impact of marketshed on genetic structure is more evident in Horro than Jarso chickens.

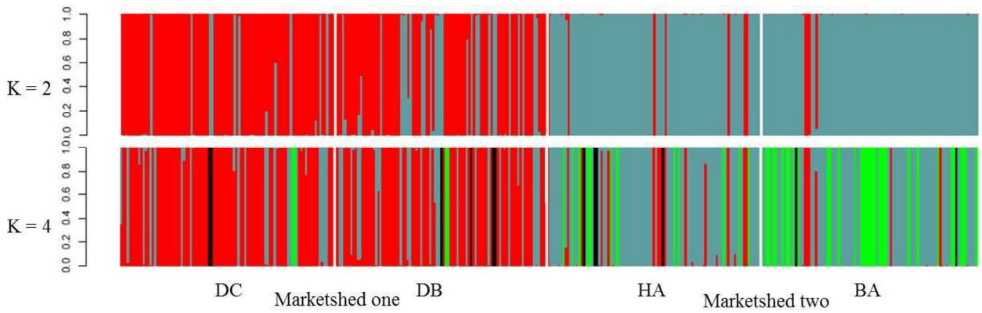


Figure S4.5a. Genetic admixture analysis of Horro chickens.

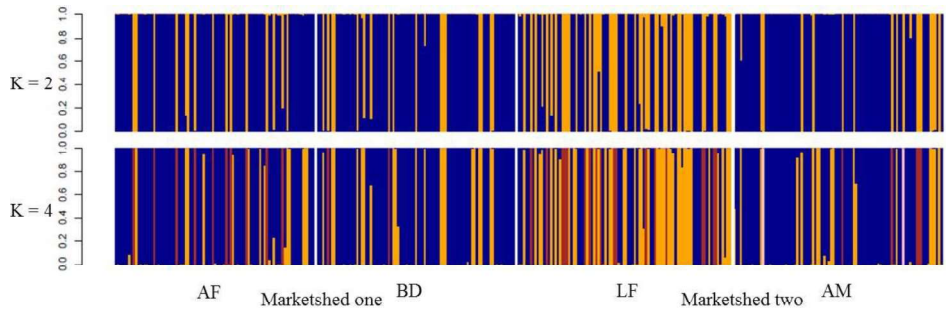


Figure S4.5b. Genetic admixture analysis of Jarso chickens.

Genetic versus geographical distances

We calculate net and between and within village genetic distance (Table S4.4 and Table S4.5) from *IBS* matrix generated from GenABEL using MEGA5 (Tamura et al. 2011). The correlation between genetic and a log transformed geographic distance is displayed in Figure S4.6a–b.

Table S4.4. Within, between and net genetic distance among chicken populations sampled across eight villages in Horro and Jarso.

Village	DC	DB	HA	BA	AF	BD	LF	AM
DC	**0.279**	0.001	0.003	0.003	0.040	0.040	0.036	0.040
DB	0.279	**0.277**	0.002	0.003	0.040	0.040	0.036	0.040
HA	0.280	0.278	**0.275**	0.002	0.041	0.041	0.038	0.041
BA	0.279	0.277	0.275	**0.272**	0.042	0.042	0.038	0.042
AF	0.309	0.308	0.308	0.307	**0.259**	0.002	0.002	0.002
BD	0.311	0.310	0.310	0.309	0.262	**0.262**	0.003	0.002
LF	0.312	0.310	0.311	0.310	0.267	0.269	**0.271**	0.002
AM	0.312	0.311	0.311	0.310	0.264	0.266	0.270	**0.265**

DC: Didibe Chistana, DB: Doyo Beriso, HA: Harro Aga and BA: Bonne Abunna belonged to Horro district; while AF: Afgug, BD: Bedhasa, LF: Lafin Fedho and AM: Aman represent Jarso.

Above the diagonal is the net genetic distance (DA) between chickens sampled from eight villages; $DA = \frac{d_{xy} - (d_x + d_y)}{2}$; DA is the mean net genetic distance, d_{xy} is between group genetic distance while d_x and d_y represent within group mean genetic

distance. On the diagonal (bold) is the mean within genetic distance among chickens sampled from the same village and below the diagonal is the mean genetic distance between chicken sampled from different villages.

Table S4.5. Within, between and net genetic distance among chickens sampled from four marketsheds in Horro and Jarso.

	H1	**H2**	**J1**	**J2**
H1	0.279	0.002	0.040	0.037
H2	0.278	0.274	0.041	0.039
J1	0.310	0.309	0.261	0.001
J2	0.311	0.310	0.267	0.269

H1: Horro marketshed one, H2: Horro marketshed two, J1: Jarso marketshed one and J2: Jarso marketshed two. On the diagonal is within marketshed, below diagonal is between marketshed and above diagonal represents net genetic distance.

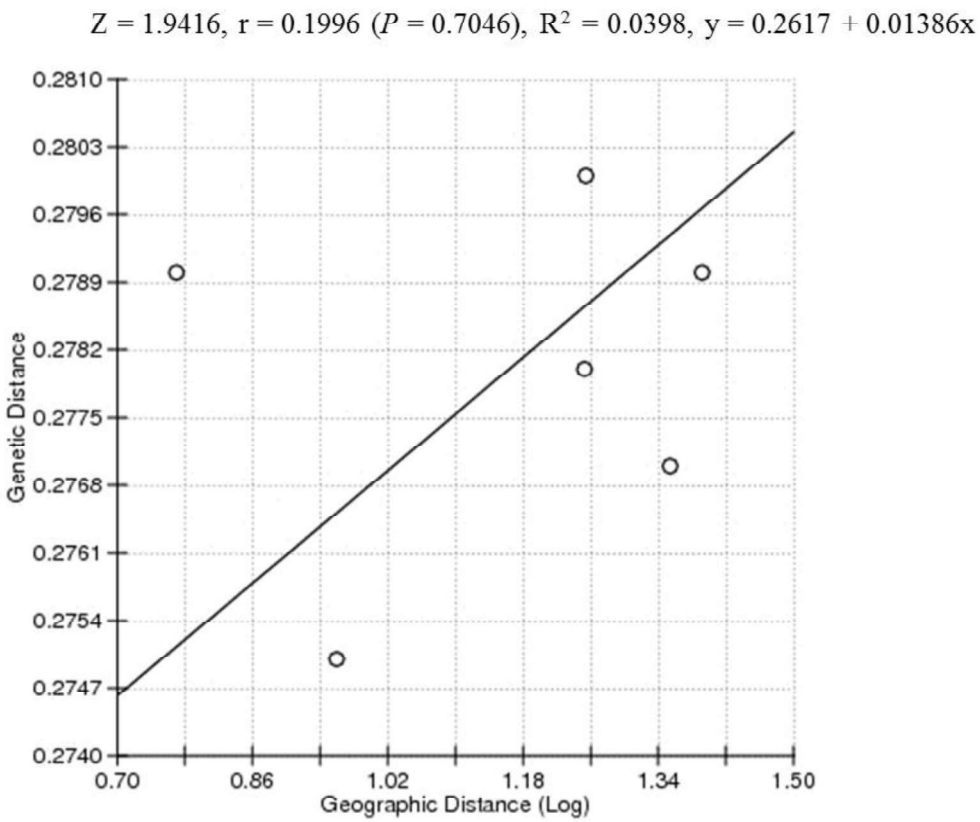


Figure S4.6a. Relationship between genetic and geographic (log transformed) distance in Horro chickens.

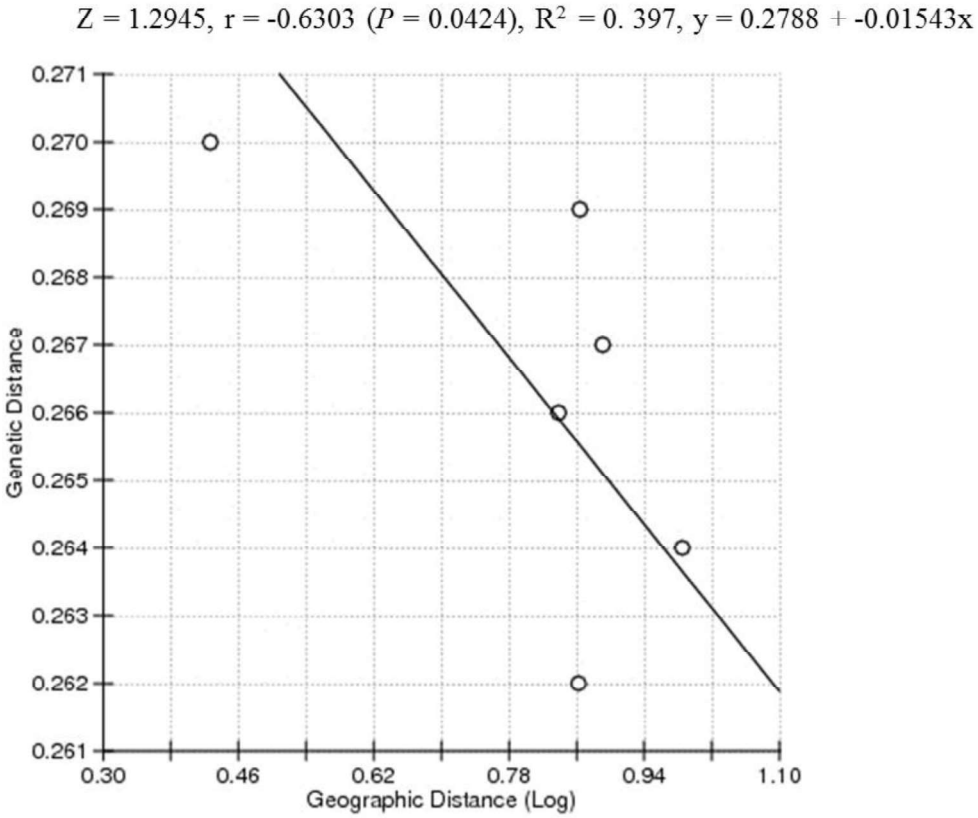


Figure S4.6b. Relationship between genetic and geographic (log transformed) distance in Jarso chickens.

