Adaptive phenotypic and genetic variation in chickens: a landscape genomics approach

Fasil Getachew Kebede

Propositions

1. Adaptive and neutral genomic variations are inseparable in Ethiopian indigenous chickens.
   (this thesis)

2. Multi-environment performance analyses (MEPA) are equally applicable in livestock as they are in crops.
   (this thesis)

3. Expectations for the development of sentient artificial intelligence (AI) are unrealistic.

4. Climate change engineering causes more harm to the ecosystem than gradual global warming.

5. Human evolution is a double-edged sword.

6. Europe has overlooked Africa as its best supplier of clean energy.

7. Africa will not grow without homegrown contemporary philosophers.

Adaptive phenotypic and genetic variation in chickens: a landscape genomics approach

Fasil Getachew Kebede
Wageningen, 29 August 2022

Adaptive phenotypic and genetic variation in chickens: a landscape genomics approach

Fasil Getachew Kebede

Thesis committee

Promotor
Prof. Dr J. Komen
Personal chair, Animal Breeding and Genomics
Wageningen University & Research

Co-promotors
Dr J.W.M. Bastiaansen
Assistant professor, Animal Breeding and Genomics
Wageningen University & Research

Dr Tadelle Dessie
Principal Scientist, Livestock Genetics
International Livestock Research Institute, Addis Ababa, Ethiopia

Other members

Prof. Dr B.J. Zwaan, Wageningen University & Research
Prof. Dr J. Sölkner, University of Natural Resources and Life Sciences, Austria
Dr B.J. Zumbach, University of Göttingen, Germany
Prof. Dr F.A. van Eeuwijk, Wageningen University & Research

Adaptive phenotypic and genetic variation in chickens: a landscape genomics approach

Fasil Getachew Kebede

Thesis

Abstract

Kebede, F.G. (2022). Adaptive phenotypic and genetic variation in chickens: a landscape genomics approach

Smallholder chicken production is an integral part of tropical farming systems and contributes significantly to sustainable livelihoods. Performance of chickens in these systems is too low to meet the growing demands for meat and eggs. Unavailability of productive and adaptive breeds is a major constraint. Knowledge on phenotypic and genetic variation among populations contributes to the design of sustainable chicken genetic improvement and development programmes. I follow a landscape genomics approach and integrate genetic, phenotypic, and environmental information in my study design. In the first part of this thesis, I aim to identify the environmental drivers of local adaptation and detect genomic footprints of natural selection in indigenous chickens. I use species distribution models (SDMs) to identify environmental predictors associated with habitat suitability. Based on higher level of matching between the presence of distinct phenotypes and availability of unique environmental niches, I classify the Ethiopian chicken populations into three ecotypes. I perform selection signatures analyses (F_{ST} and XP-EHH) and redundancy analyses (RDA) at different analytical layers (considering gradient and agroecology) to identify candidate loci and genomic regions linked mainly with local adaptation. I show that Ethiopian chicken populations differentiated the most between gradients but selection pressures leading to adaptive variation are stronger between agroecologies. I indicate that the results from RDA match the outputs from signatures of selection analyses (F_{ST} and XP-EHH). I show that RDA can be used as an alternative approach to GWAS in random mating, indigenous chicken populations which have sufficiently interacted with the environment. I demonstrate that signatures of selection analysis with the two methods (F_{ST} and XP-EHH) can be used complementarily with RDA to shed light on the relationship between genomic, phenotypic, and environmental variation in local adaptation studies in indigenous chickens. I show that phenotypic distribution models (PDMs) such as boosted generalized additive models (GAMs) are valuable tools in animal breeding to integrate environmental and phenotypic information and to predict phenotypic values. In the second part of the thesis, I evaluate the performance of improved chicken breeds introduced into smallholder systems. I show that that agroecologies defined by SDMs improve model fit in GxE predictions. I utilize the concept of phenotypic plasticity to compare yield stability of improved chicken breeds. I show that two approaches of multi-environment breed performance analysis (MEPA), namely, additive main effects and multiplicative interaction models (AMMI) and linear mixed-effects models (LMM) are applicable in chicken to identify and recommend more productive and stable breeds. Together, I demonstrate how adaptive phenotypic and genetic variation can be exploited to enhance the performance of chickens in smallholder systems.

Contents

1

General introduction

1.1 Smallholder chicken production

Smallholder chicken production system is also known as free-range, traditional, or backyard production system. The system is mainly characterized by small flock size (less than 50 birds), semi-scavenging to meet part of the daily feed requirement, limited access to veterinary health services, use of simple poultry houses, and use of mostly unimproved breeds (Sonaiya and Swan, 2007). Smallholder chicken production contributes significantly to income generation, food security, ecosystem services, and cultural values in developing tropical regions of the world. Although commercial poultry production is rapidly expanding in these regions, smallholder systems are still dominant.

1.2 Indigenous Ethiopian chickens and their productivity

Ethiopia has one of the earliest evidences for chicken domestication and dispersal in Africa (Woldekiros and D'Andrea, 2017). The main reason for this is the geographic proximity of the country to the Middle East, along the trade routes to Southeast Asia, where domestication of chickens (*Gallus gallus*) happened. Three wild species in the genus Gallus (Grey, Cylon, and Sri Lankan jungle fowls) contributed to the genetic background of domestic chicken or Red jungle fowl (Delacour, 1977). The largest introgression of chickens into Ethiopia was from Grey jungle fowl (*Gallus sonnerati*), followed by Sri Lankan jungle fowl (Lawal et al., 2020). Ethiopian chickens are distributed in all agroecologies and show substantial phenotypic and genetic diversity (Adebabay, 2018; Dana, 2011; Dessie, 2003a; Hassen et al., 2009; Mwacharo et al., 2011). Large genetic diversity of present-day Ethiopian chicken populations might be attributed to their multiple waves of introduction into the country (Lyimo et al., 2014; Mwacharo et al., 2013) and the presence of highly diverse environment (e.g., climate, vegetation, elevation) (Billi, 2015). As such, the country is generally considered as an ideal place for studying adaptive phenotypic and genetic variation in livestock.

Smallholder chicken production is an integral part of the Ethiopian farming system. About 81.7% of the country's estimated 49 million chickens are indigenous, 10.9% are hybrids, and 7.4% are exotic (CSA, 2020). The contribution of the commercial poultry sector is increasing (16.52% to meat production and 25% to egg production), currently comprising about 71,000 broiler breeder stock, 87,300-layer breeder stock and 141,700 dual-purpose parent stock (ENTAG, 2020). With regard to the flock composition and size of smallholder systems, a large proportion (37.9% or 22.7 million) are chicks, 33.6% (20.2 million) are laying hens, with eighty percent of the poultry raising households in the country keeping less than 9 birds (FAO, 2019a). Performance levels of indigenous chickens are low, with mean body weight at 16-

weeks-of-age as low as 621 grams and mean annual egg production below 60 eggs/hen (Dana, 2011; FAO, 2019b; Getachew et al., 2016). Demand for animal protein, on the other hand, is increasing rapidly. The estimated increase in egg consumption between 2005 and 2015 was 26 percent in sub-Saharan Africa, compared to only 2.4 percent in developed countries (Windhorst, 2008). Closing the projected gaps in total egg and meat consumption in countries like Ethiopia requires transforming both the smallholder and commercial poultry systems (Shapiro et al., 2015).

1.3 Enhancing smallholder chicken productivity with improved breeds

Attempts to genetically improve the productivity of smallholder chickens in Africa in the last seventy years mostly focused on the introduction of improved breeds. These attempts did not bring the desired outcome in terms of increasing egg and meat productivity. Some of the reasons for their limited success include: lack of proper matching of genotypes with environments, underdeveloped input and service delivery systems, and unavailability of breeds which are well adapted to the existing production conditions. The technological landscape (e.g. in terms of availability of vaccines and better-quality feeds) in low- and medium-income tropical countries such as Ethiopia is changing positively to justify the introduction of high producing dual-purpose breeds which can meet part of their daily feed intake requirements from scavenging (Ahuja et al., 2008b; Birhanu et al., 2021; Birhanu et al., 2022; Mebrahtu, 2017; Nasser et al., 2000; Singh et al., 2017). Multi-national chicken breeding companies are also showing interest in developing dual-purpose improved breeds suitable to smallholder systems (Duijvesteijn and Perrault, 2019; IPP, 2018).

There is considerable preference by farmers for high producing, dual-purpose tropically adapted chicken breeds, which perform better than local chickens in the smallholder poultry systems (Alabi et al., 2020; Birhanu et al., 2022). On the other hand, lack of access to productive and adaptive chicken breeds is becoming a critical challenge to increase the economic contribution of the sector (Birhanu et al., 2022; FAO, 2019b). Moreover, it is not likely that a single adaptive chicken breed will be suitable for all the different environments.

There is a strong need to evaluate candidate breeds in different environments and identify the best ones for wider use by smallholder farmers. Recently, different tropically adapted improved breeds or strains (Kuroiler, Sasso, SRIR, Potchefstroom Koekoek, and Improved Horro) were compared for their performance in different agroecologies of Africa (Alemu et al., 2021; Lozano-Jaramillo et al., 2019b; Passarelli et al., 2020). These studies showed the performance of improved varied across

agroecologies, with the interventions leading to positive outcomes on livelihoods. In addition some limitations were observed in the evaluation of chicken breeds in smallholder systems. The existing agroecological classes are defined based on environmental predictors influencing crop productivity and are not suitable to predict genotype by environment interactions (GxE) in livestock. The lack of proper agroecological definition for chickens impedes proper evaluation and identification of breeds with superior productivity and wider adaptability. The use of livestock-species-specific agroecological classes may improve model fit in GxE predictions compared to predictions based on conventional agroecologies. The common approach in estimating GxE in livestock considers performance testing environment (agroecology) as a homogenous group and compares breeds (e.g., B1, B2) across environmental classes (e.g., E1, E2) without paying attention to specific climatic predictors within each class that have an influence on phenotypes. Changes in rank or magnitude of performance between breeds based on reaction norm plots are then taken as an evidence for presence of GxE. While this approach ranks these specific breeds for the specific environmental classes, the results cannot be generalized to other environments that are not part of the study. Unavailability of appropriate analytical frameworks is another factor limiting the identification of the best performing and adaptive breeds. On-farm experimental designs and statistical methods have not been tested for their efficiency to make multi-environment breed performance comparisons in the context of smallholder systems.

In this thesis, I will use redefined chicken production agroecologies to compare breeds on growth performance and stability. In addition to redefining agroecologies into distinct environmental classes based on their influence on chicken performance, I incorporate the most important climatic predictors (i.e. having influence on habitat suitability of chickens and their phenotypic variation) to improve model fit in GxE predictions. Instead of reaction norm plots, I use machine learning based partial dependency plots (PDPs) as a visual tool to explain the relationship between specific climatic predictors and a trait in a continuum. The relationship between a climatic predictor and a trait may be complex (i.e. non-linear) and often cannot be explained by linear models (Zuur et al., 2007).

1.4 Novel approaches to studying adaptive genetic variation

Chickens kept in tropical smallholder systems are constantly subjected to environmental stressors. These stressors include extremes of temperature, solar radiation, and relative humidity; wide ranges of elevations; fluctuating feed availability; and high prevalence of diseases and parasites. Studying how the

indigenous livestock have evolved under environmental challenges will give a better understanding of molecular mechanisms and selective processes behind local adaptation (Bettridge et al., 2018; Fleming et al., 2017; Porto-Neto et al., 2014).

The evolutionary mechanisms of local adaptation for quantitative traits in Ethiopian chickens are not fully understood. Genomic studies conducted to date on these chickens (Adebabay, 2018; Bettridge et al., 2018; Desta, 2015; Gheyas et al., 2021; Hassen et al., 2009) made important scientific contributions to knowledge. They have shown that identification of key environmental drivers should precede genomic investigations to elucidate adaptive variation in domestic animals. The studies had also certain limitations that can be overcome through a more refined research methodology. For instance, the studies conducted by Desta (2015) and Bettridge et al. (2018) covered too few sample locations to capture existing genetic variation among Ethiopian chickens. Studies conducted by Adebabay (2018) and Gheyas et al. (2021) were more comprehensive and the authors did integrate environmental and genomic information to detect adaptive genetic variation. However, even in their studies, quantitative traits were not recorded on the sample chicken populations, and it was not possible to relate phenotypic variation with environmental and genomic variation. Moreover, the sampling strategy in Adebabay (2018) and Gheyas et al. (2021) did not consider environmental gradation. In this thesis, I combine recently developed techniques to study adaptive genetic variation in chickens, including landscape genomics, signatures of selection, and association analyses.

1.4.1 Landscape genomics

Identification of genomic regions which have undergone selection is one of the principal goals of theoretical and applied evolutionary genetics (Gouveia et al., 2014). Landscape genomics offers a unique opportunity to integrate phenotypic, genomic, and environmental information to understand adaptive genetic variation. It is a powerful tool to identify environmental variables related with phenotypic differentiation and to predict performances of breeds in different agroecologies. Landscape genomics seeks to understand the influences of geographic and environmental features on selectively neutral and adaptive loci, and underlying micro-evolutionary processes such as gene flow, selection, and genetic drift (Balkenhol et al., 2016; Storfer et al., 2018).

Advances in remote sensing and geographic information system (GIS), better availability of environmental and genomic data, and improvements in computational power facilitate wider application of landscape genomics in livestock. Studies which followed this approach in studying adaptive genetic variation in farm animals include Joost et al. (2007), Pariset et al. (2012), Mdladla (2016) Roffler et al. (2016), Li et al.

(2017), Cesconeto et al. (2017), Vallejo-Trujillo et al. (2018), and Gheyas et al. (2021). Landscape genomics approaches can be followed to detect signatures of selection and understand the drivers of local adaptation.

1.4.2 Signatures of selection analysis

Local adaptation refers to the response of individuals to differential selective pressure leading to higher genetic fitness in their environment than individuals from elsewhere (Kawecki and Ebert, 2004; Savolainen et al., 2013). Phenotypic and genetic differentiation along environmental gradients, or across contrasting habitat types, can be indicative of local adaptation (Conover et al., 2009; Savolainen et al., 2013; Turesson, 1922).

Positive selection leaves conspicuous footprints or selective sweeps on the genome that can be detected using several approaches (Pavlidis and Alachiotis, 2017). Cross-population Extended Haplotype Homozygosity (XP-EHH) detects differential selection between two populations (Sabeti et al., 2007). Pairwise comparison of fixation index (F_{ST}) reveals differentiation of populations in different environments due to differences in evolutionary history (Nei, 1986). The choice of tools to detect signatures of selection depends on many factors including the nature of selective sweeps we would like to investigate, the time scale over which the selection occurred, and the level of variants in which we are interested in.

1.4.3 Redundancy Analysis (RDA)

One of the most popular statistical methods used to investigate association between genomic and environmental variability is Redundancy Analysis (RDA). RDA combines regression and principal component analysis (PCA) and it is an extremely powerful tool for ecologists to model multivariate response data (Borcard et al., 2011; Legendre and Gallagher, 2001). RDA determines how groups of loci covary in response to the multivariate environment, and can better detect processes that result in weak, multilocus molecular signatures relative to univariate tests (Rellstab et al., 2015). It accounts for population structures, demographic histories, and polygenic interactions (Capblancq et al., 2018; Forester et al., 2018).

Multivariate methods like RDA, that simultaneously account for multiple drivers of phenotypic and environmental divergence, are recently being applied in landscape genomic studies to identify quantitative trait loci (QTL) associated with environment predictors (Forester et al., 2018; Harrisson et al., 2017; Kess and Boulding, 2019; Torrado et al., 2020).

1.5 Novel approaches to studying adaptive phenotypic variation

Phenotypic variation for a trait is influenced by several evolutionary processes and phenotypically plastic genotypes tend to show wide phenotypic variation (Le Corre and Kremer, 2012; Schmid and Guillaume, 2017). Genotypes that show highly variable phenotypes across environments are 'plastic', whereas genotypes that show little variability across environments are 'robust' (De Jong and Bijma, 2002). Phenotypic plasticity is the ability of a given genotype to express different phenotypes in different environmental circumstances. The term 'phenotypic plasticity' was coined by the Swedish biologist Herman Nilsson-Ehle to describe a phenomenon where a single individual produces multiple phenotypes in direct responses to different environmental circumstances (Pfennig, 2021). In case of phenotypic plasticity a single genotype expresses different phenotypes (e.g., morphological, physiological, behavioural changes) in different environments (Bradshaw, 1965; Scheiner, 1993). A plastic response is evolutionarily 'adaptive' when it improves the fitness of the individual and 'maladaptive' otherwise (Ahlgren et al., 2013; Hendry, 2016). GxE variation comprises the extent to which genotypes differ in their plastic responses to environmental changes for a given trait. The term "reaction norm" (*Reaktionsnorm*) was coined by Woltereck (1909) to describe the effect of environment on phenotypic variation of a given genotype (Costa, 2021). In animal breeding, if different genotypes have different reaction norms associated with plasticity of the trait, then there is GxE interaction (Falconer, 1996). Populations differ in their patterns of GxE interaction (phenotypic plasticity) depending on previous selection, drift, and mutation (Colautti et al., 2017). I apply recently developed techniques to study adaptive phenotypic variation in chickens, including distribution models and multi-environment performance testing.

1.5.1 Species and phenotypic distribution models

Species distribution models (SDMs), also known as environmental (ecological) niche models (ENMs) or habitat distribution models (Elith and Leathwick, 2009), use computer algorithms to analyse environmental data and to predict the distribution of a species across geographic space and time. SDMs are a popular tool in quantitative ecology because of their comparably low data requirement, availability of many software packages and guidelines, and their higher predictive abilities (Elith and Franklin, 2013; Guisan et al., 2017).

The central concept in SDMs is the niche theory formulated by Joseph Grinell and G. Evelyn Hutchinson (Soberón, 2007). Two types of niches are identified by (Hutchinson, 1957). The fundamental niche comprises all abiotic environmental

conditions where a species can survive indefinitely and have a positive population growth. The realized niche is smaller and refers to those parts of the fundamental niche where the species can survive despite the presence of competitors or negative interactions. In recent years, the conceptual framework for SDMs has been extended by livestock scientists and used to identify environmental predictors associated with habitat suitability and local adaptation in livestock populations (Gheyas et al., 2021; Kebede et al., 2021; Lozano-Jaramillo, 2019; Vajana et al., 2018; Vallejo-Trujillo et al., 2018).

Phenotypic distribution models (PDMs) use associations between phenotypes and environmental variables to map the phenotypes of populations within that species' distribution (Michel et al., 2017). PDMs refer to a family of machine learning based models (PDMs) that are fitted to capture the response of a quantitative trait as a function of environmental conditions. The most notable ones include boosted regression trees (Elith et al., 2008), boosted generalized linear models (Nelder and Wedderburn, 1972), and boosted generalized additive models (GAMs) (Hastie and Tibshirani, 1990). PDMs were used to estimate chicken growth performance at present environmental conditions (Lozano-Jaramillo et al., 2019a) and predict fish morphology in future climatic scenarios (Michel et al., 2017).

The outcomes of species distribution models can be integrated into phenotypic distribution models in studying phenotypic response at present and in future climatic scenarios (Smith et al., 2017). SDM-identified environmental predictors can be incorporated into PDMs to improve model fit in GxE predictions. GAMs relax the assumptions of linear models and achieve acceptable goodness of fit to model phenotypic distribution in response to environmental variables. The assumptions of classical statistical approaches such as Generalized Linear Models (GLM) are violated when responses are non-linear, variances change with predictors, or ecological processes operate at spatio-temporal scales (Bolker et al., 2013; Wiley and Wiley, 2019; Zuur et al., 2009) which GAMS can accommodate.

1.5.2 Multi-environment performance analysis (MEPA)

Analytical frameworks for multi-environment livestock breed performance evaluation in smallholder (extensive) systems are not readily available. Experimental designs and methods used in plant breeding to predict GxE and test yield stability are, in theory, applicable to evaluate livestock breed performance across different agroecologies. Stability refers to the ability of an individual to maintain constant yield across environments (Becker, 1981). Commonly used methods for multi-environment performance analysis (MEPA) in plant breeding include Additive Main effects and Multiplicative Interaction (AMMI) (Gauch Jr, 2013) and Linear Mixed-

effect Models (LMMs) (Piepho, 1994; Piepho and Möhring, 2005). The AMMI model is a robust multi-environment analytical technique which combines fixed-effects ANOVA with principal component analysis (PCA) to study GxE interactions (Rincent et al., 2019). LMMs are run through REML/BLUP and are reputed for higher predictive accuracy in GxE estimations (Piepho, 1994; Van Eeuwijk et al., 2016). It is imperative to investigate the applicability of different MEPA tools in identification of stable and productive livestock breeds.

1.6 Objectives and thesis outline

A solid understanding of adaptive phenotypic and genetic variation is essential to improve the productivity and resilience of livestock populations at present and help respond to future production scenarios. Knowledge on adaptive phenotypic and genetic variation among chickens kept in smallholder systems contributes to enhancing their productivity (**Figure 1.1**).
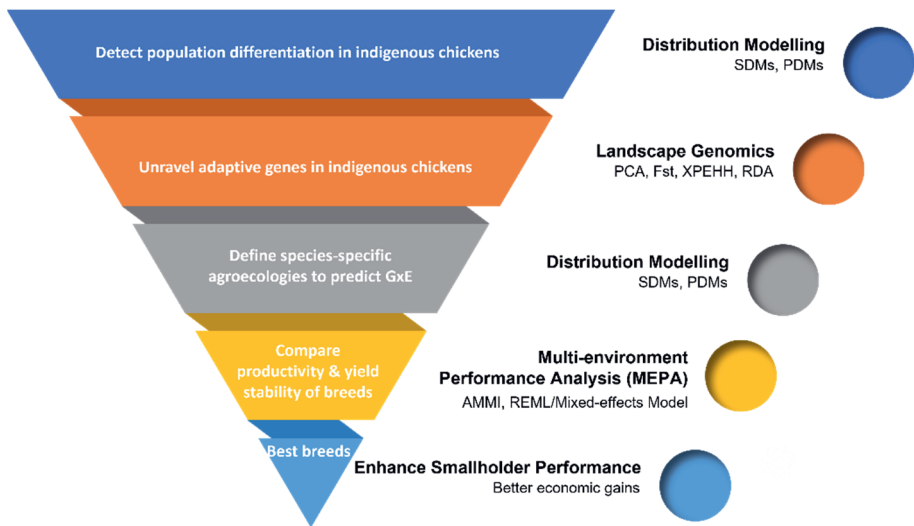
Figure 1.1. Schematic diagram explaining the conceptual framework of this thesis.

In this thesis, I integrate phenotypic, genomic, and environmental information to answer fundamental and practical questions of selective and evolutionary adaptation in chickens. The goal of the thesis is to generate knowledge and insights on adaptive phenotypic and genetic variation that will be useful to enhance chicken productivity in smallholder systems. More specifically, it addresses two research questions:

1) What are the environmental drivers of local adaptation, and phenotypic and genetic differentiation in indigenous chickens? and

2) How do improved chicken populations that are introduced into smallholder systems respond phenotypically to environmental variations?

To answer these two research questions, I shall study two types of chicken populations, namely, indigenous chicken populations and improved chicken breeds.

The objectives of this thesis are:

Part-I

1) to identify the most important environmental predictors contributing to habitat suitability and evaluate their relationships with phenotypic variability among indigenous and improved chicken populations;

2) to identify candidate genes and genomic regions related with local adaptation and phenotypic variation in Ethiopian indigenous chickens;

Part-II

3) to test applicability of SDM-identified environmental predictors to classify environmental variation at chicken performance testing sites into distinct agroecologies and improve model fit in GxE predictions;

4) to integrate information on agroecology, breed, and environmental predictors through phenotypic distribution models to evaluate growth performance and GxE; and

5) to evaluate two commonly used methods of multi-environment performance analysis (MEPA) for their applicability in livestock breed performance comparisons in smallholder systems.

Part I: Adaptive phenotypic and genetic variation among Ethiopian indigenous chickens.

In the first part of the thesis (chapters 2 and 3), I investigate the influences of the environment on adaptive phenotypic and genetic variation using indigenous chicken sample populations. Nondescript random mating indigenous (local) chicken populations have experienced the influences of selective and evolutionary processes (natural selection, genetic drift, gene flow) for many generations and meet the assumptions of statistical models that will be used to study local adaptation. In

chapter 2 of the thesis, I perform distribution modelling and correlative analysis to identify environmental variables associated with habitat suitability and phenotypic differentiation in Ethiopian indigenous chickens. In chapter 3, I follow a landscape genomic approach to search for candidate genes and regions under positive selection among populations sampled from different environmental gradients.

Part II: Adaptive phenotypic variation among improved chicken breeds introduced to smallholder farmers.

In the second part of the thesis (chapters 4 and 5) I evaluate the influences of the environment on phenotypic variation among improved chicken breeds (strains) distributed to smallholder farmers. The improved chickens used in the present study were artificially selected for specific traits (growth and egg productivity). These chickens are more useful to fit models aiming at evaluating the effects of the environment on the performance of distinct breeds. In chapter 4, I apply distribution models to integrate environmental and phenotypic information and improve prediction of breed by environment interactions for improved chicken breeds introduced into smallholder systems of Ethiopia. I test the hypothesis that classifying agroecologies considering environmental variables associated with habitat suitability and phenotypic differentiation of a livestock species improves model fit in GxE predictions. In chapter 5, I extend the concept of phenotypic plasticity to compare productive performance and yield stability among improved chickens distributed to smallholder farmers. In Chapter 6, 'General discussion', I present an overview of the significant findings from the above studies and relate them to the existing knowledge. I show implications of studying adaptive phenotypic and genetic variation with novel approaches to enhancing smallholder chicken productivity. I indicate limitations of my studies and suggest recommendations for further research.

2

Species and phenotypic distribution models reveal population differentiation in Ethiopian indigenous chickens

Fasil Getachew Kebede[1,2], Hans Komen[1], Tadelle Dessie[2], Setegn Worku Alemu[2], Olivier Hanotte[2,3], John Bastiaansen[1]

[1]Wageningen University & Research, Animal Breeding and Genomics, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands; [2]International Livestock Research Institute (ILRI), P. O. Box 5689, Addis Ababa, Ethiopia; [3]Cells, Organism and Molecular Genetics, School of Life Sciences, University of Nottingham, NG7 2UH, Nottingham, UK.

Abstract

Smallholder poultry production dominated by indigenous chickens is an important source of livelihoods for most rural households in Ethiopia. The long history of domestication and the presence of diverse agroecologies in Ethiopia create unique opportunities to study the effect of environmental selective pressures. Species distribution models (SDMs) and Phenotypic distribution models (PDMs) can be applied to investigate the relationship between environmental variation and phenotypic differentiation in wild animals and domestic populations. In the present study I used SDMs and PDMs to detect environmental variables related with habitat suitability and phenotypic differentiation among nondescript Ethiopian indigenous chicken populations. Thirty-four environmental (climatic, soil, vegetation) and 19 quantitative traits were analyzed for 513 adult chickens from 26 populations. To have high variation in the dataset for phenotypic and ecological parameters, animals were sampled from four spatial gradients (each represented by 6-7 populations), located in different climatic zones and geographies. Three different ecotypes are proposed based on correlation test between habitat suitability maps and phenotypic clustering of sample populations. These specific ecotypes show phenotypic differentiation, likely in response to environmental selective pressures. Nine environmental variables with the highest contribution to habitat suitability are identified. The relationship between quantitative traits and a few of the environmental variables associated with habitat suitability is non-linear. Our results highlight the benefits of integrating species and phenotypic distribution modelling approaches in characterization of livestock populations, delineation of suitable habitats for specific breeds, and understanding of the relationship between ecological variables and quantitative traits, and underlying evolutionary processes.

Keywords: indigenous chickens, local adaptation, agroecology, species distribution models (SDMs), phenotypic distribution models (PDMs), phenotypic differentiation, ecotype, environment, breeding programmes

2. 1 Introduction

Smallholder farmers in Africa keep scavenging poultry as a source of affordable animal protein and a means of income. The sustainability of this type of poultry production in tropical low-and medium-input systems depends on the availability of adaptive genotypes that can produce and thrive under adverse conditions such as climatic extremes, high prevalence of tropical diseases and parasites, and periodic feed shortage. The presence of selective pressures in these environments has led to adaptation of indigenous chicken populations to production constraints (Bettridge et al., 2018).

Local adaptation refers to local individuals having higher fitness in their environment than individuals from elsewhere (Williams, 1966). Environmental heterogeneity is known to be one of the main drivers of within species diversity and local adaptation (Darwin et al., 1858). Understanding the drivers of local adaptation provides essential information for designing research and development programmes aiming at improving productivity while retaining resilience. A starting point in genetic improvement of the existing local chicken populations or in considering the introduction of new genotypes is to understand how the environment is driving local adaptation (Bettridge et al., 2018). This knowledge would allow breeding of indigenous ecotypes that are more productive under village conditions while retaining locally acceptable morphological and adaptive traits (Birhanu et al., 2021; Dana et al., 2010; Muchadeyi and Dzomba, 2017).

Present day African chickens are a result of an intricate interplay between domestication and natural selection. Ethiopia is an ecological microcosm of Africa, with a rich geomorphology, where people closely interacted with the environment and practiced agriculture for millennia. Because of its cultural diversity, geographical position, complex topography, and varying climatic patterns, the country harbours rich domestic animal biodiversity. The earliest osteological evidence of domestic chicken in Africa (921-801 BCE) was recovered from Ethiopia (Woldekiros and D'Andrea, 2017).

Recent technological advances in remote sensing and GIS, increased availability of environmental data, and improved computational power facilitate the understanding of the selective forces associated with local adaptation. SDMs, implemented in MaxEnt (Phillips et al., 2006) and similar software, predict

distribution of a species based on presence-only data, estimate the contribution of environmental variables, and help identify suitable habitats in current and future environments. Lozano-Jaramillo et al. (2019b) applied SDMs to produce suitability maps of Ethiopian chickens and identify important environmental variables associated with habitat suitability in chickens, without relating ecological differences with phenotypic variation among study populations. When used alone, SDMs treat a species as an evolutionarily homogenous entity and fail to consider possible population differences pertaining to local adaptation (Hampe, 2004). SDMs also make assumptions in their modelling approach (Wiens et al., 2009) which necessitate their combined use with additional approaches, such as phenotypic distribution models (PDMs).

Phenotypic Distribution Models (PDMs) use associations between phenotypes and environmental variables to map the phenotypes of populations within that species' distribution (Michel et al., 2017). These phenotype-environment associations (PEAs), are well documented for natural populations of several wild plant and animal species (Bergmann, 1848; Cain and Sheppard, 1954; Clausen et al., 1940; Langerhans, 2008; Maloney et al., 2012; Mayr, 1942; Michel et al., 2017; Phillimore et al., 2010; Smith et al., 2017) and can be applicable to predict phenotype distribution among domestic animals.

Phenotypic differentiation represents the fraction of phenotypic variance between populations over the total phenotypic variance and helps understand evolutionary processes shaping populations (Leinonen et al., 2006; Schmid and Guillaume, 2017; Storz, 2002). With the exception of (Lozano-Jaramillo et al., 2019a) who applied PDMs to predict performance of improved chicken strains, distribution models have seldom been applied in indigenous livestock to identify environmental factors associated with phenotypic differentiation and to define their ecotypes. In contrast to introduced strains which have been subjected to intense artificial selection in a relatively short period of time, indigenous populations have been exposed to natural selection over multiple generations which permits a better understanding of evolutionary processes. Even with natural populations of animals, correlation between a phenotype and environment could be spurious if PDM are used on their own (Etterson and Shaw, 2001; Michel, 2011; Michel et al., 2017) and this requires their combination with additional analytical approaches, such as SDMs.

To overcome possible limitations in the use of SDMs in domesticated species like livestock, where humans may have interfered in the geographic distribution of the study species, we have taken corrective measures in our study design. Our sampling

strategy was elaborate enough to ensure environments potentially habitable by chickens are included in sufficient sample size, while those uninhabitable are excluded in the sampling frame. We targeted random mating, nondescript indigenous chicken populations from separate livestock market-sheds, clustered along environmental gradients, to maximize ecological and phenotypic variation between sample populations. We followed a novel approach integrating SDMs with PDMs through generalized additive models (GAMs) to identify the most important environmental variables contributing to habitat suitability and evaluate their relationships with phenotypic differentiation among Ethiopian indigenous chicken populations.