Phylogenetic relationship

We construct a phylogenetic tree at population and marketshed level. Figure S4.7a & b display dendrogram constructed for two marketsheds of Horro and Jarso chickens respectively. Figure S4.7c–f present dendrogram constructed at village level in the four marketsheds.

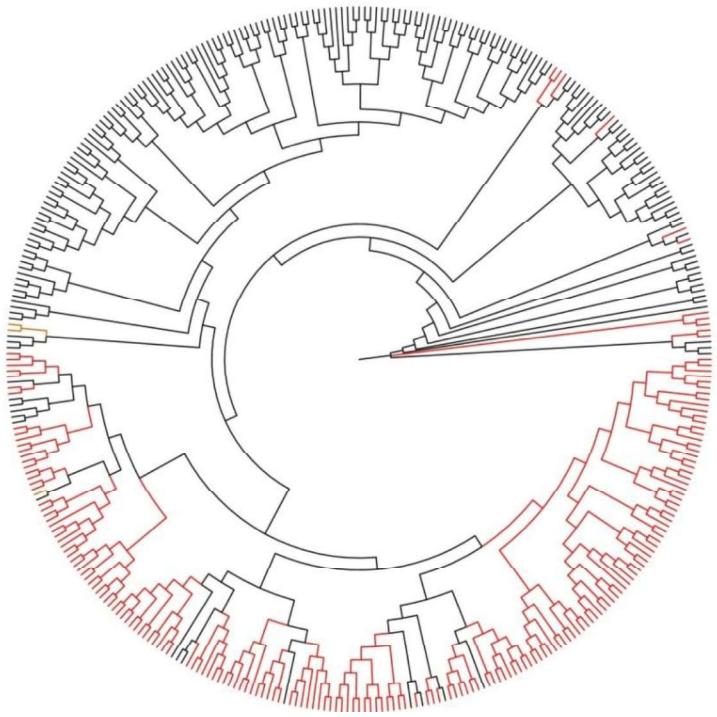


Figure S4.7a. The dendrogram for chickens sampled from two marketsheds in Horro (black represents marketshed one and orange marketshed two).

Figure S4.7b. The dendrogram for chickens sampled from two marketsheds in Jarso (black represents marketshed one and red marketshed two).

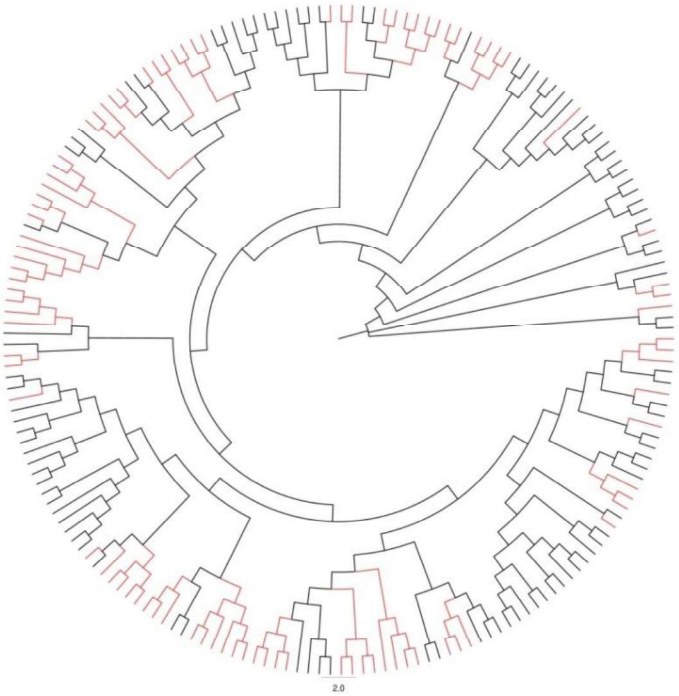


Figure S4.7c. The dendrogram for chickens sampled from marketshed one in Horro (black represents Didibe Chistana village and red Doyo Beriso village).

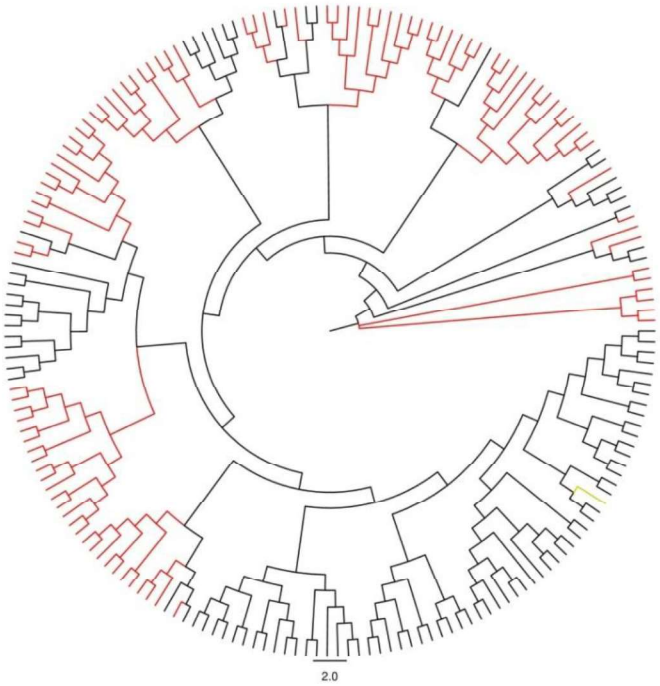


Figure S4.7d. The dendrogram for chickens sampled from marketshed two in Horro (black represents Harro Aga village and red Bonne Abunna village).

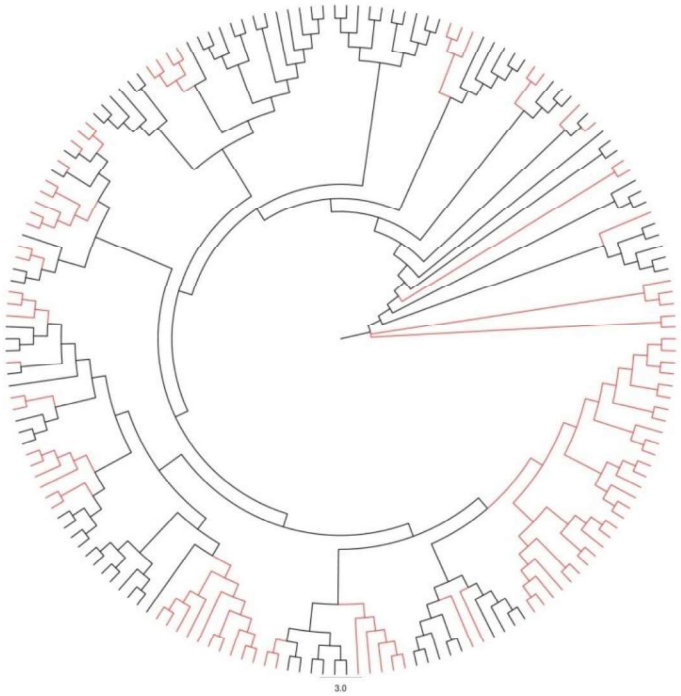


Figure S4.7e. The dendrogram for chickens sampled from marketshed one in Jarso (black represents Afgug village and red Bedhasa village).

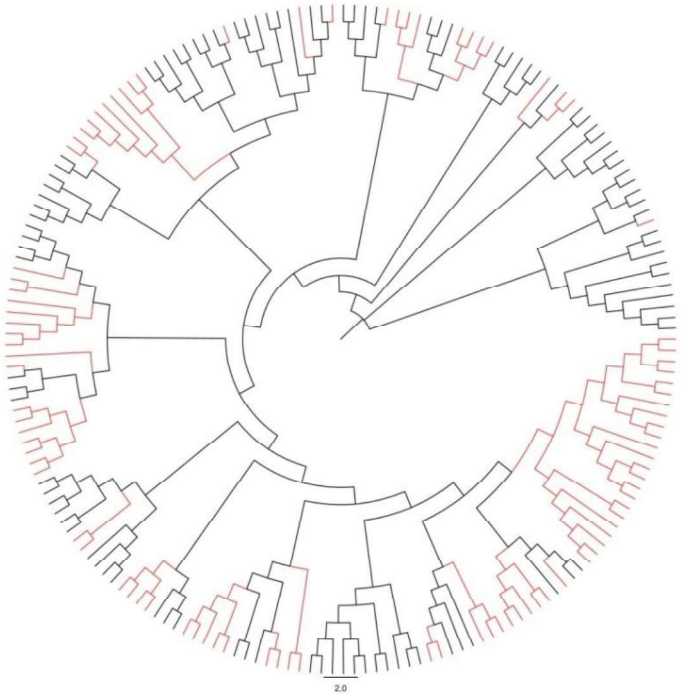


Figure S4.7f. The dendrogram for chickens sampled from marketshed two in Jarso (black represents Lafin Fedho village and red Aman village).

Linkage disequilibrium

Linkage disequilibrium was calculated using the r2fast function of GenABEL at population and autosome level and the binned values are presented in Figure S4.8a–c.

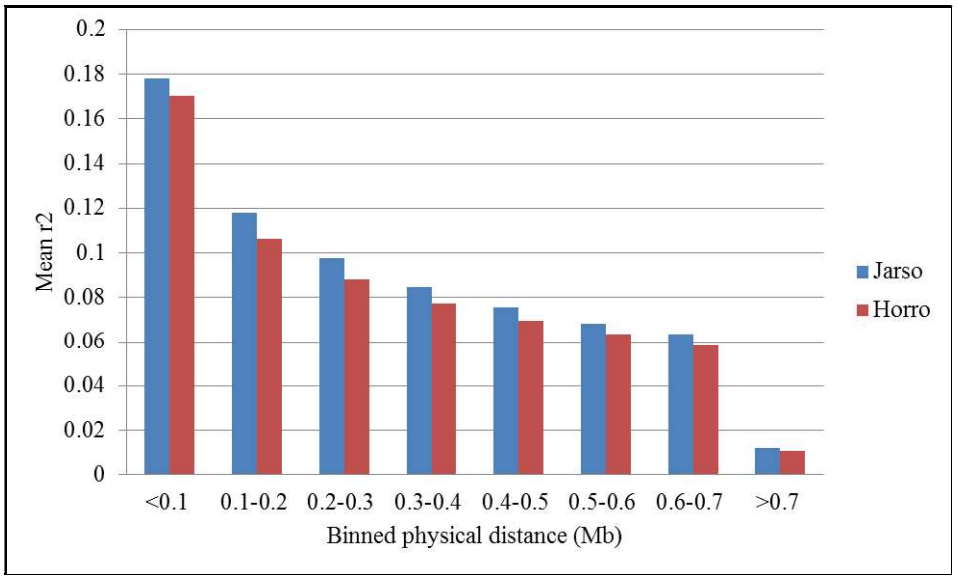


Figure S4.8a. Mean r^2 for macrochromosomes (GGA1–5) in Horro and Jarso chickens.