2. 2 Materials and Methods
Sampling strategy

A hybrid strategy, maximizing both environmental and geographical representativeness of sampling sites, increases statistical power by reducing false discovery rates caused by demographic processes and confounding effects (De Mita et al., 2013; Lotterhos and Whitlock, 2015; Selmoni et al., 2020).
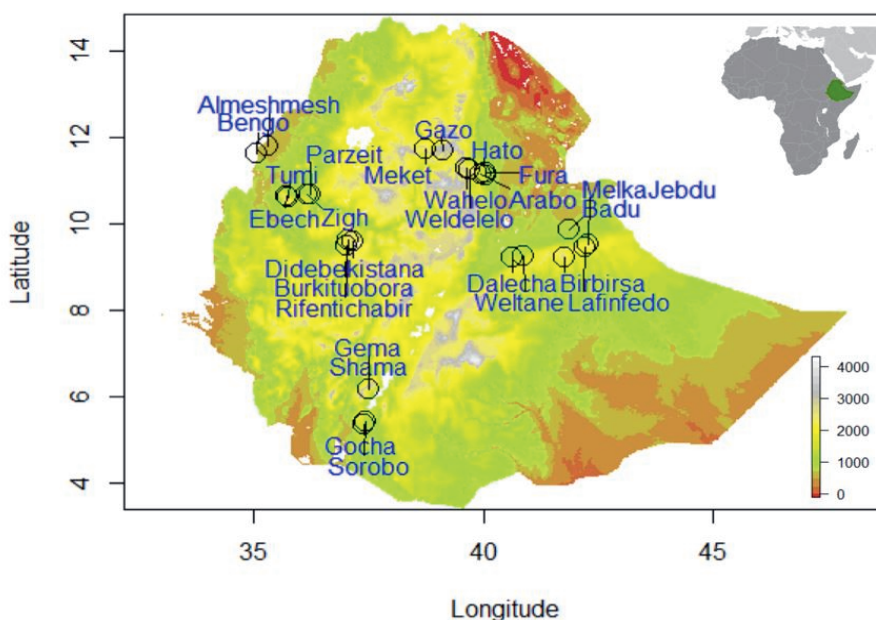

Figure2.1. Topographic map of Ethiopia depicting the 26 indigenous chicken sample populations

We used a hybrid sampling strategy that covered the target area, ensuring high environmental variability, wide geographic distributions, and considering the demographic and biotic processes influencing the Ethiopian indigenous chicken populations (**Figure 2.1**). 513 chickens and geolocations of 26 populations were sampled from four spatial gradients with a minimum distance between gradients of 500km. Each gradient comprised three environmental clusters, primarily delineated based on elevation (400-1800; 1800-2400; 2400-3500 m.a.s.l.).

While we did not consider administrative regions in Ethiopia in our sampling strategy, we would like to describe the four regions covered in the present study (Amhara, Oromia, Benishangul-Gumuz, and Southern regions) to give a brief view of the geographic landscape. Gradient-I stretched from the Rift valley lowlands of northeastern Ethiopia (McConnell, 1967), along the territories of Afar region to the highlands of Wollo province within Amhara region. Gradient-II, starts from the Rift valley lowlands in central Ethiopia, crosses the highlands of Hararghe, including Mount Gara Muleta, and stretches to eastern Ethiopia within Oromia region. Gradient-III stretches from the highlands of northwestern Ethiopia and goes down to the lowlands along the Ethiopian-Sudanese border within Benishangul-Gumuz region. Gradient-IV spans from the highlands of western Ethiopia in Oromia region to the lowlands along the Ethiopian-Kenyan border in Southern region. Areas around the national borders of Ethiopia have low elevation, which gradually culminate to highland plateaus in the center of the country creating a striking contrast in agroecology (Authority, 1988). Geographic coordinates and phenotypic measurements were not taken from areas which are not habitable by chickens because of their unconducive environments (below 400 and above 3500 m.a.s.l.).

We made sure that clusters within a gradient were distant by at least 100km and the target chicken populations were sampled from households which visit isolated i.e., not connected livestock market-sheds. The concept of market-shed refers to a geographic area, where households therein are in sufficient proximity to exchange their animals in various ways (e.g., sale, gift), most commonly travelling on foot. Each cluster along the spatial gradient constituted 2-3 populations. A total of 26 populations were sampled (**Figure 2.1 and Supplementary Table 2.1**). The sampling frame is spatially evenly spread to capture high inter- and intrapopulation environmental and quantitative trait variability. The research design integrating species and phenotypic distribution models is presented in **Figure 2.2**.
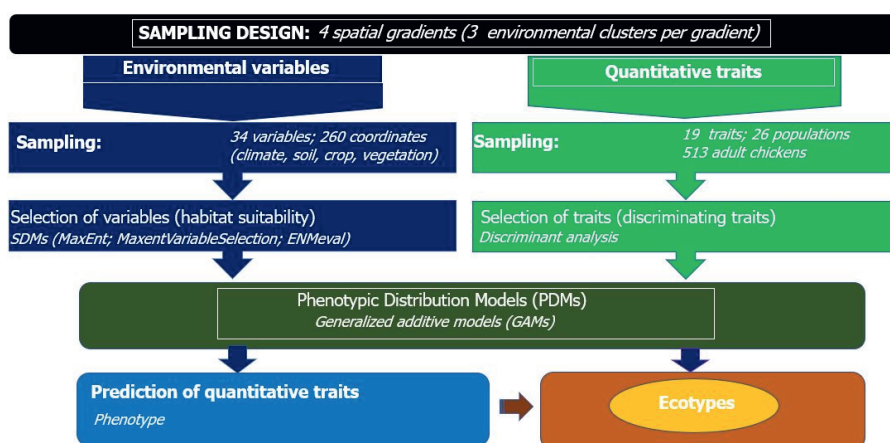
Figure 2.2. Workflow for integrated species and phenotypic distribution modelling to detect population differentiation and define ecotypes of indigenous chickens

The sample locations in our study covered 13 out of the total of 18 agroecological zones (MoA, 1998; Tadesse et al., 2006) in Ethiopia. Agroecological zonation utilizes biophysical attributes of soil, terrain, and climate to organize land-use types or production systems into relatively homogenous units (Hurni, 1998). The five agroecologies that were not covered, are areas where chickens have either not been reared due to extreme climates, cannot be kept at all (e.g., water bodies, undisturbed forests), or have only been recently introduced.

2.2.1 Environmental data

A total of 34 environmental variables were selected for their potential effects on chicken adaptive evolution. Data on these variables was extracted from online databases (**Supplementary Table2.2**). The environmental data included bioclimatic (n=24), vegetation (n=2) and soil (n=8) variables. Values for bioclimatic variables (temperature, precipitation, soil radiation, and water vapour pressure) in different seasons were obtained from WorldClim database (http://www.worldclim.org/; version 2) at a spatial resolution of 30 seconds (~1Km^2) (Fick and Hijmans, 2017) based on mean values of 30 years (1970-2000). Cropland extent at 30-meter (m) resolution was attained from Global Food Security Analysis-Support Data (Xiong et al., 2017). The SoilGrids system at 250m resolution with standard numeric soil properties (organic carbon, bulk density, Cation Exchange Capacity (CEC), pH, soil texture fractions at 15-30 cm depth) was accessed from ISRIC database(Hengl et al., 2017; Hengl et al., 2015). In addition to the 34 environmental variables elevation

data was downloaded from DIVA-GIS (http://www.diva-gis.org/gdata)(Farr et al., 2007; Hijmans et al., 2001) at a spatial resolution of 30 seconds (~1Km^2).

2.2.1.1 Species distribution models (SDMs)

SDMs (also called niche, envelope, or bioclimatic models) associate georeferenced observations of a biotic response variable – typically species occurrence or abundance – with multiple environmental predictors using a broad array of statistical learning methods to describe species' niches (Elith and Franklin, 2013; Elith and Leathwick, 2009; Franklin, 2010; MacKenzie et al., 2017). MaxEnt is a general-purpose machine learning algorithm developed to model species distributions from presence-only records (Phillips et al., 2006).

For every population, a single geographic coordinate was taken at the center of the village during sampling of chickens. Coordinates from nine additional grids (1.44km^2), covering a total of 12.96km^2, were then drawn around a recorded location and extracted using Google Earth Pro v 7.3.2 to ensure high representation of environmental variability affecting the population (Gheyas et al., 2021). This way the total number of 'presence' or 'occurrence' points used in SDMs for the 26 sample populations comprised 260 coordinates. Different R software packages: '*sp*' (Pebesma et al., 2012), *'raster'* (Hijmans et al., 2015)*, 'rgdal'* (Bivand et al., 2021a)*, 'maptools'* (Bivand et al., 2021b)*, 'rgeos'*(Bivand et al., 2017) *,* and *'dismo'* (Hijmans et al., 2017)*,* were used to extract, read, and visualize geospatial data. Dimension and extent of the grids were corrected and homogenized for 1 km^2 based on the WGS84 geodetic reference system (Decker, 1986).

Selection of environmental variables

To constrain model complexity and increase the performance of species distribution models, the highest contributing set of uncorrelated environmental variables were identified and Maxent's regularization multiplier was fine-tuned using the R package '*MaxentVariableSelection*' (Jueterbock et al., 2016). The predictive performance of the most important environmental variables was measured using test gain in MaxEnt v.3.4.1(Phillips et al., 2006) (Phillips and Dudík, 2008).

2.2.1.2 Configuration of model parameters

Species-specific tuning of model parameters can improve the performance of MaxEnt model compared to the default settings (Elith et al., 2011) (Radosavljevic and Anderson, 2014). The large set of feature types was subsequently reduced to the

optimal subset to improve model fit and the optimum regularization multiplier for model training was identified by the R package '*ENMeval*' (Muscarella et al., 2014) by using spatial blocks method (Radosavljevic and Anderson, 2014). Regularization refers to smoothing the model, making it more regular, to avoid fitting too complex a model. It is a common approach in model selection and penalizes coefficients (the betas) to values that allow both accurate prediction and generality (Elith et al., 2011; Tibshirani, 1996).

Species' responses to environmental covariates or independent variables (e.g., temperature, elevation) tends to be complex and usually requires fitting of nonlinear functions (Austin, 2002). In machine learning algorithms this is achieved by applying transformations of the original covariates into feature. MaxEnt currently has six feature classes: linear, product, quadratic, hinge, threshold and categorical (Elith et al., 2011). We built models with regularization multiplier values ranging from 0.5 to 4.0 (increments of 0.5) and with six different feature combinations (H, LQH, HQP, HQC, LQHP, LQHPT; where L = linear, Q = quadratic, H = hinge, P = product and T = threshold); this resulted in 48 individual model runs. The parameter configuration with the lowest delta AICc value was chosen to run the model (**Supplementary Table 2.3**). To reduce the influence of sampling bias, we included a bias file (Philips et al 2009) and preferentially sampled pseudo background points from areas near our presence points based on kernel density function (Venables and Ripley, 2002).

2.2.1.3 Tests of niche similarity

A niche is a description of the conditions in which a species maintains a viable population. Populations in a species that are adapted to a specific local habitat or niche show genetically induced phenotypic differences in response to environmental selective pressures and are regarded as 'ecotypes' (Knüpffer et al., 2003; Müntzing, 1971). Niche similarity between one or more pairs of populations was measured according to Warren et al. (2008). Raster files (.ASCII) of predicted habitat suitability produced by MaxEnt in logistic output (no probability and complete probability of presence designated by 0 and 1, respectively) were used as inputs to perform correlation test by ENMTools (Warren et al., 2011; Warren et al., 2010). Correlation tests were used to cluster sampling sites on the selected environmental variables and build dendrogram through hierarchical clustering with R package *cluster* (Maechler et al., 2013)**.** The grouping of sampling locations into environmental niches was based on "Euclidean" distance. Different clustering methods (Ward's minimum variance method, complete linkage, average linkage, and single linkage) were compared. Visualization of the cluster memberships of locations of populations

based on niche similarity, measured by correlations tests on the most important environmental variables, was accomplished using the R package *factoextra* (Kassambara and Mundt, 2017).

2.2.2 Quantitative trait data

A total of 19 phenotypic traits **(Supplementary Table 2.4),** selected for their potential role in adaptation in chicken based on available literature, were measured on 513 adult chickens (380 hens and 133 cocks) from the 26 nondescript indigenous chicken sample populations. We had three environmental clusters (lowland: 400-1800; midaltitude: 1800-2400; and highland: 2400-3500 m.a.s.l.) stretching across each of the four elevational gradients in this study. A total of 12 environmental clusters from the four elevational gradients were included. Each environmental cluster is represented by two randomly selected chicken populations, except in two instances where we took three populations. A population refers to the total number of nondescript indigenous chickens available in an administrative village. Adult chickens were selected randomly for phenotyping through transect walk across villages. This method entailed walking along a defined path (transect) across a village and sampling one chicken from each farming household until a total of 15 hens and 5 cocks were measured.

The age of the chickens was estimated by interviewing owners to confirm that females were in their second clutch (7 to 8 months-of-age) and males were above 12 months-of-age. The researchers also visually appraised cocks (roosters) for presence of well-developed spurs. One chicken was sampled per household. Under rare circumstances (n = 9 households), two chickens were sampled per household when farmers proved their animals have no family relationship.

Live bodyweight of individuals was taken in the morning on fasting chickens. Accurate morphological measurements were made by digitally analyzing the pictures of individual chickens photographed in a sheltered environment using a software (Rasband, 1997). To reduce systematic error, the same operator measured all chickens, which were held in the same position by a technician. A steel ruler was placed in the background of every picture as a distance reference.

2.2.2.1 Selection of quantitative traits

A multivariate test of differences between populations with stepwise selection (Klecka et al., 1980) was performed through linear discriminant function analysis

(SAS, 2002) to identify the traits which were most useful in classifying populations. Principal Component Analysis (PCA) was run with R '*stats*' package on quantitative trait data to see how much percent of the variation is explained by the first nine principal components (PCs).

2.2.2.2 Clustering of nondescript chicken populations into ecotypes

The 26 nondescript Ethiopian chicken populations sampled in this study are heterogenous in terms of qualitative traits (e.g., coat colour, comb shape, feather pattern) and quantitative traits. We used the most discriminant quantitative traits, which are most useful because of their variability, to group populations into ecotypes. We expect that populations of chickens within the same niche are affected by similar environmental variables and cluster into the same ecotype. The phenotypic values of these traits were analysed by the average silhouette method to decide on the optimal number of clusters. The average silhouette method measures how well each experimental unit lies within its cluster and is less ambiguous than the elbow method to decide on the number of clusters (Rousseeuw, 1987) (Kaufman and Rousseeuw, 2009).

Different hierarchical clustering methods (Ward's minimum variance method, complete linkage, average linkage, and single linkage) were compared via R packages '*cluster*' (Maechler et al., 2013) and '*factoextra*' (Kassambara and Mundt, 2017) to make a valid comparison of population memberships between dendrograms produced on similarity of phenotypes. We used the same approach for clustering of environmental and phenotypic data to avoid any possible bias associated with the use of different tools.

2.2.3 Phenotypic distribution models (PDMs)

While species can vary genetically and phenotypically across their range and populations can be locally adapted, SDMs assume that all populations respond homogenously to the range of environmental conditions experienced by the whole species (Atkins and Travis, 2010; Bolnick et al., 2003; Fitzpatrick and Keller, 2015; Hällfors et al., 2016). Phenotypic distribution models (PDMs) on the other hand, do capture the response of quantitative traits as a function of environmental conditions (Lozano-Jaramillo et al., 2019a; Michel et al., 2017; Smith et al., 2017). We used PDMs to study variation within quantitative traits in response to the most important set of environmental variables identified by SDMs. The association of these environmental variables with habitat suitability were evaluated for their individual

effects on each of the discriminating the trait. The relationship between quantitative traits and environmental variables was expected to be non-linear (Oddi et al., 2019; Zuur et al., 2007). The assumptions of classical statistical approaches such as Generalized Linear Models (GLM) are violated when responses are non-linear, variances change with predictors, or ecological processes operate at spatio-temporal scales (Bolker et al., 2013; Zuur et al., 2009).

Exploration of phenotypic and environmental data was initially carried out to understand their distribution, variance structure, and linearity or non-linearity of trend and to choose appropriate analytical methods. Generalized additive models (GAMs) were selected because they are particularly useful for analyzing relationships explained by complicated shapes, such as hump-shaped curves (Crawley, 2012).The R package '*mgcv*' (Wood and Augustin, 2002) was used to fit GAMs (Hastie and Tibshirani, 1990). Model validation was made based on Akaike Information Criterion (AIC) values.

The response of each quantitative trait was predicted as a function of ecotype, niche, and the six SDM-selected environmental variables. The GAM included ecotypes and their respective niches as linear terms and the environmental covariates as smoothing parameters. The notation for the GAM smoothing in a Gaussian model is as follows (Hastie and Tibshirani, 1990; Wood and Augustin, 2002).

$$g(E(y_i)) = \alpha + \beta_j + \gamma_m + f_k(X_{ki}) \ldots,$$

Where $(E(y_i))$ is one of n observations of the response trait, g is the Gaussian distributed exponential family with identity link function, α is the intercept, β_j is a linear parameter for ecotype (1,2,3), γ_m is a linear parameter for environmental niches (1,2,3), f_k are the smoothing terms based on non-parametric predictor covariates X_{ki} (the shape of the predictor funtions which will be fully determined by the data structure).

Estimation of smoothing parameters effects (environmental variables) was done by restricted maximum likelihood (REML) as random effects (Wiley and Wiley, 2019) with Gaussian process smooth (bs='gp') in the GAMs model (Wood, 2012).

Partial dependence plots (PDPs) (Friedman, 2001) are the most popular approach for visualizing the effects of the predictor variables on the predicted outcome during supervised machine learning applications (Apley and Zhu, 2020). A partial dependence plot can show whether the relationship between the target and a

feature is linear, monotonic, or more complex. PDPs exhibiting the effects of environmental factors with estimated p-value on a phenotype were produced by using the R package '*mgcViz*'(Fasiolo et al., 2020) at 95% confidence interval.

2.3 Results

2. 3.1 Environmental variables contribute to habitat suitability
2.3.1.1 Optimum model parameters

ENMeval identified HQP (Hinge, Quadratic, and Product) features with regularization-multiplier=3.0 as the best parameter combination. This had the lowest deltaAICc value and was chosen to produce suitability maps by MaxEnt (**Figure 2.3A**). Compared to the default (**Figure 2.3B),** the model fit with the optimum parameters predicted larger areas as most suitable for poultry production (**Figure 2.3C**). The areas least populated by chickens include the extreme lowlands (below 400m.a.s.l.), with prohibitively high temperature, high solar radiation, low precipitation, and high relative humidity; and the extreme highlands (above 3400m.a.s.l), with prohibitively low temperatures. The extreme highlands are frosty and hence not habitable both by livestock and humans. Ethiopian lowland pastoral areas are affected by recurrent drought and have generally sparse livestock population (Tilahun and Schmidt, 2012). Agreement between the results of the present study and the census report (CSA, 2017) and other literature indicating the distribution of livestock (Tilahun and Schmidt, 2012) confirm that those areas in the country which are shown as least suitable in the habitat suitability maps produced by SDMs are indeed unsuitable for the study species.
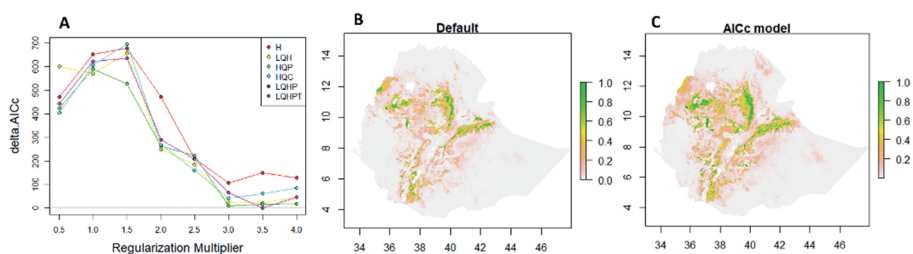
Figure 2.3. Model configuration and habitat suitability maps for Ethiopian indigenous chicken populations. **(A**) AICc values for analysed feature combinations using different regularization-multipliers ranging from 0.5 to 4.0. Feature combinations include one or more of the following types: L = linear, Q = quadratic, H = hinge, P = product, and T = threshold. **(B**). Map produced using default settings of MaxEnt. **(C)**. Map produced using optimum parameters (HQP features with regularization-multiplier=3.0) identified by ENMeval

Sedentary systems in central Ethiopia have conducive environmental conditions for chickens while pastoral systems (hot, dry areas, with strong solar radiation) along the borders of the country do not (Bayou and Assefa, 1989; CSA, 2017; Gebrechorkos et al., 2019; Getahun, 1978; Mirkena et al., 2018). The choice of livestock species to rear is also culturally embedded over generations.

2.3.1.2 Most contributing environmental variables

SDMs identified the most important environmental variables associated with distribution of chickens (**Figure 2.4**). Correlated variables (|r| > 0.6) and those with a relative contribution score below 4% were removed to restrict multicollinearity driven effects in projecting species ranges (Brun et al., 2019; Dormann et al., 2013). Out of 34 environmental variables, 9 were retained as most important in determining habitat suitability and can be regarded as potential drivers of local adaptation in Ethiopian indigenous chickens. The first five variables with the highest contribution included soil clay content, precipitation of the warmest quarter, precipitation of the coldest quarter and temperature seasonality.
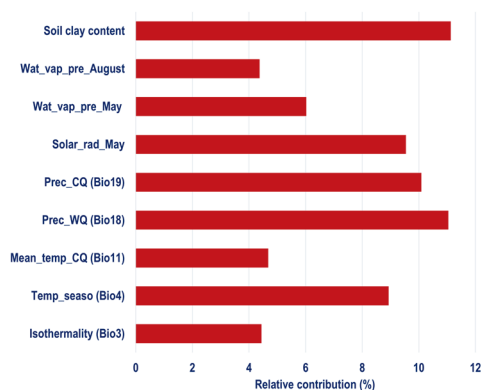
Figure 2.4. Environmental variables of importance and their percent contribution predicted by MaxentVariableSelection

Jackknife test was run to compare the relative importance of the 9 selected environmental variables (**Figure 2.5**). The test showed that precipitation of the coldest quarter and water vapour pressure in May have the highest gain when used in isolation, and therefore are the most useful variables for predicting the distribution of the species on occurrence data. On the other hand, the environmental variable that decreases gain the most when omitted is solar radiation in May, meaning it has the most important information that is not present in other variables.

Figure 2.5. Gains of the variables in the Maxent model (Jackknife test) for Ethiopian indigenous chickens. Torques bars: model gain without corresponding variables; blue bars: model gain with only the corresponding variables; red bars: total gain using all the variables.

2.3.2 Distinct niches are associated with distinct ecotypes

Populations of animals adapted to a specific environment or niche are regarded as ecotypes. Clustering of sample chicken populations into phenotypically homogenous groups and an overlap of the clustered populations with niche classificati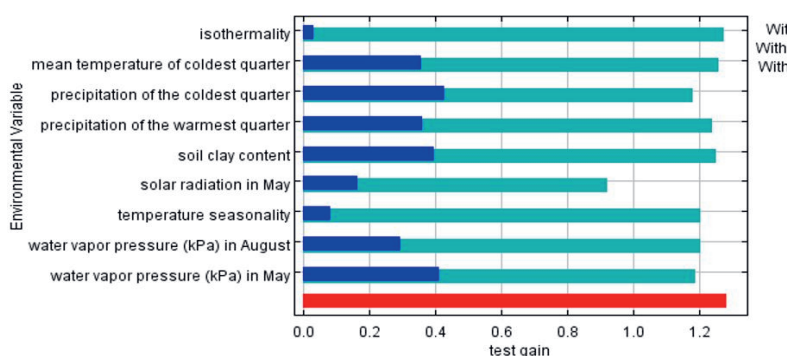on based on their respective environments was used as a basis to define ecotypes. The number of chicken ecotypes was determined through Silhouette method using phenotypic data (**Figure 2.6A**). The optimal cluster in the present study, the one that maximized the average silhouette from a range of possible **k** values, was **k**=3. The same clustering method (Ward Jr, 1963) was used to make a valid comparison of population memberships between dendrograms produced on similarity of niches (**Figure 2.6B**) and on similarity of phenotypes (**Figure 2.6C**).
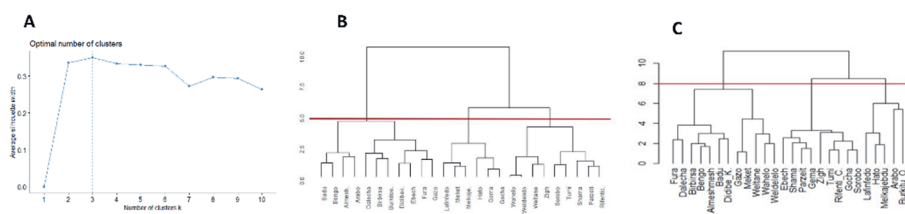
Figure 2.6. Dendrogram of clusters to group 26 Ethiopian indigenous chicken populations. **(A)** Based on niche overlap statistic (I) between suitability maps. **(B)** Based on the most discriminating phenotypes (hierarchical agglomerative clustering, Ward's minimum variance method). The red line at a cutoff value of 5.0 produces three distinct niches. **(C)** Plot of statistics for deciding appropriate number of clusters based on phenotype. The red line at a cutoff value of 8.0 produces three distinct ecotypes.

Populations were clustered into three environmental niches based on correlation test (**Figure 2.6B**). Ward's method had the strongest clustering structure for clustering on niche overlap (Ward = 0.89). The agglomerative coefficients for the other approaches (complete linkage = 0.78; average linkage (UPGMA) = 0.68; and single linkage = 0.36) was lower. At a cutoff value of 5.0, reading the plot from left to right, niche-I comprised eleven sampling locations, while niche-II and niche-III comprised 6 and 9 locations, respectively.

2.3.4 Variation in quantitative traits

Before classifying the 26 sample chicken populations into ecotypes through hierarchical clustering based on similarity for quantitative traits, we reduced the number of traits through discriminant analysis (**Table 2.1**). Out of 19 quantitative traits (**Supplementary Table 2.4**), 8 (BL, WS, CL, CW, BW, EW, WW, KL) had the highest discriminant function because of their high variation between populations. Except wattle width ($p<0.05$), the remaining 7 of these 8 discriminant traits showed highly significant phenotypic variation ($p<0.0001$ to $p<0.01$) between female sample chicken populations. The GLM analysis combining data from both sexes revealed all the discriminating quantitative traits varied significantly between sexes ($p<0.0001$) except for beak length ($p=0.1738$). The partial r-square indicates body length had the highest discriminatory effect out of all traits retained in the models in both sexes. Only two quantitative traits (BL and BW) were found useful for classifying male sample ($p<0.0001$). This might be related with their lower sample size or a different structure of morphological variation among male sample populations compared to females.

A subset of quantitative traits that best revealed the differences among chicken populations (**Table 2.1**) were then used for clustering. Ward's hierarchical clustering rendered the highest agglomerative coefficient (Ward = 0.81) for clustering of populations on phenotypic similarity compared with the other approaches (complete linkage = 0.71; average linkage (UPGMA) = 0.58; and single linkage = 0.49) (**Figure 2.6C**). The cutoff value at 8, indicated by horizontal line, resulted in three clusters. A Principal Component Analysis (PCA) on quantitative trait data showed that the first three PCs explain 75.7% of the phenotypic variation among populations (PC1=43.1%, PC2=19.5%, and PC3=13.2%) supporting our grouping of chicken populations into three ecotypes (**Supplementary Table 2.5**).

Table 2.1. Stepwise selection summary indicating most discriminating traits for adult male and female Ethiopian indigenous chicken sample populations

Sex	Quantitative trait	Partial R-sq.	F value	Pr > F
Hens	BL	0.4761	13.51	<.0001
	WS	0.2934	6.15	<.0001
	CL	0.2274	4.34	<.0001
	CW	0.1766	3.15	<.0001
	BW	0.1741	3.08	<.0002
	EW	0.1677	2.93	<.0003
	WW	0.1184	1.63	<.0214
	KL	0.1534	1.93	<.0014
Cocks	BL	0.7756	14.52	<.0001
	BW	0.4856	3.9	<.0001

BL=body length(cm); WS=wingspan(cm); CL=comb length(mm); BW=body weight(kg); EW=earlobe width(mm); WW=wattle width(mm); KL=beak length(mm)

A summary of cluster analyses (**Table 2.2**) shows that most of the populations of a specific ecotype are distributed within the same niche while only a few of them distributed elsewhere. Eight out of twelve populations from ecotype-I, three out of five populations from ecotype-II, and six out of nine populations from ecotype-III were correctly classified into their respective niches.

Table 2.2. Ecotype of Ethiopian indigenous chicken populations defined on phenotype and their respective niches as identified by species distribution models

Ecotype	Populations	Distributed within the same niche	Distributed outside the niche
I	Fura, Dalecha, Birbirsa, Bengo, Almeshmesh, Badu, Didibe Kistana, Gazo, Meket, Weltane, Wahelo, Weledelelo	Fura, Dalecha, Birbirsa, Bengo, Almeshmesh, Badu, Didibe Kistana, Gazo	Meket, Weltane, Wahelo, Weldelelo
II	Lafinfedo, Hato, Melkajebdu, Arabo, Burkitu Obora	Lafinfedo, Hato, Melkajebdu,	Arabo, Burkitu Obora
III	Ebech, Shama, Parzeit, Gema, Zigh, Tumi, Rifenti Chabir, Sorobo, Gocha	Shama, Parzeit, Zigh, Tumi, Rifenti Chabir, Sorobo	Ebech, Gema, Gocha

Matching between chicken ecotypes and different environmental classification methods was performed to establish a logical association between phenotypic

distinctiveness and environmental selective pressures (**Table 2.3**). The environmental classification methods included species distribution models, conventional (Dove, 1890), Official (MoA, 1998), and gradient-based agroecological classifications. The highest level of correct classification was performed by SDMs (64.5%), followed by environmental gradient (elevational cline) classification (57.3%). The higher correct classification level obtained by the SDM approach, suggests the potential influence of the selected environmental variables (n=9) on shaping adaptive variation among Ethiopian indigenous chicken ecotypes.

Table 2.3. Comparison of methods to classify environments of Ethiopian indigenous chicken ecotypes (n=3)

Classification method	Criteria for classification		No. of population correctly classified (%)			Total correctly classified (%)
		Classes	Ecotype-I	Ecotype-II	Ecotype-III	
SDM	niche similarity	3	8 (66.7)	3 (60.0)	6 (66.7)	17 (64.5)
*Conventional AEs	climatic classes (altitude, temperature, precipitation)	3	6 (66.7)	3 (33.3)	4 (50.0)	13 (50.0)
§Official AEs	temperature, soil type, plant growing period/moisture condition, land use	13	3 (33.3)	3 (33.3)	2 (25.0)	8 (30.7)
Gradient	elevational clines in distinct geographies	4	5 (55.5)	6 (66.6)	4 (50.0)	15 (57.3)

*Conventional agroecological classes (AEs) comprise three groups measured in m.a.s.l.: I=lowlands (400-1800); II=1800-2400; III=2400-3500 (Dove, 1890). § Official AEs represent standard agroecologies of Ethiopia (MoA, 1998).

2.3.5 Environmental variables contribute to phenotypic differentiation

Having noticed that populations have differentiated distinctly in specific environments, we focused on predicting phenotypic values of ecotypes for the most discriminant quantitative traits within their respective niches under the influence of the selected environmental variables. Prediction of quantitative traits with Generalized Additive Models (GAMs) in each of the three Ethiopian indigenous chicken ecotypes is presented in **Table 2.4**. Significant *p*-values were obtained for all the nine SDM identified environmental variables except for soil clay content. Five environmental variables (Bio18, Bio19, WVPM, and WVPA) had significant effect on differentiation of multiple traits. The traits selected by discriminant function for their usefulness in classification of populations into ecotypes had also the highest model fit (R-square adjusted values) explaining their importance in studying the influence of environmental variables on adaptive phenotypic variation.

Table 2.4. Prediction of quantitative traits with Generalized Additive Models (GAMs) in Ethiopian indigenous chicken ecotypes (n=3) fitted with ecotype and niche as fixed effects.