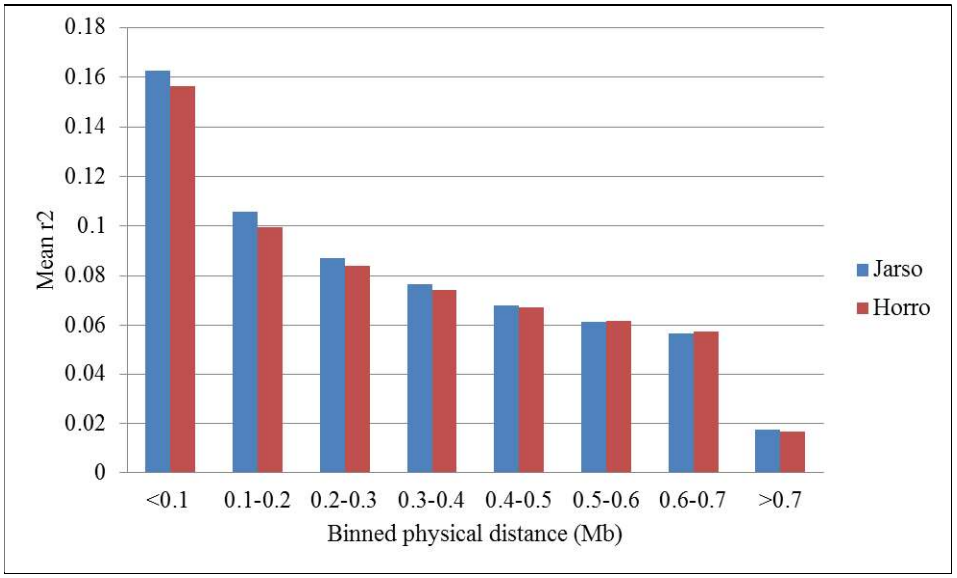


Figure S4.8b. Mean r^2 for intermediate autosomes (GGA 6–10) in Horro and Jarso chickens.

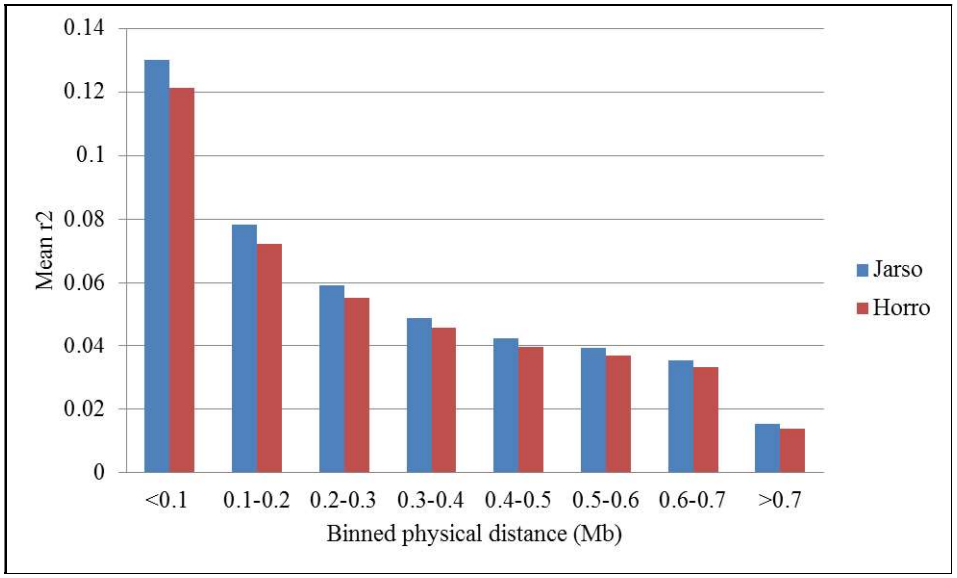


Figure S4.8c. Mean r^2 for microchromosomes (GGA 11–28) in Horro and Jarso chickens.

Effective population size estimate at marketshed level

Marketshed level effective population estimate is presented in Figure S4.9. Two trends are produced from Ne calculated in Jarso chickens, which shows a contrasting demographic history of Jarso chickens.

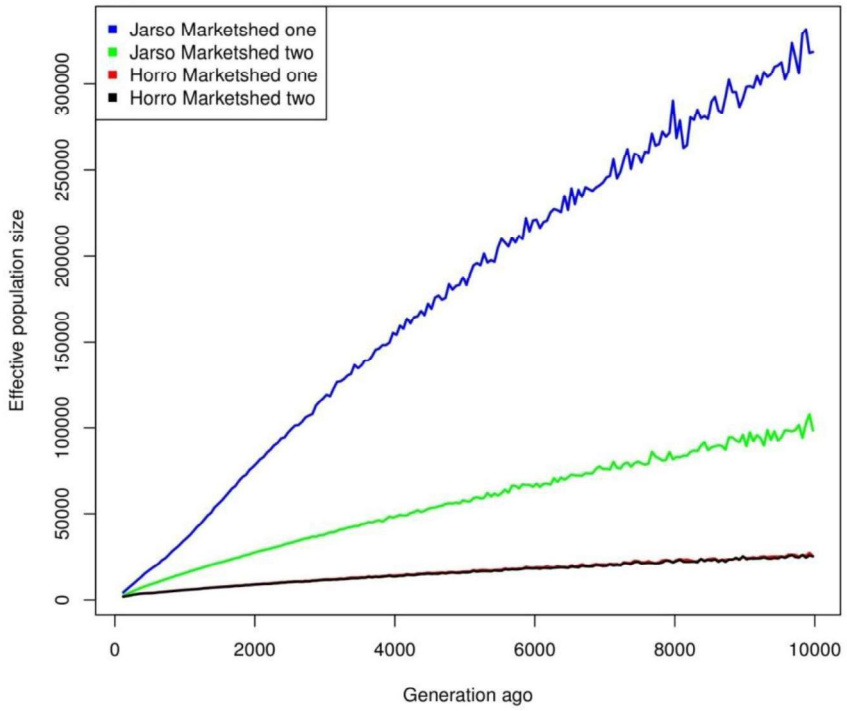


Figure S4.9. Effective population size estimates of Horro and Jarso chickens at marketshed level.

References

AULCHENKO, Y. S., RIPKE, S., ISAACS, A. & VAN DUIJN, C. M. GenABEL: an R package for genome-wide association analysis. Bioinformatics, 2007, 23(10), 1294–1296.

DRAY, S. & DUFOUR, A. B. The ade4 package: implementing the duality diagram for ecologists. J STAT SOFTW, 2007, 22(4), 1-20.

JOMBART, T. adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics, 2008, 24, 1403–1405.

TAMURA, K., PETERSON, D., PETERSON, N., STECHER, G., NEI, M. & KUMAR, S. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution, 2011, 28, 2731–2739.

WARNES, G. R., BOLKER, B., BONEBAKKER, L., GENTLEMAN, R., HUBER, W., LIAW, A., ... & VENABLES, B. gplots: Various R programming tools for plotting data. R package version, 2009, 2(4).

Appendix 5

Genetic cluster definition

The optimal number of principal components need to be retained to analyze the genetic structure of village and commercial chickens and the junglefowl species found to be twenty (Figure S5.1). We used the BIC statistics as implemented in the DAPC of the R package adegenet (Jombart 2008) to identify the optimal number of genetic clusters (Figure S5.2).

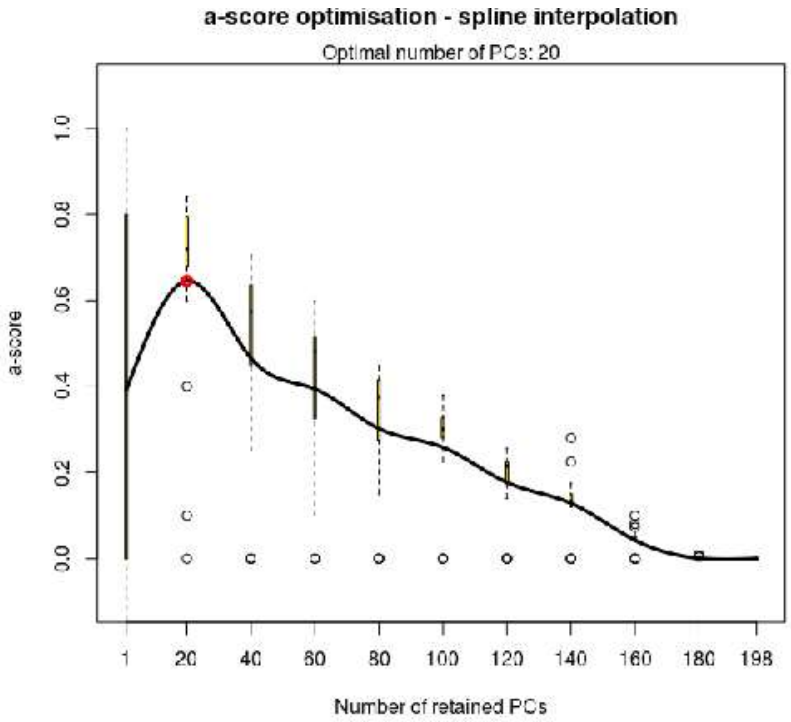


Figure S5.1. The optimal number of principal components required to analyse the genetic structure of domestic chickens and the junglefowl species.