Trait	Random effects/ Smoothing term									Model fit			
	Bio3	Bio4	Bio11	Bio18	Bio19	SRM	WVPM	WVPA	SCC	df	[1]AIC	R-sq. (adj)	Deviance explained (%)
BL	*			*	**	*	**	**		14.9	1539	0.65	66.7
					*		*	*					
WS			**	**	*	*	**			13.6	1644	0.55	56.5
			*	*			*						
CL				*						8.0	2583	0.21	22.3
CW							**	**		9.3	2140	0.10	11.9
							*	*					
BW	*			**	*	**	**			12.5	-121	0.45	46.5
						*	*						
EW		***	.				**	**		16.3	1657	0.25	28.2
							*	*					
WW	*			**				**		10.0	2035	0.12	14.3
KL				*	**	.		**		11.9	1724	0.05	7.6

[1]Akaike information criterion (AIC) is a goodness of fit measure (likelihood or log-likelihood) that penalizes for complexity number of parameters or degree of freedoms). BL=body length; WS=wingspan; CL=comb length CW=comb width; BW=body weight; EW=earlobe width; WW=wattle width; KL=beak length; Bio3=Isothermality; Bio4=Temperature seasonality; Bio11=Mean temperature of coldest quarter; Bio18=Precipitation of warmest quarter; Bio19=Precipitation of coldest quarter; SRM=solar radiation of May; WVPM=water vapour pressure of May; WVPA= water vapour pressure of August; SCC=soil clay content. Linear effect of ecotype is significant for all discriminating phenotypes. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Ethiopian indigenous chicken ecotypes identified by SDMs showed significant quantitative trait variation (**Table 2.5**). Populations in ecotype-I had the smallest measurement for all traits while ecotype-II had the largest measurements for most traits. It is not possible to tell from the present results alone whether the performance exhibited by ecotypes is primarily attributable to their niche or their genetic background.

Table 2.5. Quantitative trait variation in Least Square Mean (Standard Error) among adult female Ethiopian indigenous chickens of different ecotypes defined by integrating SDMs with PDMs

Ecotype*	LSMean (S.E.)									
	Hens (*n=380*)								Cocks (*n=133*)	
	BW	BL	WS	CL	CW	EW	WW	KL	BW	BL
I	1.01	35.46	38.78	21.3	7.95	8.96	16.89	16.32	1.31	38.89
	(0.01)[b]	(0.22)[c]	(0.22)[b]	(0.62)[c]	(0.36)[b]	(0.2)[b]	(0.31)[c]	(0.20)	(0.05)[b]	(0.52)[b]
II	1.31	39.13	41.88	30.21	10.18	10.9	19.4	16.46	1.78	44.34
	(0.02)[a]	(0.23)[a]	(0.24)[a]	(0.65)[a]	(0.37)[a]	(0.2)[a]	(0.32)[a]	(0.21)	(0.05)[a]	(0.54)[a]
III	1.28	38.48	42.03	25.22	8.82	10.4	18.48	16.65	1.82	44.49
	(0.02)[a]	(0.24)[b]	(0.24)[a]	(0.66)[b]	(0.38)[b]	(0.20)[a]	(0.33)[b]	(0.22)	(0.05)[a]	(0.57)[a]

BW=body weight(kg); BL=body length(cm); WS=wingspan(cm); CL=comb length(mm); CW=comb width(mm); EW=earlobe width(mm); WW=wattle width(mm); KL=beak length(mm). [a,b,c]Means with different superscripts within the same column and model are significantly (*P* <0.05) different. * Ecotypes were highly significant from each other (p<0.0001) for all phenotypic measurements except for KL in hens (*P*=0.5393).

Habitat suitability maps for Ethiopian indigenous chicken ecotypes (**Figure 2.7**) illustrate ideal environmental conditions that vary spatially between ecotypes. Chickens of ecotype-I (**Figure 2.7A**) are mainly distributed in central and northwest Ethiopia, ecotype-II (**Figure 2.7B**) are distributed in the west and southwest, while ecotype-III (**Figure 2.7C**) are distributed in eastern and northeastern Ethiopia. Areas of the country characterized by adverse environmental conditions due to their extreme temperature, high solar radiation, and low precipitation are shown as least suitable. This result conforms to the available census data which shows regions in the country with more friendly climate to chickens are more populated by the species (CSA, 2017).
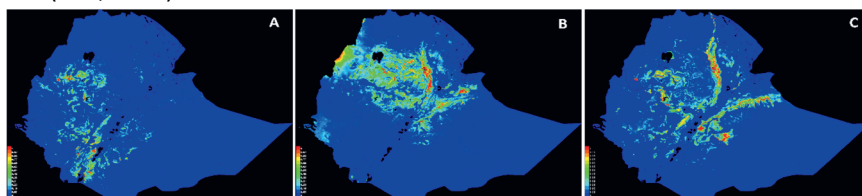

Figure 2.7. Suitability maps of three Ethiopian chicken ecotypes. Colours towards red spectrum indicate more suitable conditions. **(A)**. Ecotype-I; **(B)**. Ecotype-II; **(C)**. Ecotype-III.

The response of adult live body weight (BW) and body length (BL) in female indigenous chickens to some of the significant environmental variables ($p<0.001$) are presented in **Figures 2.8 and 2.9.** The relationship between body weight and solar radiation, and body weight and water vapour pressure in May (kPa) is linear while its relationship with isothermality is non-linear (**Figure 2.8**). Isothermality quantifies how large the day-to-night temperatures oscillate relative to the annual oscillations. An isothermal value of 100 indicates the diurnal temperature range is equivalent to the annual temperature range, while anything less than 100 indicates a smaller level of temperature variability within an average month relative to the year (O'Donnell and Ignizio, 2012). Our results suggest that body weight is less influenced by smaller temperature fluctuations within a month relative to the year. On the other hand, solar radiation above 18000Kj/m^2/day is stressful and has negative and linear effect on female body weight. The relationship between bodyweight and mean temperature of the coldest quarter is more complex, showing that the mean temperatures during the coldest three months of the year is less useful to examine how this variable affects adult live body weight.

Figure 2.8. Generalized additive model partial dependence plots for live body weight (kg) in female indigenous chickens. Each plot shows a covariate and their partial dependence on adult live body weight in the context of the model. The y axis shows the mean of observed change in live body weight and the x axis the covariate interval. The *blue line* represents the 95% confidence interval; *Red line* = mean of observed change in live body weight; *s* = smoothed variable; and *( )* =effective degrees of freedom

A non-linear relationship is noted between body length (BL) and water vapour pressure in August (kPa), and between body length and precipitation of the coldest quarter (mm/m²). Temperature seasonality had a negative and linear relationship with this trait. Temperature seasonality is a measure of temperature change over the course of the year. Our result indicates that higher standard deviation in the mean monthly temperature is associated with smaller body length, a trait which is strongly correlated with live body weight. Precipitation of the coldest quarter is a quarterly index which approximates the total precipitation that prevails during the three months of the year. Accelerated mean change in body length, in the context of the model was seen up to 700mm/m² of precipitation in the coldest quarter. Precipitation above this threshold might be related with less availability of scavenging feed resources and more prevalence of diseases and parasites, having adverse effects on this trait. Biologically speaking, water vapour pressure is a function of temperature and pressure. Negative relation is noted between this environmental variable and body length, probably because of the stressful situation (e.g., lower feed intake) it creates on the animals. A non-linear reduction was observed in body length for higher soil content above 20 percent which may have a relationship with the type of vegetation and land use pattern in those areas (**Figure 2.9**).
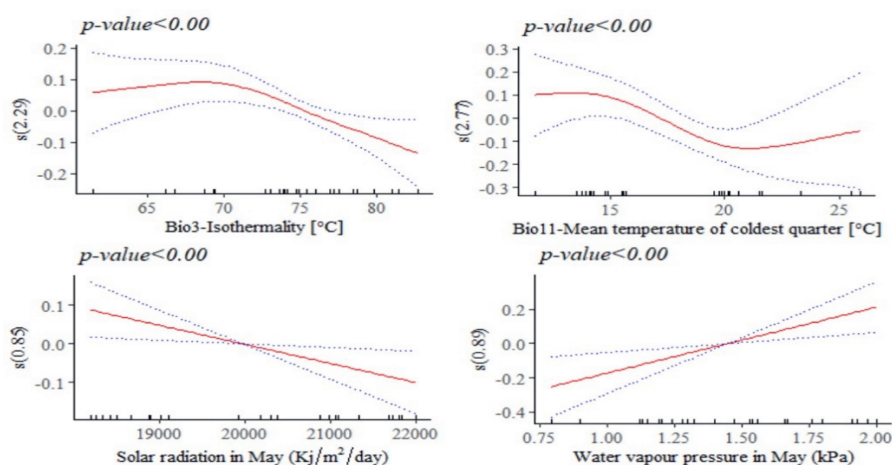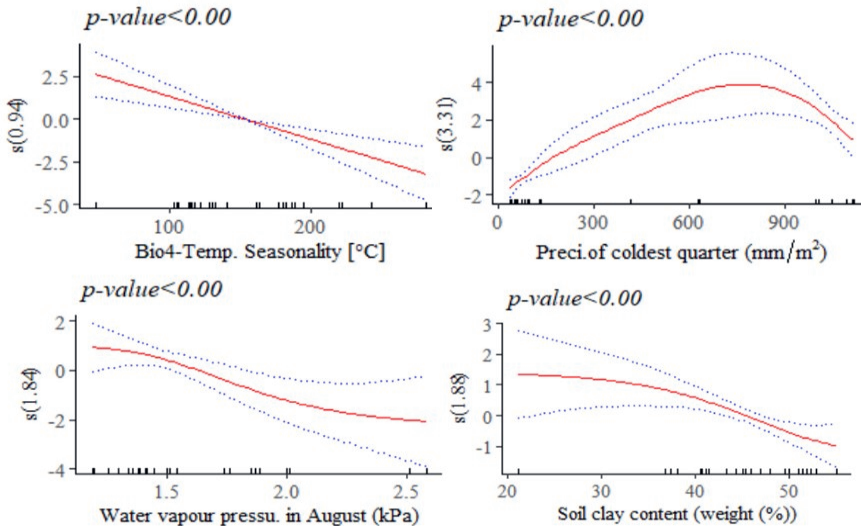

Figure 2.9. Generalized additive model partial dependence plots for body length (mm) in female indigenous chickens. Each plot shows a covariate and their partial dependence on adult live body weight in the context of the model. The y axis shows the mean of observed change in live body length and the x axis the covariate interval. The *blue line* represents the 95% confidence interval; *Red line* = mean of observed change in live body weight; *s* = smoothed variable; and *( )* =effective degrees of freedom.

2.4 Discussion

Sustainable livestock production particularly in the tropics requires adaptive genotypes which can withstand the undesirable effects of climate change and produce optimally (Fleming et al., 2017) (Bettridge et al., 2018). Ecological variables vary in terms of their influences on organisms as inducers of local adaptation. Knowledge of ecological factors responsible for adaptive variation should be the first step to design selective breeding programmes on indigenous livestock, plan crossbreeding with improved genotypes, or introduce new genotypes from a different environment (Bettridge et al., 2018; Birhanu et al., 2021; Fleming et al., 2017).

We have applied distribution models to identify the most important environmental factors associated with habitat suitability and phenotypic differentiation in indigenous populations of chickens. Previous studies indicated that populations differentiate phenotypically and genetically in response to the environment (Schmid and Guillaume, 2017; Smith et al., 2017). A tight relation is expected between environmental elements (e.g., precipitation, temperature, radiation, elevation) and livestock population dynamics (Alemayehu and Getu, 2016; Getachew et al., 2016) in Ethiopia.

Precipitation of the warmest and the coldest quarters, soil clay content, temperature seasonality, solar radiation, water vapour pressure, and mean temperature of the coldest quarter, were identified by SDMs as the most important variables associated with habitat suitability in Ethiopian indigenous chickens. Precipitation is associated with types and amounts of crops cultivated; availability of scavenging feed resources and edible soil fauna; disease prevalence, and predation. Precipitation and temperature were also identified as most important contributors to local adaptation in African chickens (Bettridge et al., 2018; Fleming et al., 2017; Gheyas et al., 2021). The body weight of Horro, Koekoek, Sasso, and SRIR chickens distributed to different regions of Ethiopia was best predicted by variables associated with temperature and precipitation (Lozano-Jaramillo et al., 2019a). Clay content is a proxy for soil fertility and has impacts on feed availability for scavenging chickens. Through their physical and chemical properties, clay minerals can be expected to have more nutrient reserves in the tropics (Kome et al., 2019; Landon, 2014).

All the nine environmental variables for their association with habitat suitability by SDMs had significant effects on differentiation of quantitative traits. The influence of isothermality (Bio3), temperature (Bio4 & Bio11), precipitation (Bio19), solar

radiation, and water vapour pressure on trait differentiation may be related with adaptive physiology of chickens, in terms of their biological response to extremes in relative humidity and heat stress. (Alemu et al., 2021); Lozano-Jaramillo et al. (2019a) have also observed effects of precipitation and temperature on improved chicken breeds introduced to smallholder farmers in Ethiopia.

We classified the Ethiopian indigenous chicken sample populations into three ecotypes and compared their respective performances. Homogenous clusters for measured quantitative traits and their overlaps with distinct niches were used to define ecotypes. Unlike previous efforts made to group Ethiopian indigenous chicken populations on qualitative phenotypes such as comb shape, and feather colour (Deneke Negassa, 2014; FAO, 2012; Getachew et al., 2016; Melesse and Negesse, 2011; Overdijk, 2019) , the definition of ecotypes in the present study integrated phenotypic and environmental information. This process included identification of the most contributing environmental variables for habitat suitability, grouping of sample locations into specific niches based on their environmental similarity, and selection of the most useful quantitative traits for population classification purposes.

PDMs, in a form of non-linear generalized additive models (GAMs) were demonstrated as an innovative approach to integrate environmental and phenotypic information and study their relationships. GAMs relax the assumptions of linear models such as GLMs and achieve acceptable goodness of fit. Such a non-linear data structure would have been missed otherwise (Wiley and Wiley, 2019). Phenotypic distribution models were used to complement predictions of species distribution models in studying responses of prairie grass to climate change (Smith et al., 2017).

The use of SDMs is unchartered territory for livestock scientists. Limitations are expected in their use on domesticated species because of human interference influencing the natural distribution of the study populations. While existing SDMs alone do not seem appropriate to study breeds recently introduced into a new environment artificially, the models are applicable for those studying local adaptation among indigenous populations of livestock which have lived in their environment for hundreds of generations or more and have experienced significant selective pressures. Predictive ability of machine learning algorithms on domesticated species can be improved if they incorporate more data in addition to presence-absence information and harness sophisticated algorithms. Boosted regression trees and random forests as well as generalized additive and linear mixed models have improved prediction of SDMs in other species (Shirk et al. 2018).

Several evolutionary processes shape genetic and phenotypic differentiation, including the joint effects of environment (phenotypic plasticity), gene flow, and natural selection (Schmid and Guillaume, 2017). It is not clear from the present study whether the phenotypic differentiation that ensued between indigenous chicken ecotypes is the result of differentiation in allele frequencies. An integrated framework including environmental, phenotypic, and genomic analysis is needed to unravel the genetic basis of phenotypic differentiation among populations and ecotypes of these chickens. If the phenotype is directly influenced by the environment, genetic and phenotypic differentiations can be decoupled (Crispo, 2008; Schmid and Guillaume, 2017). Improvements in predictive ability of models is also achieved when SDMs are used along with phenotypic and genomic information in landscape genetics and genomics studies (Gotelli and Stanton-Geddes, 2015; Joost et al., 2007; Razgour, 2015).

The present study demonstrated how SDM-identified environmental information can be integrated with PDMs to define ecotypes, predict quantitative traits, and understand the ecological roots of phenotypic differentiation. Considering the environmental influences of economically important quantitative traits, such as live body weight, improves the estimation of breeding values and assists in the development of improved breeds suited to smallholder farmers. Differences in performance among ecotypes in the different niches will also mean evaluations of performance and yield stability across environments are pertinent in breeding and development programmes designed for low- and medium-input poultry production systems of the tropics. Prospects of further use for SDMs and PDMs in livestock include definition of agroecologies, estimation of genotype by agroecology interactions (GxE), multi-environment performance evaluations, and prediction of performance under present and future production scenarios (e.g., climate change).

2.5 Acknowledgements

2.6 Data availability statement

The online version of this article (doi: 10.3389/fgene.2021.723360) contains supplementary materials, which are available to authorized users. Datasets are available on request. The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

2.7 Ethics statement

The animal study was reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) of the International Livestock Research Institute, with compliance number IACUC2019-12.

2.8 Author contributions

FGK, JB, HK, OH and TD: conceived the ideas and designed the study. FGK selected sample populations, collected metadata, and performed the phenotyping of chickens. FGK and JB: performed data analysis. FGK, JB and SWA did the Generalized Additive Models (GAMS). FGK wrote the manuscript. HK and HB provided useful comments and suggestions and helped draft the manuscript. TD, HK, OH and JB secured funding. All co-authors critically revised the manuscript and gave final approval for publication.