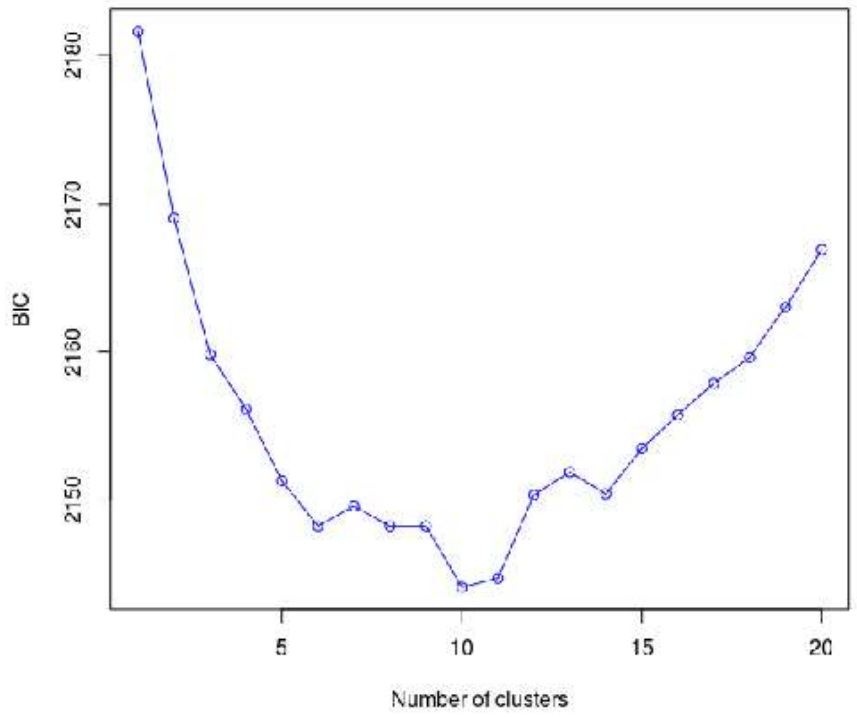


Figure S5.2. The optimal number of genetic clusters found using the find.clusters function of adegenet.

Population structure

Principal component analysis of village chickens and the junglefowl species

We reduced the number of Ethiopian chickens included in the PCA to a random subset of five from each of Horro and Jarso to minimize the impact of sample size variation. Principal component one that explains 10.01% of the total genetic variation separates village chickens from the junglefowl, while PC2, which accounts for 4.62% of the variation separates the junglefowl species from village chickens. The junglefowl are widely scattered along the PC2 axis, which may implicate the genetic divergence among the four wild species. We found that *G. lafayetti*, *G. sonnerattii* and *G. varius* are relatively genetically distant from domestic chicken than *G. gallus* subspecies. Despite their geographical heterogeneity, other villages chickens made a tight cluster and clearly diverged from Ethiopian chickens, which implies that genetic relatedness may not linearly correlate with geographical proximity.

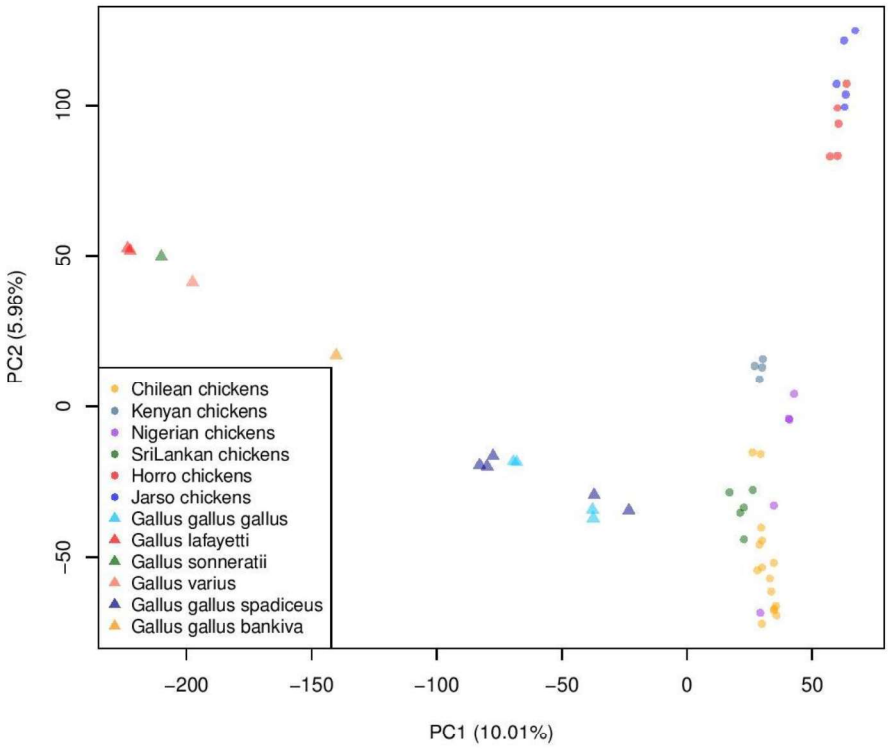


Figure S5.3 Principal component analysis of village chickens and the junglefowl species.

Principal component analysis of Ethiopian chickens and commercial layers

We used fifty Ethiopian chickens and commercial layers in this analysis. Ethiopian chickens were clearly diverged from the commercial layers (Figure S5.3) by PC1 axis that explains 12.68% of the total genetic variation, while PC2 axis (explaining 8.34% of the variation) separated white egg layers from brown egg layers and Ethiopian village chickens. Commercial layers also tended to genetically diverge from each other following their selection history for brown and white egg production.

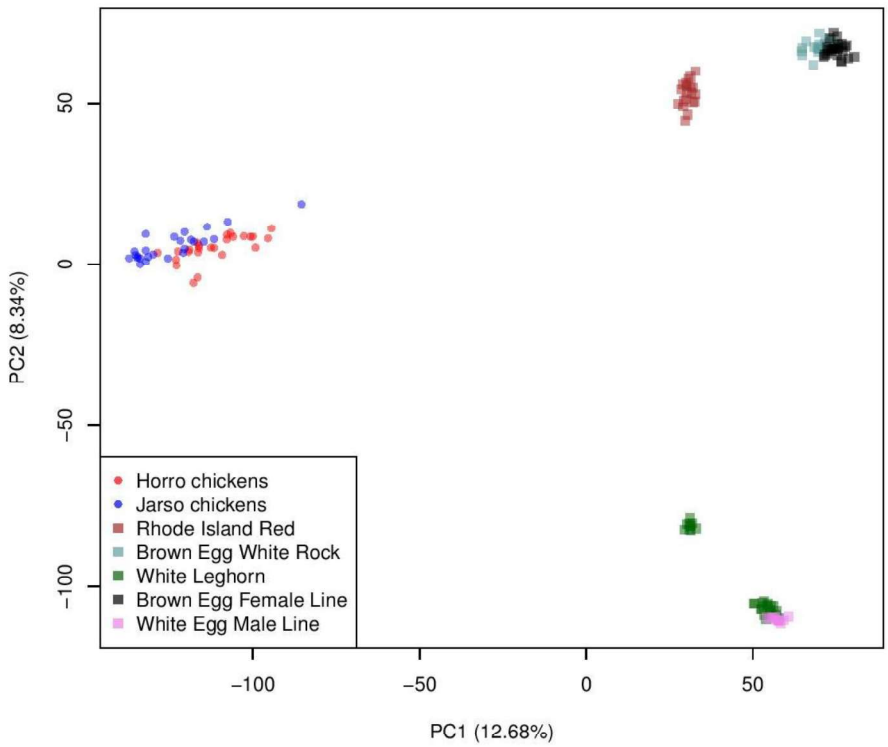


Figure S5.4. PCA plot of randomly selected fifty Ethiopia chickens and commercial layers.

Principal component analysis of Ethiopian and other village chickens

While we analyse ten Ethiopian chickens (five from each Horro and Jarso), they are separated from other village chickens (Figure S5.5) by the first PC that explains 8.69% of the total genetic variation. The PC2 axis that accounted for 4.93% of the genetic variation separates Kenyan and some of the Nigerian chickens from Ethiopian, Sri Lankan, Chilean and the remaining Nigerian chickens.

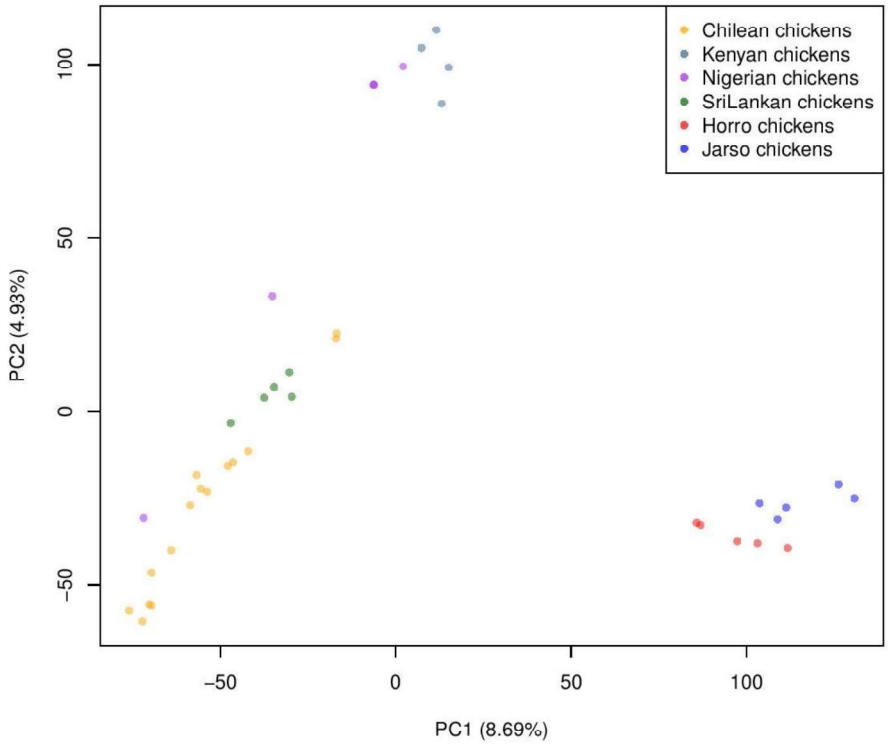


Figure S5.5. PCA plot of randomly selected ten Ethiopian and other village chickens.

Principal component analysis of Ethiopian chickens and the junglefowl species

As to our expectation, Ethiopian chickens were separated by the first PC that explains the highest proportion of the genetic variation (19.66%) from the junglefowl species (Figure S5.6). The PC2 axis that explains 10.71% of the total genetic variation separated non-red junglefowl and Ethiopian village chickens from *G .g. gallus, G. g. spadiceus* even though Ethiopian chicken are closer to the red junglefowl on PC1.

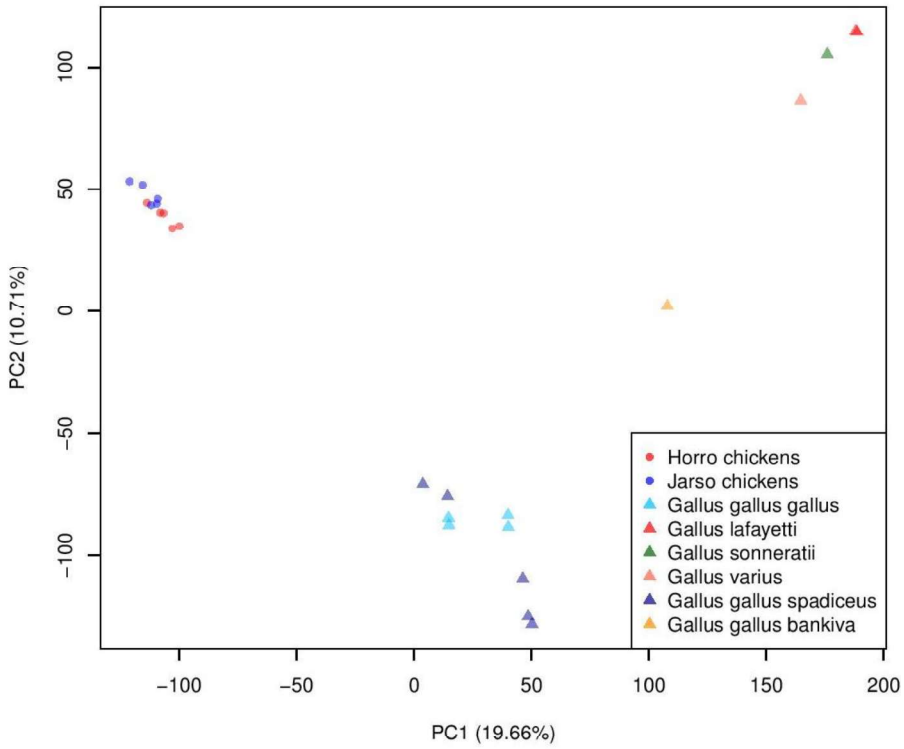


Figure S5.6. PCA plot of randomly selected ten Ethiopian chickens and the junglefowl species.

Principal component analysis of other village chickens and the junglefowl species

Our analysis shows that other village chickens were genetically diverged from the junglefowl species by the first PC (Figure S5.7). Based on PC1, Other village chickens also closely related to *G. g. gallus* and *G. g. spadiceus* as Ethiopian chickens are (Figure S5.6).

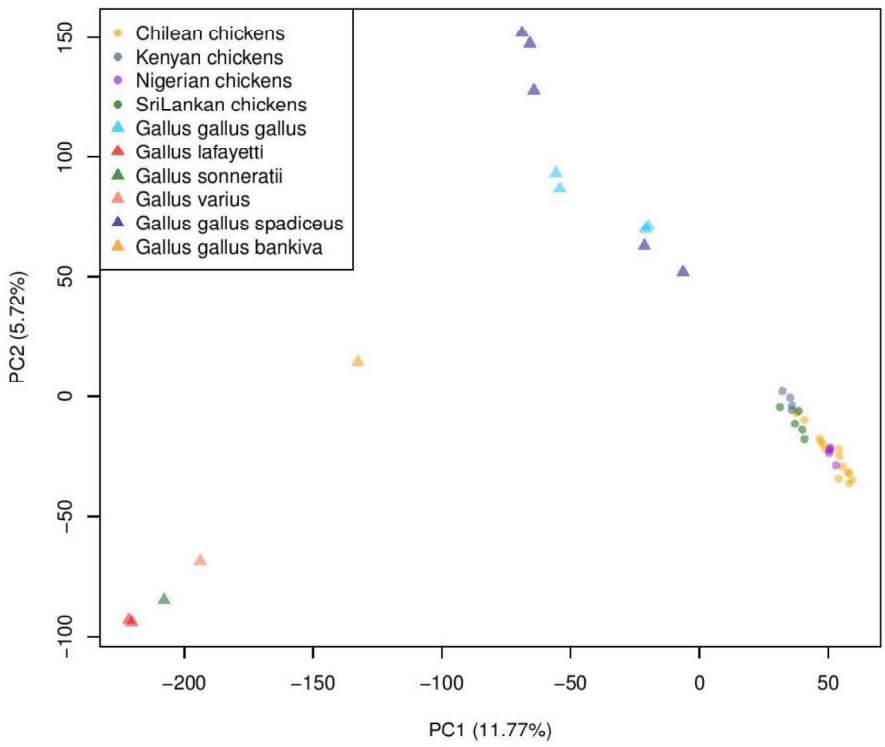


Figure S5.7. PCA plot of other village chickens and the junglefowl species.

References

DRAY, S. & DUFOUR, A. B. The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software, 2007, 22(4), 1–20.

JOMBART, T. adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics, 2008, 24, 1403–1405.

Appendix 6

Variability in number of ROH segments and sum of ROH tracts

There is a considerable variation in the number of ROH segments and their sum length between hen and cock groups in both Horro and Jarso chickens. The variabilty is more pronounced in Jarso than Horro chickens (Figure S6.1 & 2), indicating the genetic homogenity of Horro chickens. Both the minimum and the maximum number of SNPs that made 1Kb ROH tract (DENSITY) were mapped to two Jarso chickens on GGA22 and GGA5 respectively. The high DENSITY found in this study implicates the adequate genomic coverage of the SNPs included in ROH mapping. The longest ROH tract made by a single SNP was ~ 2.6Kb.

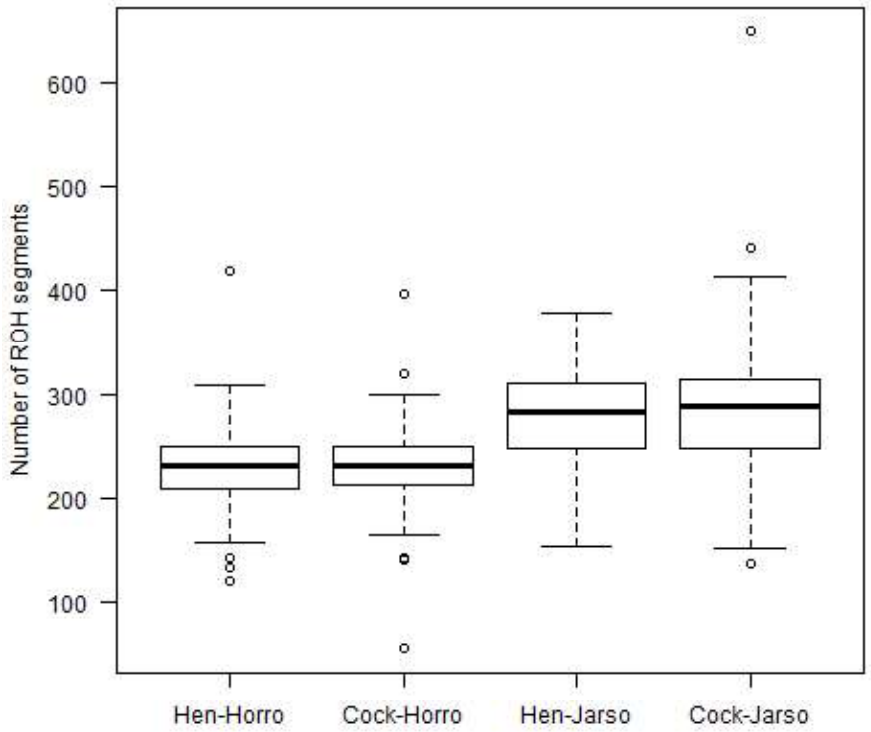


Figure S6.1. Variabilty in the number of ROH segments among hen and cock groups at population level.

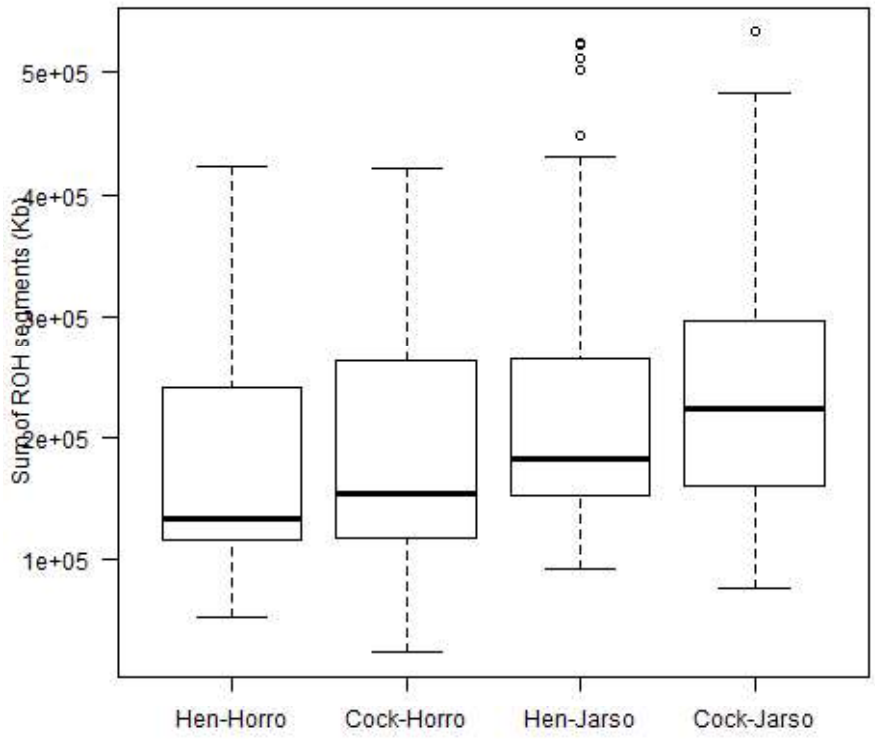


Figure S6.2. Variabilty in the sum length of ROH segments among hen and cock groups in each population.