2.9 Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

3

Landscape genomics reveals regions associated with morphological and environmental adaptation in Ethiopian indigenous chickens

F.G. Kebede[1,2*], M.F.L. Derks[1], T. Dessie[2], O. Hanotte[2,3], C. Pita Barros[1], R. P. M. A. Crooijmans[1], H. Komen[1], J. W. M. Bastiaansen[1]

[1]Wageningen University & Research, Animal Breeding and Genomics, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands; [2]International Livestock Research Institute, P. O. Box 5689, Addis Ababa, Ethiopia; [3]Cells, Organism and Molecular Genetics, School of Life Sciences, University of Nottingham, NG7 2UH, Nottingham, UK

To be submitted

Abstract

Local adaptation is relevant to sustainable livestock production in low- and medium-input tropical livestock systems. Randomly mating indigenous livestock populations are raised in stressful environmental conditions for many generations. Apart from selective forces, these populations are influenced by neutral evolutionary processes such as drift and gene flow. We followed a landscape genomic approach to integrate environmental, genomic, and phenotypic information and investigate phenotypic and genetic variation among Ethiopian indigenous chicken populations. A hybrid sampling strategy was applied to ensure populations from all environments and geographies are represented. Chickens were sampled from four environmental gradients, each of them representing six populations along an elevational cline in different geographies of the country. The environment in each cline was clustered into three agroecologies (two populations representing each agroecology). Signatures of selection analyses (F_{ST} and XP-EHH) were performed to detect footprints of natural selection, and redundancy analyses (RDA) were applied to detect genotype-environment and genotype-phenotype-associations. Structure analysis with PCA revealed a strong genetic grouping of six chicken populations distinct from the other 20 populations. The genetic structure of sample populations was carefully considered in the analysis. A total of 1909 outlier SNPs associated with six environmental predictors were identified by partial RDA. Using the same method, a total of 2430 outlier SNPs associated with five quantitative traits were detected. Eighty-three candidate SNPs that had moderate to high correlation ($|0.3| > r \leq |0.6|$) with mature body weight were identified. Higher average F_{ST} values were obtained for comparisons between any two gradients while lower values were obtained for comparisons between any two agroecologies. In contrast, stronger XP-EHH signals were observed for comparisons between agroecologies than between gradients. A large overlap was also found between signatures of selection identified by F_{ST} and XP-EHH, showing that both methods target similar selective sweep regions. Ethiopian chicken populations differentiated the most between gradients but selection pressures leading to adaptive variation are stronger between agroecologies. Higher genetic divergence between gradients suggests that evolutionary processes other than natural selection, such as gene flow and drift, have influenced sample populations in different geographies. Our results also show environmental and phenotypic predictors are informative to explain genomic variation in Ethiopian indigenous chickens. Signatures of selection analysis with the two methods (F_{ST} and XP-EHH) can be used complementarily with RDA to shed light on the relationship between genomic, phenotypic, and environmental variation in local adaptation studies in indigenous chickens.

3. 1 Introduction

Genetic and phenotypic differentiation in indigenous livestock populations is influenced by several evolutionary processes. Differentiation along environmental gradients, or across contrasting habitat types, can be indicative of local adaptation (Conover et al., 2009; Savolainen et al., 2013; Turesson, 1922). Understanding the genetics of phenotypic variation in livestock, and local adaptation in response to environmental variation helps to improve productivity and to address climate change (Kelly, 2019; Merilä and Hendry, 2014; Rovelli et al., 2020; Sgro et al., 2016). Environmental differences acting as a natural selective force can result in exceptionally strong genetic differentiation in genomic regions containing loci subjected to selection (Lewontin and Krakauer, 1973). Recent studies in non-domesticated animals (Bergland et al., 2014; Foucault et al., 2018; Lafuente and Beldade, 2019) suggest that phenotypes associated with environmental differences are in fact due to genetic differentiation as a result of selection. For instance, alleles providing adaptation to high elevation are found in high frequency in populations at high elevation but in low frequency in populations at low elevation in humans (Hackinger et al., 2016; Lorenzo et al., 2014).

The effects of environmental selective pressures as drivers of local adaptation and specially their influences on phenotypic and genetic differentiation in Ethiopian chicken populations has not been investigated comprehensively. Certain phenotypes in indigenous chickens (e.g., comb shape, parasitic resistance) are related with local adaptation (Bettridge et al., 2018). Genomic regions conferring adaptation to environmental challenges (e.g., elevation, temperature, water scarcity, and feed availability) have been identified in indigenous chickens (Elbeltagy et al., 2019; Fleming et al., 2017; Gheyas et al., 2021). Important insights on local adaptation of Ethiopian indigenous chickens were obtained in previous studies regarding genotype-environment and environment-phenotype associations. Kebede et al. (2021) detected the association between environmental predictors and differentiation of quantitative traits as an evidence for adaptative variation. Gheyas et al. (2021) investigated the relationship between genomic and environmental variation.

The present study follows a landscape genomics approach to integrate environmental, phenotypic, and genetic variation to analyse and explain population differentiation among Ethiopian indigenous chickens. Landscape genomics provides an analytical framework useful to investigate the underlying evolutionary processes behind phenotypic and genetic differentiation of random mating indigenous populations raised in heterogenous environments. We combine Species Distribution Modelling (SDMs), signatures of selection analyses, and association analyses to study

adaptive variation in Ethiopian chickens. SDMs are useful to identify environmental predictors associated with habitat suitability and local adaptation in indigenous livestock populations (Gheyas et al., 2021; Kebede et al., 2021; Lozano-Jaramillo, 2019; Vajana et al., 2018; Vallejo-Trujillo et al., 2018). Signatures of selection analysis identify regions of the genome that have differentiated between populations, possibly in response to selective pressure (Sabeti et al., 2006; Voight et al., 2006).

Multivariate methods that simultaneously account for multiple drivers of phenotypic and environmental divergence, are recently being applied in landscape genomic studies to identify quantitative trait loci (QTL) associated with environment predictors (Forester et al., 2018; Harrisson et al., 2017; Kess and Boulding, 2019; Torrado et al., 2020) and with phenotypic variables (Carvalho et al., 2021; Kess and Boulding, 2019; Talbot et al., 2017; Valette et al., 2020; Vangestel et al., 2018). Multivariate ordination methods such as RDA have outperformed mixed-model-based methods and machine learning-based methods (e.g., Random Forest) in detecting loci associated with environmental variation (Capblancq et al., 2018; Forester et al., 2018). Despite its ability to investigate genotype-phenotype associations RDA is mostly neglected in GWAS studies, while it became a standard in genotype-environment association studies (Jombart et al., 2009; Valette et al., 2020).

We implemented a robust sampling strategy, considering environmental gradation (e.g., elevational clines) and geographic (latitudinal and longitudinal) variation in the country (Kebede et al., 2021). This enabled us to survey all possible agroecologies and ecotypes. By adopting a hybrid sampling strategy, environmental and geographic representativeness of sampling sites is maximized, and statistical power is increased by reducing false discovery rates (De Mita et al., 2013; Lotterhos and Whitlock, 2015; Selmoni et al., 2020). Our analytical methods investigated the association between three sources of variation (environment, genotype, and phenotype). More specifically, the present study was undertaken in Ethiopian indigenous chickens to accomplish the following three objectives: 1) to identify candidate genes and genomic regions linked with environmental adaptation; 2) to explain variations in the genome by using environmental variables influencing habitat suitability as predictors (detect genotype-environment association); and 3) to explain variations in the genome by using quantitative traits associated with population phenotypic differentiation as predictors (detect genotype-phenotype association).

3.2 Materials and methods

3.2.1 Sampling strategy

The topographic map of Ethiopia (**Figure 3.1**) shows the 26 Ethiopian sample locations for indigenous chicken populations, and their environmental gradients. A hybrid sampling strategy was used to collect data, capturing environmental, phenotypic, and genetic variability (**Figure 3.1** and **Figure 3.2)**. The spatial distribution of samples considered environmental (e.g., geography, climate) and biotic processes (e.g., domestication, routes of introduction) influencing the chicken populations. For species with limited dispersal, sample sizes above 200 units are generally sufficient to detect most adaptive signals in landscape genomics, while in random mating populations this threshold should be increased to 400 units (Selmoni et al., 2020). A total of 513 chickens were sampled from four environmental gradients (*gradient-I, -II, -III, and -IV*) with a minimum distance between gradients of 500km. A gradient refers to an elevational cline located in a specific geography of the country. *Gradient-I* stretches from the Rift valley lowlands of northeastern Ethiopia along the territories of Afar region to the highlands of Wollo province within Amhara region. *Gradient-II,* starts from the Rift valley lowlands in central Ethiopia, crosses the highlands of Hararghe, including Mount Gara Muleta, and stretches to eastern Ethiopia within Oromia region. *Gradient-III* stretches from the highlands of northwestern Ethiopia and goes down to the lowlands along the Ethiopian-Sudanese border within Benishangul-Gumuz region. *Gradient-IV* extends from the highlands of western Ethiopia in Oromia region to the lowlands along the Ethiopian-Kenyan border in Southern region. Areas around the national borders of Ethiopia have low elevation, which gradually culminates to highland plateau in the center of the country creating a striking contrast in agroecology.

Each gradient comprised three environmental clusters or agroecologies, primarily delineated based on elevation. These are *lowland* (400-1800 m.a.s.l); *midaltitude/midland* (1800-2400 m.a.s.l.); and *highland* (2400-3500 m.a.s.l.) according to the conventional agroecological classification in Ethiopia (Dove, 1890; MoA, 2000). Clusters within a gradient were distant by at least 100km and farmers keeping target chicken populations within a cluster visited separate livestock markets. Each cluster along the spatial gradient constituted of 2-3 populations. The metadata of 513 individual samples is presented in **Supplementary Table 3.1**.
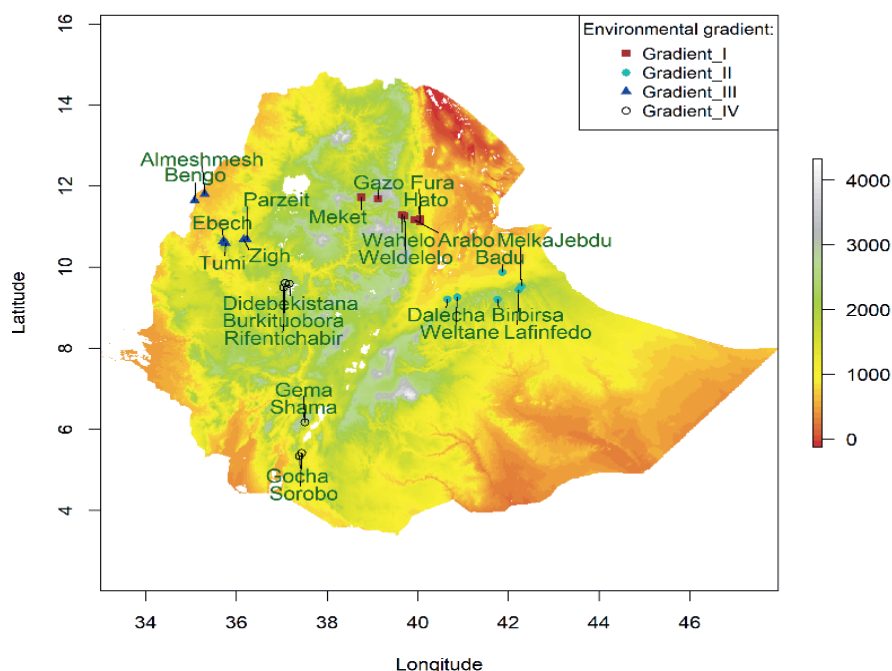
Figure 3.1. Topographic map of Ethiopia depicting the 26 Ethiopian indigenous sample populations and their environmental gradients. Range of numbers with different colours in the legend indicate elevation (m.a.s.l.).

Figure 3.2 shows the sampling and analytical framework used in the present landscape genomic study. The chicken populations from different geographies of Ethiopia may be the result of different evolutionary histories. We controlled for the potential confounding effects between demographic processes (e.g., domestication history, migration) and adaptive variation in our analysis by performing signatures of selection analyses at three different analytical layers (*layer-I, layer-II, and layer-III*).

Figure 3.2. Sampling and analytical framework in landscape genomics study to detect adaptive phenotypic and genetic variation in Ethiopian indigenous chicken populations LL=lowland; MA=midaltitude; and HL=highland

3.2.1.1 Environmental data

For every population, a single geographic coordinate was taken at the center of the village during sampling of chickens. Coordinates from nine additional grids (1.44km^2), covering 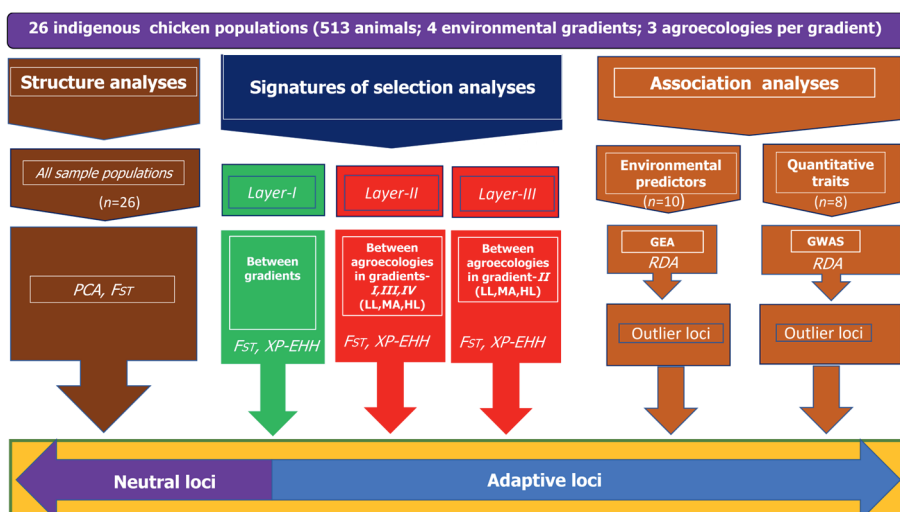a total of 12.96km^2, were then drawn around a recorded location and extracted using Google Earth Pro v 7.3.2 to ensure high representation of environmental variability affecting the population. Out of 34 environmental predictors, 9 predictors identified through species distribution models (SDMs) for their association with habitat suitability of chickens in Ethiopia (Kebede et al., 2021) were included in the present study for genotype-environment association analysis **(Supplementary Table 3.2)**. Additionally, considering its importance in the conventional definition of agroecologies in Ethiopia, its link with certain adaptive traits in chickens (Huang et al., 2017), and our sampling design along elevational clines, we incorporated elevation as a tenth environmental predictor. All the ten predictors were used to produce habitat suitability maps for the 26 sample chicken populations with MaxEnt computer algorithm (Phillips et al., 2006). Configuration of model parameters for MaxEnt was based on the settings used by Kebede et al. (2021).

3.2.1.2 Quantitative trait data

A population refers to all the nondescript indigenous chickens available for phenotyic and genetic sampling in an administrative village. Collection of phenotypic data was performed on adult chickens. These chickens were selected randomly through transect walk across villages. This method entailed walking along a defined path (transect) across a village and sampling one chicken from each farming household until a total of 15 hens and 5 cocks (roosters) were measured. The age of the chickens was estimated by interviewing owners to confirm that females were in their second clutch (7 to 8 months-of-age) and males were above 12 months-of-age. The researchers also visually appraised roosters for the presence of well-developed spurs. One chicken was sampled per household. Under rare circumstances ($n = 9$), two chickens were sampled per household when farmers proved their animals have no family relationship. 19 quantitative traits were initially measured on each of the 513 adult chickens. Out of these 19 quantitative traits, we used the 8 traits identified by (Kebede et al., 2021) for their putative roles in local adaptation and usefulness in phenotypic classification of Ethiopian chicken populations (**Supplementary Table 3.3**). These are livebody body, beak length, comb width, wattle width, earlobe width, wingspan, body length, and comb width. Live bodyweight of individuals was taken in the morning on fasting chickens. Accurate morphological measurements were made by digitally analysing the pictures of individual chickens photographed in a sheltered environment using ImageJ software (Rasband, 1997). To reduce systematic error, the same operator measured all chickens, which were held in the same position by a technician. A steel ruler was placed in every picture as a distance reference.

3.2.1.3 Blood sampling

Whole blood samples were taken from the wing vein of individual chickens in line with standard procedures (Grimes, 2002). 50 - 250 µl of whole blood with anticoagulant (K2EDTA) per sample was put into a cryo-tube filled with 1.5 ml absolute ethanol (100%). Samples were preserved at -20^{0}C until DNA extraction and processing.

3.2.2 Whole genome sequence and data processing

WGS data was generated on Illumina HiSeq2000 platform in paired-end mode with a read length of 150bp. Reads were quality trimmed using (Bolger et al., 2014). The average depth of coverage was 8.63 (range: 5.47-14.12) with an average mapping rate of 99.2% (97.05-99.6) and a mapping quality of 33.6 (28.77-34.45) to the GRCg6a reference assembly (Ensemble Gallus_gallus.GRCg6a.dna.toplevel.fa). Genomic analysis was performed on autosomes and non-autosomes.

3.2.3 Variant calling and functional annotation

Freebayes was used to call variants with the following settings: min-base-quality 10 --min-alternate-fraction 0.2 --haplotype-length 0 --ploidy 2 --min-alternate-count 2 (Garrison and Marth, 2012). Post processing was performed using BCFtools (Li, 2011). Variants with low phred quality score (< 20), low call rate (<0.7) and variants within 3 bp of an insertion-deletion (indel) were discarded.

3.2.4 Population structure analysis

PCA was performed using the Eigenstrat method in Eigensoft v 6.1.4 software (Patterson et al., 2006; Price et al., 2006) to understand the structure of the 26 populations.

3.2.5 Signatures of Selection Analysis (SSA)

The search for signals of positive selection was carried out on SNP data (*n*=25M) from WGS. Haplotypes were phased using FastPhase software prior to signatures of selection analyses. We performed signatures of selection analyses (F_{ST} and XP-EHH) under three different analytical layers (**Figure 3.2**). In *layer-I*, we classified the indigenous chicken populations into four gradients (without regard to their agroecologies) and compared them. In *layer-II*, we classified sample chicken populations into three agroecologies (lowland, midaltitude, and highland) across *gradients (-I, -III, and -IV)*, and compared them. In *layer-III*, we classified sample chicken populations into three agroecologies (lowland, midaltitude, and highland) within *gradient-II* and compared them.

Fixation test (F_{ST}) analysis was conducted using VCFtools v0.1.16 Danecek et al. (2011) to identify regions of increased genomic differentiation between the classifications defined in the analytical layers. We calculated the average F_{ST} values with overlapping windows of 50kb (25kb overlapping). We calculated the average XP-EHH values for the classifications defined in each of the analytical layers.

The same size of overlapping bins (50kb) was used for XP-EHH analysis to allow comparison with F_{ST} . First, the average (F_{ST} or XP-EHH) values for all bins in each pairwise comparison in an analytical layer were sorted on their significance. Empirical P-values were calculated for both F_{ST} and XP-EHH by ranking the windows based on each metric and dividing the rank by the total number of windows. Only the 1% most significant windows (*p* < 0.01 F_{ST} or XP-EHH) were retained as significant. Significant windows which were commonly identified by the two methods were counted as overlapping.

3.2.6 Association Analyses

Association analyses were performed using Redundancy Analysis (RDA) with the R package 'vegan' (Oksanen et al., 2013) to identify environmental predictors and quantitative traits associated with genomic variation. Environmental predictors and quantitative traits were analysed separately according to Forester (2019). Genotypes were filtered for SNPs not in Hardy-Weinberg equilibrium ($p < 5x10^{-6}$), with MAF <5%, and by LD pruning of SNPs. Individuals with missing genotypes > 5% were removed, resulting in a cleaned dataset with 1,070,305 SNPs and 466 individuals. The genetic dataset was structed as a matrix of 466 chickens by ~1 million SNP markers.

3.2.6.1 Genotype-environment association (GEA) analysis with RDA

Correlated predictors cause problems for regression-based models like RDA and variable reduction was done using the $r > |0.7|$ "rule of thumb" (Dormann et al., 2013) to retain ecologically relevant but not highly correlated environmental predictors. We fitted partial RDA with the 10 selected environmental predictors conditioned on (i.e. controlling for the effects of) geography as explanatory variables and the genetic dataset as response variable (Rellstab et al., 2015). SNPs exhibiting RDA loadings greater than 3.5 standard deviations (two-tailed p-value = 0.0005) from the mean were identified as selection signals This threshold is very conservative and helps to identify loci under strong selection (i.e. minimizes false positive rates (Forester et al., 2018). After a visual inspection of the scree plots, we extracted SNP loadings from the first three canonical axes.

3.2.6.2 Genotype-phenotype association analysis with RDA

We fitted partial RDA with the 5 least correlated and most explanatory quantitative traits selected by correlation analysis. The RDA were fitted with the quantitative traits as explanatory variables, conditioned on geography, and the genetic dataset as response variable. SNPs exhibiting RDA loadings greater than three and half standard deviations from the mean were identified as association signals in line with Forester et al. (2018). After a visual inspection of the scree plots, we extracted SNP loadings from the first three canonical axes.

3.3. Results

3.3.1 Habitat suitability

The suitability of an environmental niche for a population depends on which environmental predictors are influencing the species. The habitat suitability maps produced by species distribution models (SDMs) suggests that the 26 populations

have different niches. They likely went through different environmental selective pressures which may give rise to phenotypic and genetic differentiation (**Figure 3.3**).
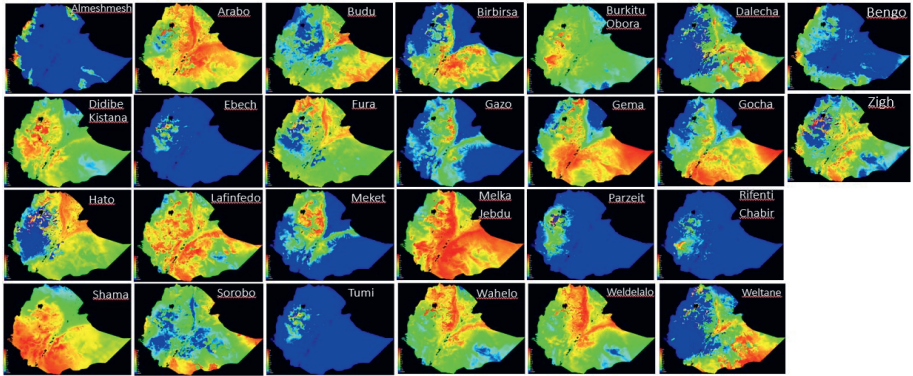

Figure 3.3. Habitat suitability maps of the 26 Ethiopian chicken populations. Colours towards red spectrum indicate more suitable conditions.

3.3.2 Genomic diversity of Ethiopian indigenous chickens

Analysis of the WGS data resulted in 25M SNPs. Information on genome coverage, mapping rate and quality of samples is presented in **Supplementary Table 3.4**. The genetic structure of the 26 Ethiopian indigenous chicken populations analysed by PCA clearly shows separation between the populations of *gradient-II* (*n*=6) and the populations of the other three gradients (*n*=20) (**Figure 3.4A** and **Figure 3.4B**). We therefore decided that sample populations from *gradient-II* should not be analysed together with the other populations. Some admixture is seen among sample populations taken from *gradients -III* and *-IV*. *Gradient-I* showed some distinctiveness from the rest, with a slight mix with *gradient-IV*.
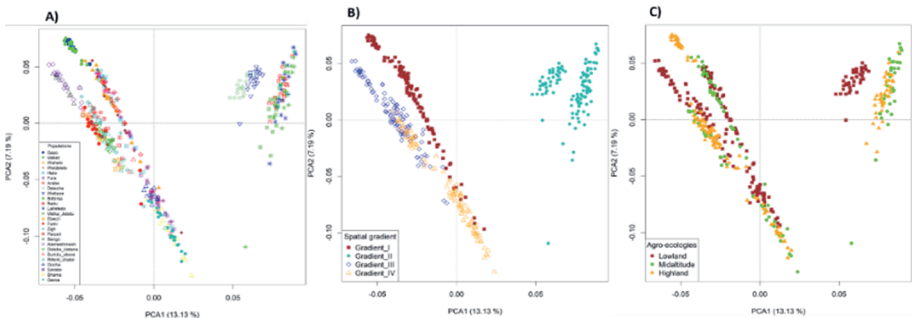

Figure 3.4. PCA plots of Ethiopian indigenous chicken populations based on 25 million SNPs. **A)** PCA plot labelled with 26 sample chicken populations; **B)** PCA plot labelled with four environmental gradients (elevational clines in four different geographies); **C)** PCA plot labelled with three agroecologies (lowland, midaltitude and highland).

Figures 3.4B and **C** illustrate that samples from lowland agroecologies of gradient-*II* and -*III* are distinct with in their gradient while considerable mixing was observed between populations sampled from midaltitude and highlands in all gradients. In *gradient-I* a mix was also seen between lowland and highland populations.

3.3.3 Signatures of selection for environmental adaptation

3.3.3.1 F_{ST}

Genetic differentiation between gradients (*analytical layer -I*)
The Manhattan plots of F_{ST} analyses show pairwise comparison between populations sampled from environmental *gradients -I, -III,* and -*IV* (**Figure 3.5**).

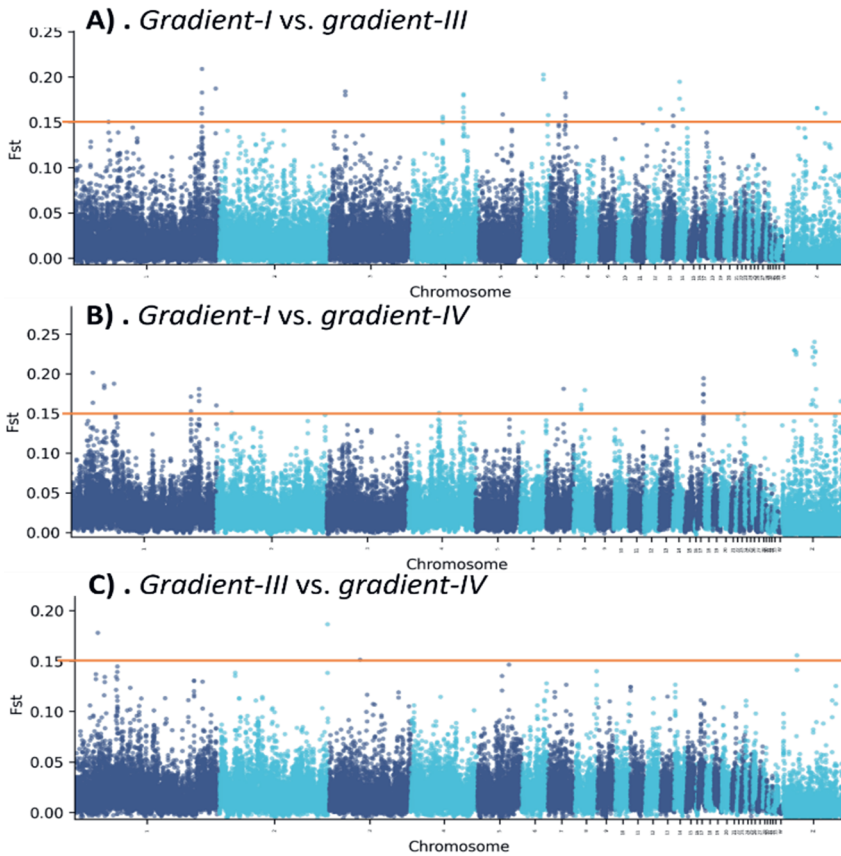**Figure 3.5.** Manhattan plots of Fixation index (F_{ST}) for overlapping bins of 50kb showing pairwise comparisons of Ethiopian indigenous chicken populations between *gradients (-I, -III, and -IV)*. Points above the horizontal line surpass the 0.15 average F_{ST} value. **A).** *Gradient-I* vs *gradient-III* **B).** *Gradient-I* vs *gradient-IV*; and **C).** *Gradient-III* vs *gradient-IV*.

An F_{ST} value greater than 0.15 can be considered as significant in differentiating populations (Frankham et al., 2002). Thus, genomic regions with significant divergence were found between populations sampled from any two gradients. Higher average F_{ST} value between *Gradient-I* vs *gradient-III* suggests that these populations have differentiated the most (**Figure 3.5A)**. Local signatures of selection are seen on 10 different chromosomes for comparison between these two gradients. On the other hand, significant divergence between populations of *gradient-I* and *gradient-IV* was mainly seen due to differentiation of regions on 4 different chromosomes (1,8,17 and Z) (**Figure 3.5B)**. Low differentiation was found between populations of *gradient-III* and *gradient-IV* (**Figure 3.5C)**. F_{ST} scores for comparisons between gradients and the identified genes in the significant bins are presented in **Supplementary Table 3.5.**

Genetic differentiation between agroecologies across gradients (*analytical layer-II*)

The F_{ST} scores for comparisons across agroecologies (**Figure 3.6**; **Supplementary Table 3.6**) are lower (average F_{ST} value below 0.15) compared to the scores between gradients, suggesting relatively low genetic differentiation between populations sampled from any two agroecologies. The populations sampled from the lowlands and the highlands had the lowest population differentiation **(Figure 3.6A)**. The results show clear signals of differentiation in the lowland vs midland comparison on chromosomes 1 and 2 (F_{ST} values close to 0.12; **Figure 3.6B)**. Similarly, two clear signals of differentiation were identified. SNPs with F_{ST} values closer to 0.12 were observed on chromosomes 2 and 20 in the midland vs highland comparison.**(Figure 3.6C)**.

Figure 3.6. Manhattan plots of Fixation index (F_{ST}) for overlapping bins of 50kb showing pairwise comparison between agroecologies (lowland, midaltitude, highland). for Ethiopian indigenous chicken populations sampled from gradients -*I,-III,* and -*IV* together **A).** Lowland vs highland; **B).** Lowland vs midland; and **C).** Midland vs highland.

Genetic differentiation between agroecologies within *gradient-II* (*analytical layer-III*)

The F_{ST} scores for comparisons between agroecologies within *gradient-II* (**Figure 3.7, Supplementary Table 3.6**) are generally higher than the scores for agroecological comparisons across the other gradients (*Layer-II*).
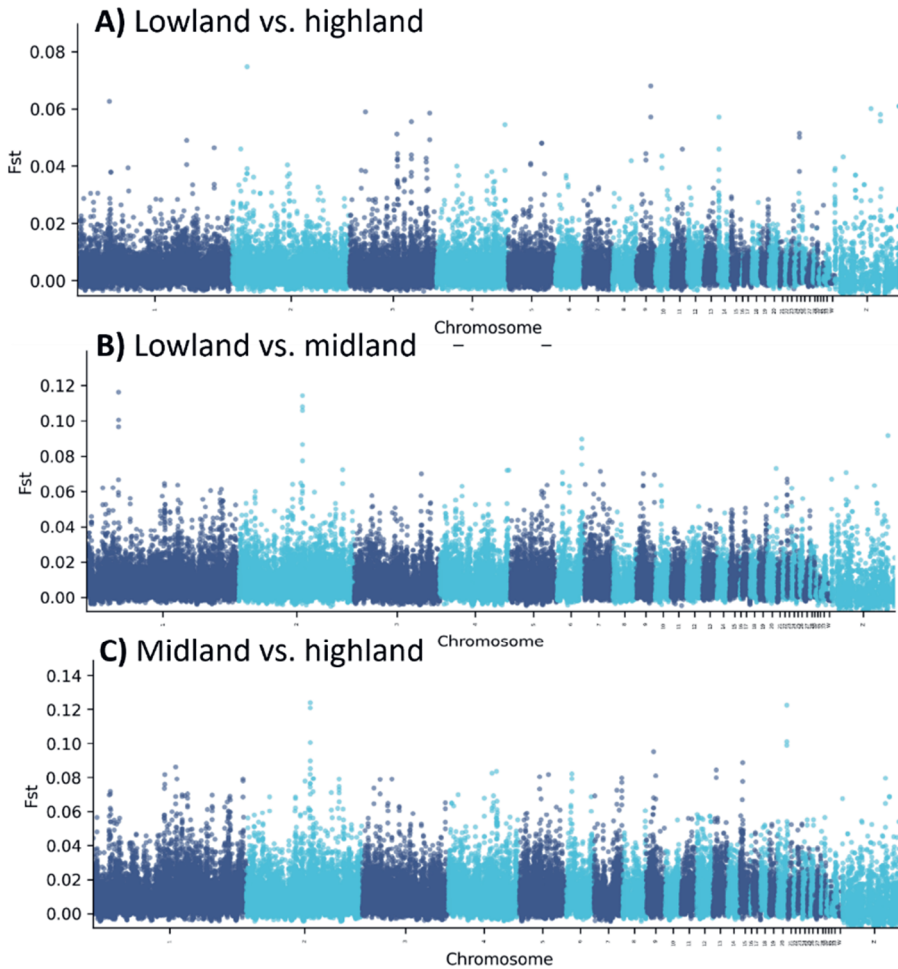
60

Figure 3.7. Manhattan plots of Fixation index (F_{ST}) for overlapping bins of 50kb showing pairwise comparison of Ethiopian indigenous chicken populations between agroecologies in *gradient-II* (lowland, midaltitude, highland). **A).** Lowland vs highland; **B).** Lowland vs midland; and **C).** Midland vs highland.