Chromosome-wide runs of homozygosity

Chromosome-wide runs of homozygosity (ROH) summary statistics was performed for each autosome, population and sex group using SPSS 21.0.0.0 (IBM Corp., 2012) is presented in Table S6.1. A Pearson correlation test performed between chromosome-wide ROH parameters shows that both the number of ROH segments (NSEG) versus the sum length of ROH segments (KB_{CHR}) (r = 0.796, $P < 0.0001$) and NSEG versus the sum length of ROH averaged for NSEG, *i.e.* KB_{AVG} (r = 0.039, $P < 0.0001$) show positive and significant correlation. Similarly, NSEG versus the physical size of autosomes shows statistically significant positive correlation (r = 0.698, $P < 0.0001$). ROH was more frequent in bigger than smaller autosomes. KB_{CHR} show positive and significant correlation with KB_{AVG} (r = 0.325, $P < 0.0001$) and with physical size of the chromosome (r = 0.608, $P < 0.0001$). KB_{AVG} versus chromosome physical size also show positive correlation (r = 0.158, $P < 0.0001$). The mean comparison performed between the two chicken populations shows that both NSEG ($t_{21166} = 9.094$) and KB_{CHR} ($t_{21166} = 8.324$) were significantly lower in Horro than in Jarso chickens ($P < 0.0001$), however there is no significant difference

in KB_{AVG} between the two chicken populations (t_{21166} = 0.663, P = 0.507). A comparative analysis performed between hen and cock populations showed no significant difference in NSEG (t_{21166} = 0.134, P = 0.893), however, there is significant difference in KB_{CHR} (t_{21166} = 2.924, P = 0.003) and in KB_{AVG} (t_{21166} = 4.608, $P < 0.0001$). Both KB_{CHR} and KB_{AVG} were more extensive (*i.e.* ROH burden) in cock than hen population. Analysis of variance shows that chromosome physical size has statistically significant effect on NSEG ($F_{(27, 21140)}$ = 5790.406, $P < 0.0001$), KB_{CHR} ($F_{(27, 21140)}$ = 1593.108, $P < 0.0001$) and KB_{AVG} ($F_{(27, 21140)}$ = 39.408, $P < 0.0001$). ROH was relatively more frequent and more extensive in bigger autosomes, *i.e.*, macrochromosomes and the intermediate-sized autosomes compared to the microchromosomes. Summary statistics for chromosome-wide ROH for population and sex groups, and for each autosome are presented in Table S6.1 & Table S6.2 respectively.

Proportion of chromosomes having ROH tracts

Among 383 chickens included in this analysis from Horro only in three chickens ROH was detected in all the twenty eight autosomes, while 9 out of 373 Jarso chickens have ROH in all the autosomes, however, these proportions are not signficantly different (χ^2_1 = 0.2214, P = 0.3620). In Horro chickens the mean (standard deviation) number of chromosomes that have at least a single ROH was 22.97 (±2.167), while this was 24.94 (±1.685) in Jarso chickens and this shows statistically significant difference (t_{754} = 13.982, $P < 0.0001$). The NSEG moderately increases as the number of chromosomes to which ROH detected increases (r = 0.425, $P < 0.0001$) and a comparable trend was observed for KB_{CHR} (r = 0.632, $P < 0.0001$) and for KB_{AVG} (r = 0.409, $P < 0.0001$).

Table S6.1. Summary statistics for chromosome-wide ROH burden by population and sex groups.

Population	Statistics	NSEG	KB	KB_{AVG}
Horro (n = 10724)	Mean(sd)	8.23(13.091)	6494.724(12360.398)	720.324(1346.422)
	Median	3	1798.795	429.237
	Range	(0, 127)	(0, 132942)	(0, 59545.2)
Jarso (n = 10444)	Mean (SD)	9.97(14.694)	8012.345(14126.212)	731.173(1005.590)
	Median	4	2467.045	488.492
	Range	(0, 131)	(0, 136515)	(0, 15920.8)
Total (n = 21168)	Mean(SD)	9.09(13.931)	7243.497(13282.421)	725.677(1190.503)
	Median	4	2106.94	458.546
	Range	(0, 131)	(0, 136515)	(0, 59545.2)

Table S6.2. Summary for number of ROH tract detected across the twenty eight autosomes (n = 192315).

Population	Statistics	KB_{IND}	NSNP	DENSITY	PHOM	PHET
Horro(n = 88232)	Mean(SD)	789.39(1799.789)	422.94(942.601)	1.958(0.527)	0.989(0.0085)	0.007(0.0055)
	Median	358.253	193	1.962	0.991	0.006
	Range	200.005–59545.19	100–29394	0.388–11.064	0.917–1	0–0.033
Jarso(n = 104083)	Mean(SD)	803.983(1628.652)	433.79(853.62)	1.933(0.574)	0.989(0.0084)	0.0068(0.0053)
	Median	398.725	219	1.947	0.991	0.006
	Range	200.005–47854.8	100–23280	0.386–11.31	0.909–1	0–0.035
Total (n =192315)	Mean	797.288(1709.307)	428.82(895.555)	1.944(0.553)	0.989(0.0085)	0.0069(0.0054)
	Median	378.709	206	1.954	0.991	0.006
	Range	200.005–59545.19	100–29394	0.386–11.31	0.909–1	0–0.035

Absence of ROH

Absence of ROH across the autosomes (Table S6.3) was more frequent in Horro than Jarso chickens ($t_{742} = 13.912$, $P < 0.0001$), however there is no significant difference between hen and cock populations ($t_{742} = 0.059$, $P = 0.953$). Absence of ROH was not detected for GGA1–6. There is highly significant difference among the 22 autosomes (GGA7–28) for absence of ROH ($\chi_{221} = 4730.971$, $P < 0.0001$). From all the chickens included in the analysis 91.8% (694/756) do not have ROH on chromosome 16 and 65.21% (493/756) on chromosome 25. Absence of ROH is more frequent in smaller micro-chromosomes (GGA16–28) than in the intermediate-sized autosomes and bigger micro-chromosomes (GGA7–15).

Table S6.3. Summary statstics for absence of ROH in the autosomes.

Population	**Statistics**	**Autosomes lacking ROH**
Horro (n = 380)	Mean (SD)	5.07(2.129)
	Median	5
	Range	(1–15)
Jarso (n = 364)	Mean (SD)	3.13(1.634)
	Median	3
	Range	(1–11)
Total (n = 744)	Mean (SD)	4.12(2.136)
	Median	4
	Range	(1–15)

ROH islands and uniparental disomy

The number of overlapping ROH segments mapped in each autosome is presented in Table S6.4. Large numbers of overlapping regions were found in macro-chromosomes and this might be associated with their large physical size and low recombination rate. The putative isodisomy detected in the chicken population (Table S6.5) indicates the commonness of this chromosomal mis-segregation.

Table S6.4. The number of ROH tracts detected in overlapping ROH regions.

Chromosome	Pool	Overlapped segments	Horro	Jarso
1	9047	1544719	716442	828277
2	5855	978294	444899	533395
3	4997	842329	365702	476627
4	3772	648761	296405	352356
5	2756	475338	211954	263384
6	1658	253500	113437	140063
7	1578	267872	118215	149657
8	1226	188498	88982	99516
9	1184	184775	82217	102558
10	1011	142349	72305	70044
11	792	111015	51051	59964
12	942	140736	62348	78388
13	862	139238	56525	82713
14	706	108199	51476	56723
15	567	89943	42529	47414
16	1	62	35	27
17	487	66876	26657	40219
18	446	56780	25959	30821
19	453	60925	24772	36153
20	581	82723	37786	44937
21	268	28536	11441	17095
22	158	21132	7667	13465
23	242	25969	8848	17121
24	273	29848	11128	18720
25	66	5652	2325	3327
26	214	22461	8958	13503
27	175	20356	10095	10261
28	174	18361	8844	9517

Table S6.5. The summary statistics of isodisomy detected based on ROH analysis.

ID	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET	Sex
HA1B02B	5	154	59545345	59545.19	29394	2.026	0.993	0.003	F
HB2B16B	7	17045	36212750	36195.71	20729	1.746	0.995	0.003	M
HB1B22B	11	12472	19393519	19381.05	13094	1.48	0.994	0.003	M
HB2A03A	12	17210	19874086	19856.88	14013	1.417	0.995	0.003	F
HA2B07B	13	4783	17759759	17754.98	10661	1.665	0.994	0.003	M
HB1B22B	14	11989	15157385	15145.4	12391	1.222	0.994	0.003	M
HA1A08A	15	5387	12650049	12644.66	9783	1.293	0.994	0.003	F
HB2A05B	15	5387	12650049	12644.66	9783	1.293	0.995	0.002	F
JA2B17B	15	5387	12650049	12644.66	9783	1.293	0.995	0.003	M
JA2B24B	15	5387	12650049	12644.66	9783	1.293	0.996	0.002	M
HA2A22A	16	1253	496589	495.336	438	1.131	0.979	0.007	F
HB2A11A	16	1253	496589	495.336	438	1.131	0.989	0.005	F
JB1B15A	16	1253	496589	495.336	438	1.131	0.984	0.014	F
JB1A13A	20	46623	14299658	14253.04	8920	1.598	0.996	0.002	M
JB2B12B	20	46623	14299658	14253.04	8920	1.598	0.993	0.003	M
JB2B05B	21	2407	6791011	6788.604	8272	0.821	0.994	0.003	F
JB2B07A	21	2407	6791011	6788.604	8272	0.821	0.993	0.003	F
JB1B18B	22	3898	4077139	4073.241	4061	1.003	0.995	0.003	F
HA2A21B	26	1127	5321460	5320.333	5733	0.928	0.993	0.003	F

Table S6.5 cont'd

JA1B06A	26	1127	5321460	5320.333	5733	0.928	0.995	0.003	M
HB1B08B	27	2694	5205232	5202.538	5213	0.998	0.993	0.004	M
HA2A11A	28	6695	4742489	4735.794	4999	0.947	0.991	0.003	F
HB2A19B	28	6695	4742489	4735.794	4999	0.947	0.989	0.006	F

References

IBM CORP. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp, 2012.

Appendix 7

Figure S7.1a and S7.1b display observed and Gaussian distribution of integrated haplotype score in Horro and Jarso chickens respectively. Figure S7.2 and S7.3 show observed versus Gaussian distribution of Rsb score and FST respectively between Horro and Jarso chickens. Figure S7.4a & S7.4b present SweeD analysis plots produced from the smaller dataset in Horro and Jarso chickens respectively. A chain of plots that are presented in Figure S7.5a–e show the standardized varLD score in each autosome.

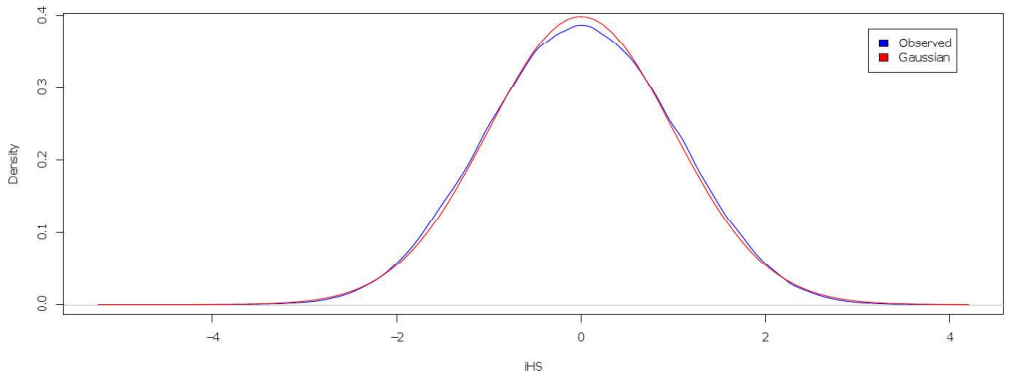


Figure S7.1a. Observed versus Gaussian distribution of iHS score in Horro chickens.

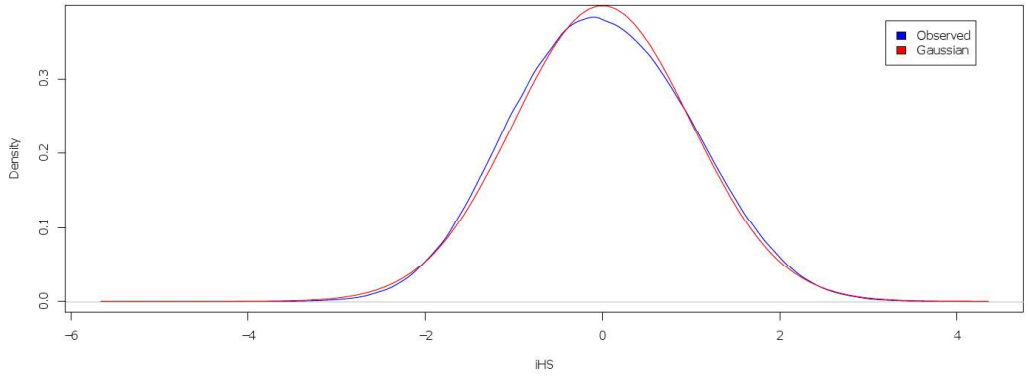


Figure S7.1b. Observed versus Gaussian distribution of iHS score in Jarso chickens.

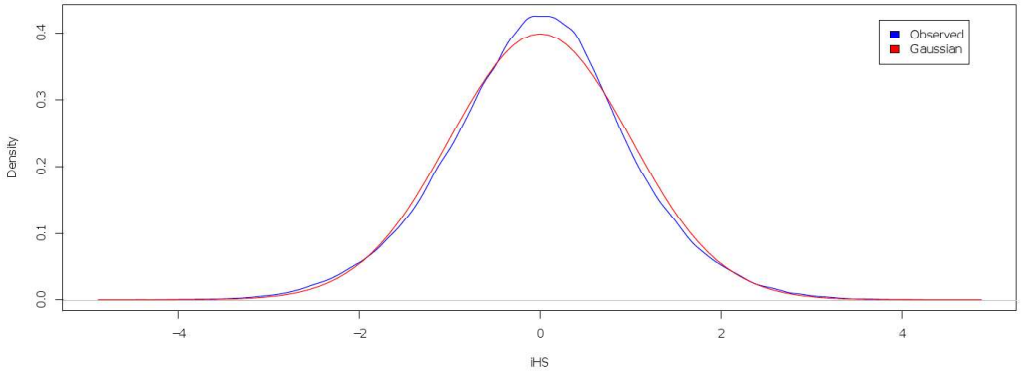


Figure S7.2. Observed versus Gaussian distribution of Rsb in Horro and Jarso chickens.

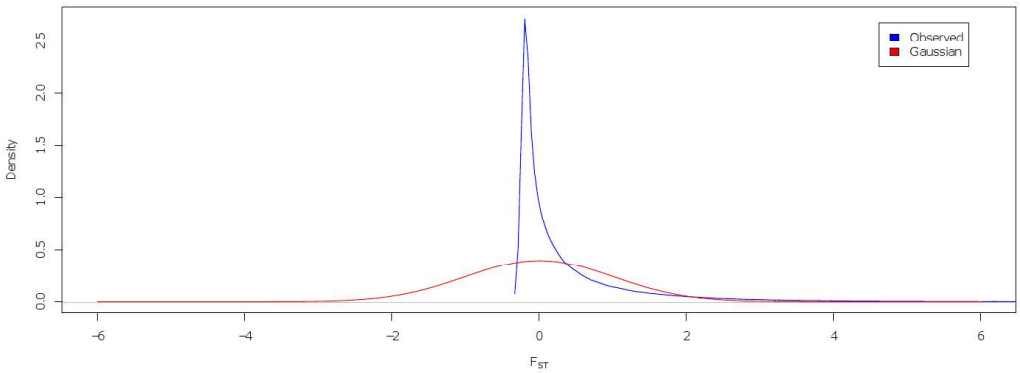


Figure S7.3. Observed versus Gaussian distribution for F_{ST} value.

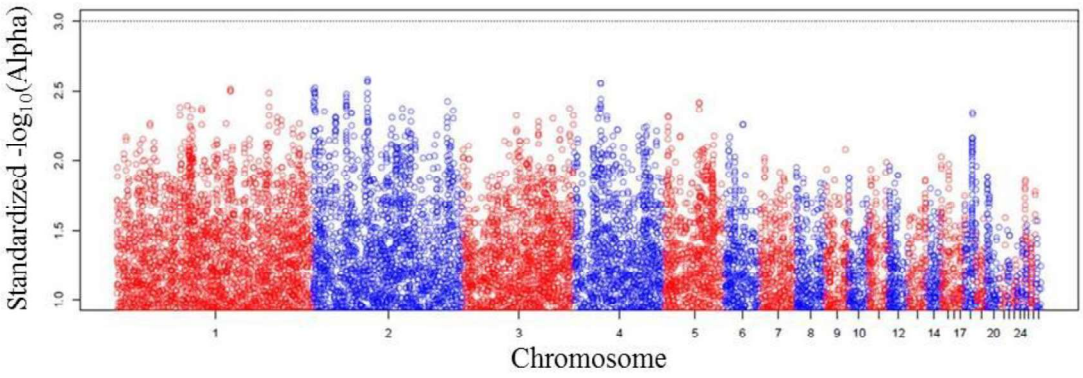


Figure S7.3a. SweeD analysis plot for the smaller dataset in Horro chickens.

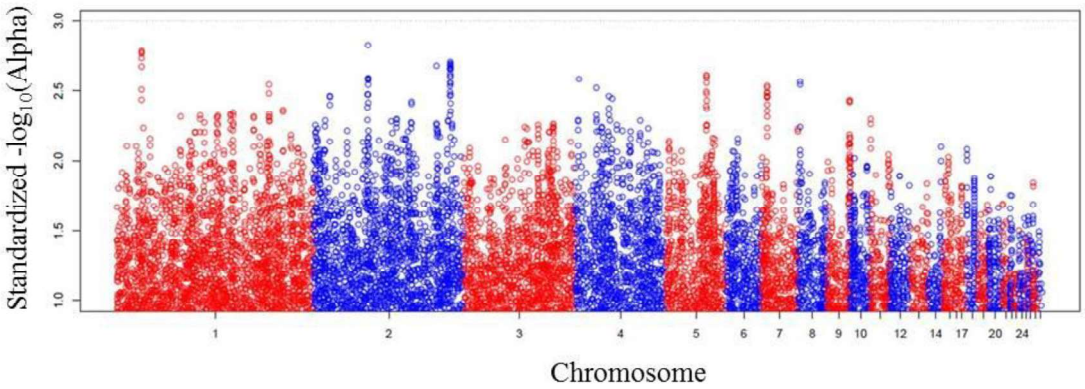


Figure S7.3b. SweeD analysis plot for the smaller dataset in Jarso chickens.

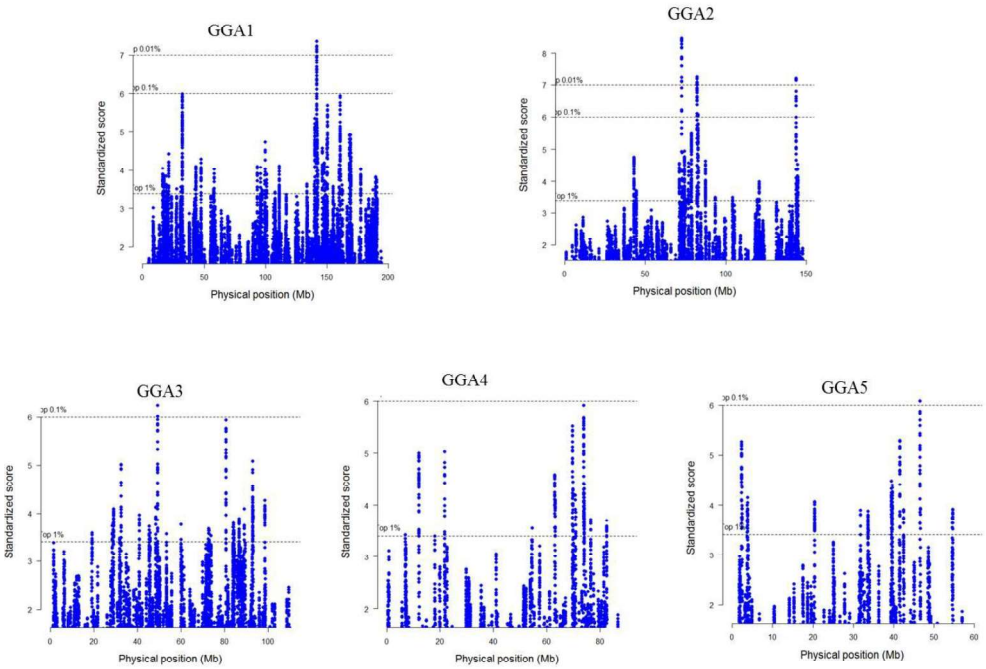


Figure S7.4a. The varLD plots for GGA1–5.

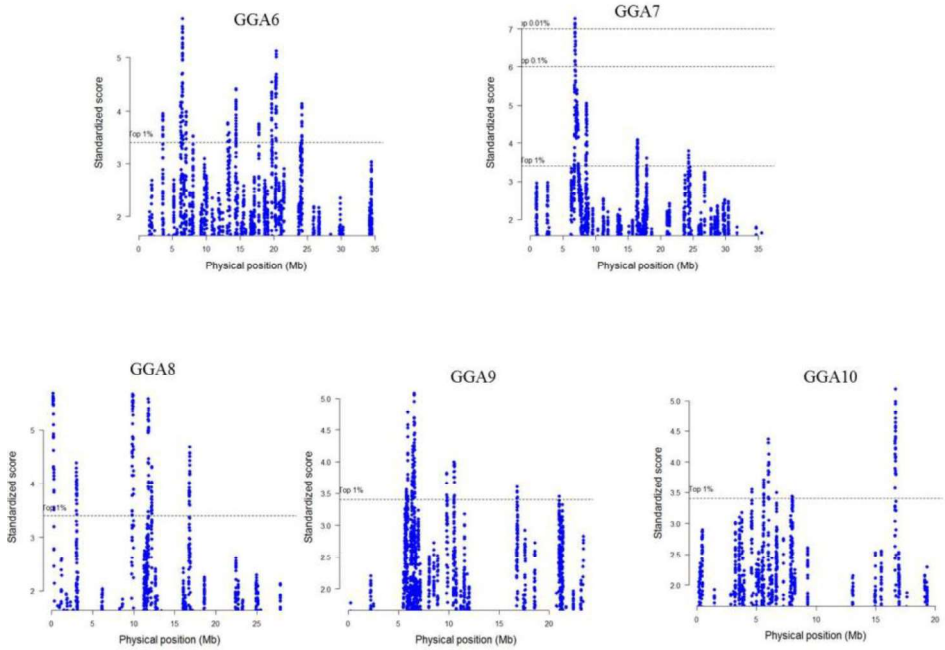


Figure S7.4b. The varLD plots for GGA6–10.

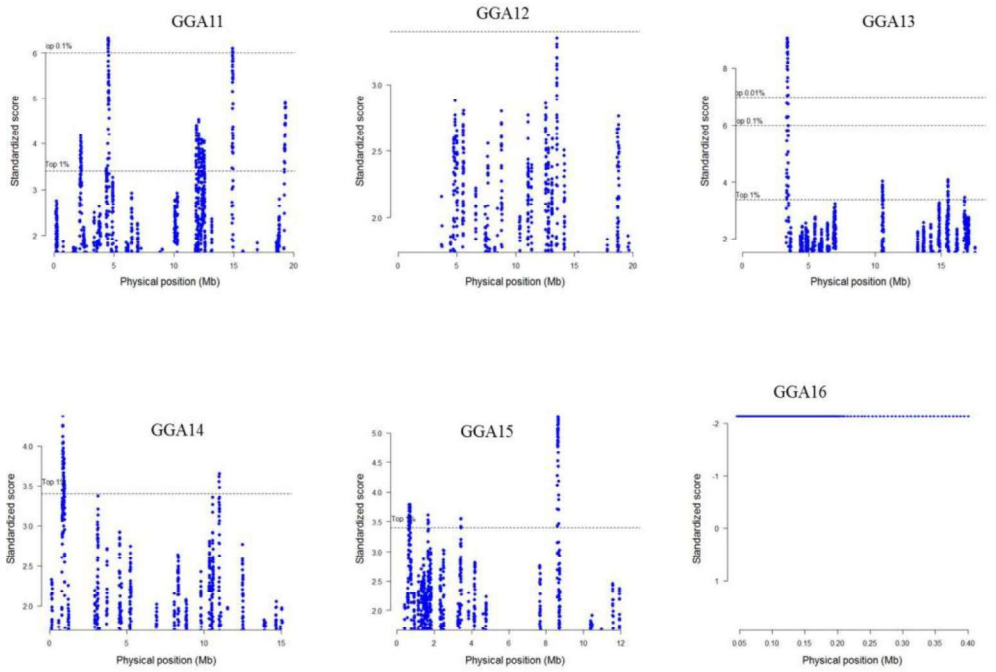


Figure S7.4c. The varLD plots for GGA11–16.

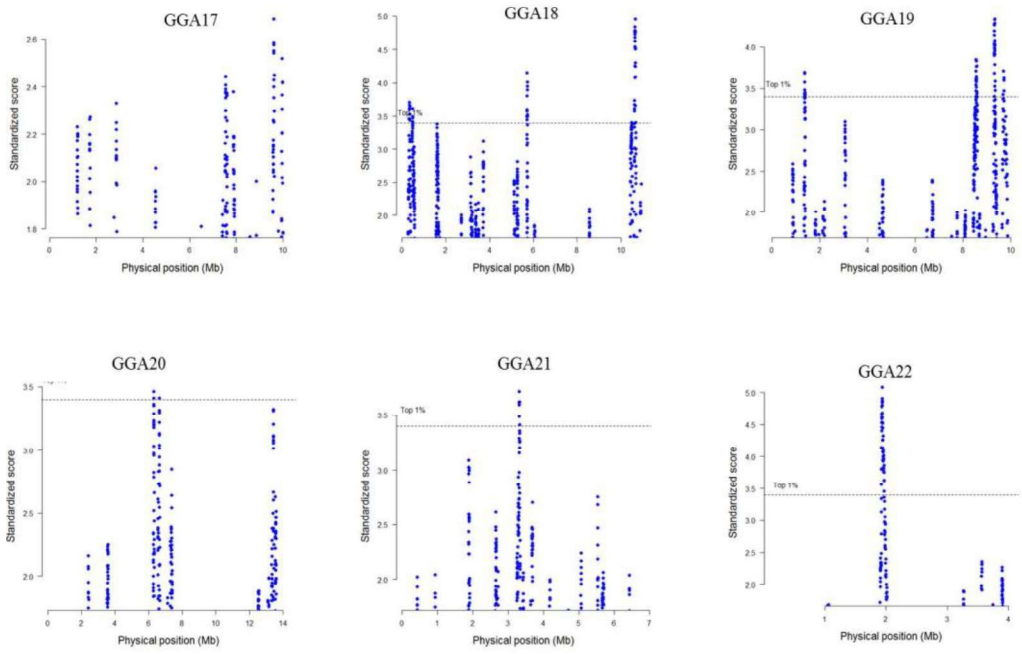


Figure S7.4d. The varLD plots for GGA17–22.

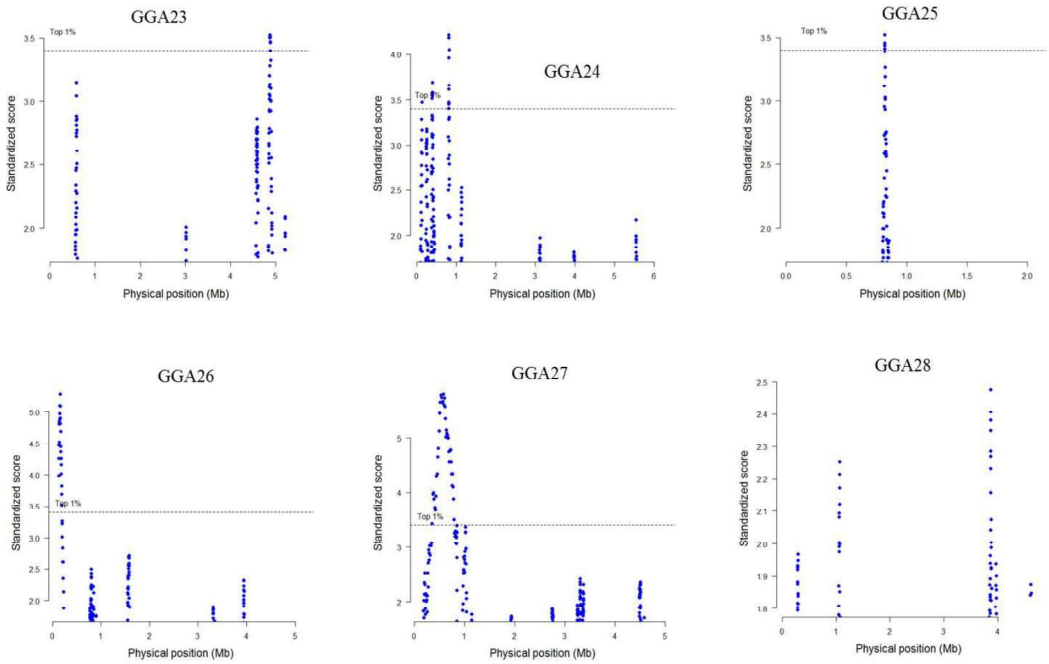


Figure S7.4e. The varLD plots for GGA23–28.