Lowland and highland populations have differentiated the most on regions on chromosomes 7,8 and Z (**Figure 3.7A**). Regions on chromosomes 1,3,5,7,8,9, and Z have shown differentiation in the lowland-midland comparison (**Figure 3.7B**). Midland vs highland comparison yielded low average F_{ST} scores across the genome (below 0.15) suggesting the presence of low genetic differentiation between the populations sampled from these two agroecologies of *gradient-II* (**Figure 3.7C**).

3.3.3.2 XP-EHH
Selection signatures between gradients (*analytical layer -I*)

XP-EHH detected signatures of selection between populations sampled from two gradients **(Figure 3.8)**.

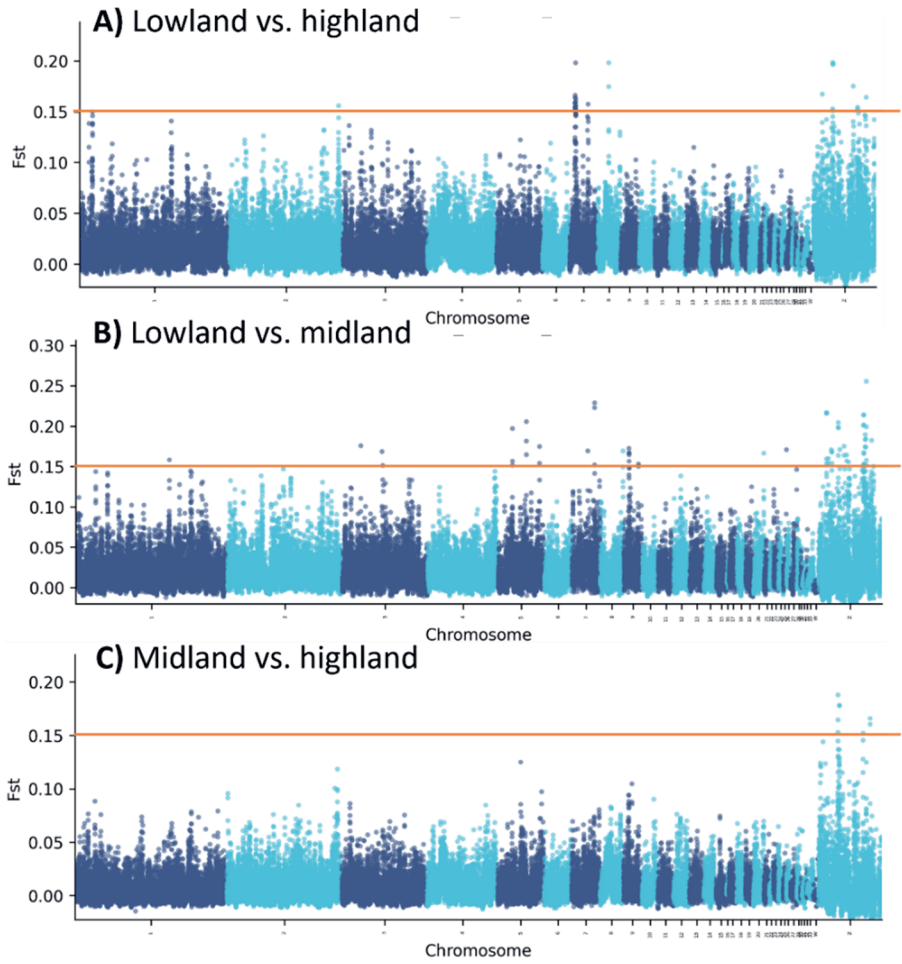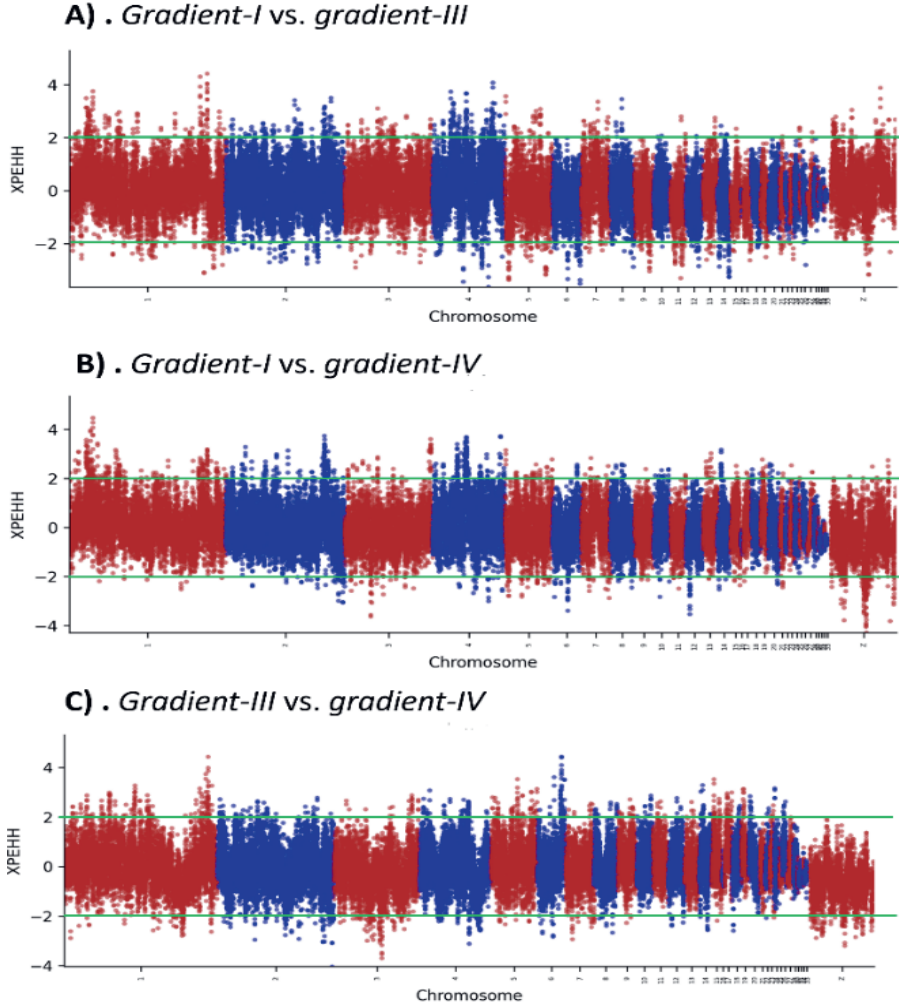**Figure 3.8.** Manhattan plots of XP-EHH for overlapping bins of 50kb showing pairwise comparisons for Ethiopian indigenous chicken populations between *gradients (-I, -III, and -IV)*. The y-axis shows -log (p-value), with positive values identifying extended homozygosity in populations from the first gradient relative to populations from the second gradient, and vice versa for negative values. All SNPs with a -log(p-value) above 2 or below -2 from the green line are significantly selected (p <0.01) in one gradient but not in the other. **A)**. *Gradient-I* vs *gradient-III*; **B)**. *Gradient-I* vs *gradient-IV*; **C)**. *Gradient-III* vs *gradient-IV*.

XP-EHH values greater than 2 (above the green line) or -2 (below the green line) on the plots indicate a difference in the top 1% and is considered as significant in differentiating the populations. Haplotypes on 14 out of 35 chromosomes were selected in *gradient-I* compared to *gradient-III* **(Figure 3.8A)**. On the other hand, haplotypes on 21 out of 35 chromosomes were selected in *gradient-III* compared to *gradient-I*. Regions on 20 chromosomes were under selection in *gradient-I* compared to *gradient-IV* and 17 regions were under selection in *gradient-IV* compared to *gradient-I* **(Figure 3.8B)**. The cross-population comparisons of gradients also show 22 regions were selected in *gradient-III* while 15 were selected in *gradient-IV* **(Figure 3.8C)**. XP-EHH scores for gradient-based comparisons (*Layer-I*) are presented in **Supplementary Table 3.7**.

Selection signatures between agroecologies across gradients (*analytical layer-II*)
XP-EHH detected strong signatures of selection between populations sampled from two different agroecologies (**Figure 3.9**).

Figure 3.9. Manhattan plots of XP-EHH for overlapping bins of 50kb showing pairwise comparison between agroecologies (lowland, midaltitude, highland) for Ethiopian indigenous chicken populations sampled across three gradients (-*I,-III,* and -*IV*). The y-axis shows -log(p-value), with positive values identifying extended homozygosity in populations from the first agroecology relative to populations from the second agroecology and vice versa for negative values. All SNPs with a -log(p-value) above 2 or below -2 from the green line are significantly selected (p <0.01) in one agroecology but not in the other. **A).** Lowland vs highland; **B).** Lowland vs midland; and **C).** Midland vs highland.

XP-EHH scores for comparisons between agroecologies across the three gradients (*Layer-II*) are presented in **Supplementary Table 3.8**. Haplotypes on 20 chromosomes were selected in highland and 17 were selected in lowland (**Figure 3.9A**); 22 haplotypes were selected in lowland and 15 were selected in midland (**Figure 3.9B**); and 18 haplotypes were selected in midland and 14 were selected in highland (**Figure 3.9C**).

Selection signatures between agroecologies within *gradient-II* (*analytical layer-III*)

Sharp peaks were seen on many chromosomes across the genome for comparisons between any two agroecologies within *gradient-II* suggesting regions under selection **(Figure 3.10)**. XP-EHH scores for comparisons between agroecologies in *gradient-II* are presented in **Supplementary Table 3.8**. More number of chromosomes with regions under selection were found during agroecological comparison within *gradient-II* than agroecological comparison across gradients. Haplotypes on 23 chromosomes were selected in highland and 21 were selected in lowland (**Figure 3.10A**); 19 haplotypes were selected in lowland and 21 were selected in midland (**Figure 3.10B**); and 20 haplotypes were selected in midland and 22 were selected in highland (**Figure 3.10C**).
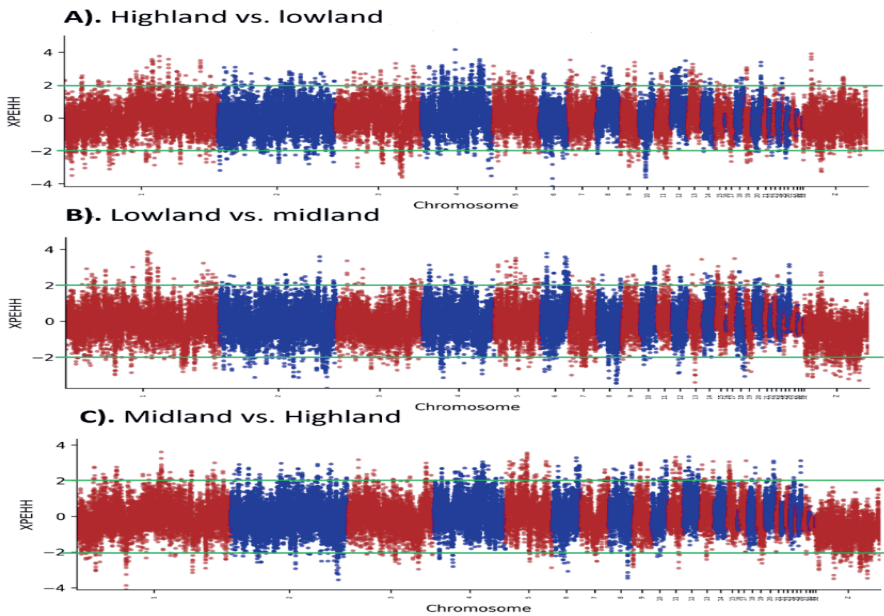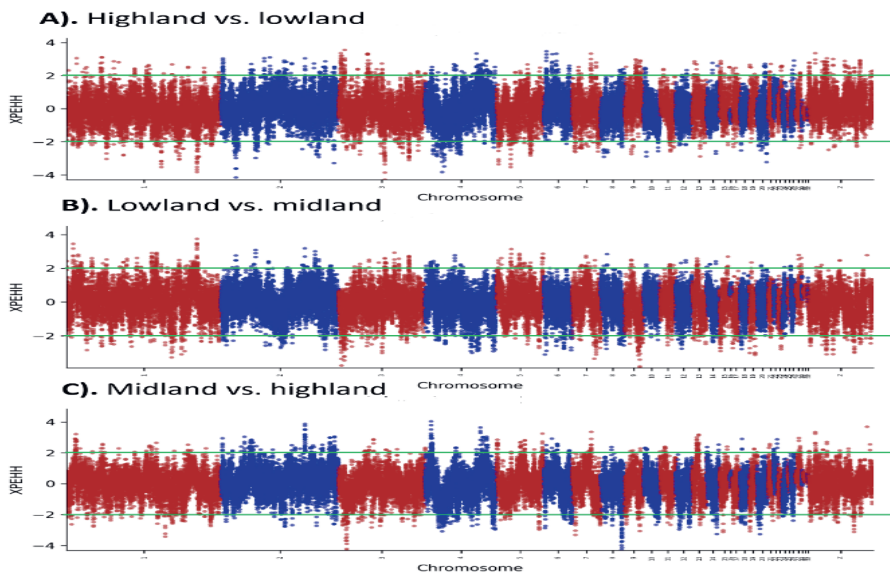

Figure 3.10. Manhattan plots of XP-EHH for overlapping bins of 50kb showing pairwise comparison between agroecologies (lowland, midaltitude, highland) for Ethiopian indigenous chicken populations sampled from gradients *-II.* The y-axis shows -log (p-value), with positive values identifying extended homozygosity in populations from the first agroecology relative to populations from the second agroecology and vice versa for negative values. All SNPs with a -log(p-value) above 2 or below -2 from the green line are significantly selected (p <0.01) in one agroecology but not in the other. **A).** Lowland vs highland; **B).** Lowland vs midland; and **C).** Midland vs highland.

3.3.4 The overlap between signatures of selection analyses (F_{ST} and XP-EHH)

3.3.4.1 Overlaps between gradients *(analysis layer -I)*

The overlap between significant ($p < 0.01$) F_{ST} and XP-EHH windows in pairwise gradient comparisons are presented in Venn diagrams (**Figure 3.11).** A relatively large overlap (19.2%) was observed between significant windows identified by F_{ST} and XP-EHH analyses in the pairwise gradient comparisons. There was 18.4 % overlap between *gradient- I* and *-III*, 16.8% between *gradient- I* and *-IV*, and 22.3% overlap between *gradient- III* and *-IV* comparison.

The overlap is the highest between *gradient-I* and *-IV.* A complete list of significant genes ($p < 0.01$) from overlapping windows jointly identified by F_{ST} and XP-EHH in each gradient-wise comparisons is presented in **Supplementary Table 3.9.**

Figure 3.11. Venn diagrams showing overlaps between significant ($p < 0.01$) F_{ST} and XP-EHH windows per pairwise gradient comparisons. **A)** Between *gradient- I* and *-III* **B)** Between *gradient- I* and *-IV* **C)** Between *gradient- III* and *-IV.*

There were 18 genes present in overlapping bins from more than one pairwise comparison between gradients (**Table 3.1**.) Most of these genes (n=13) were present in windows identified in comparisons involving *gradient-I*.

Table 3.1. Genes in windows that are significant for both F_{ST} and XP-EHH in more than one gradient comparison.

Gradient Comparisons	No. of windows	Windows	Genes
I_III; I_IV	16	1_174025001, 4_70325001, 4_42175001, 1_173725001, 4_42200001, 8_14300001, 1_174050001, 1_173700001, 1_23025001, 4_42100001, 1_173750001, 1_174075001, 4_42275001, 4_42300001, 3_41375001, 1_23000001,	AGA, NEIL3, ENSGALG0000005423 1, SMAD9, RFXAP, EXOSC8, CO, FNBP1L, DR1, ENSGALG0000003231 9, DCLK1, AASS, PTPRZ1
III_IV; I_IV	6	4_31000001, 1_58225001, 3_108550001, 2_147200001, 1_58250001, 2_147225001,	TSNARE1, MTPN, ANAPC10, ABCE1, OTUD4

3.3.4.2 Overlaps between agroecologies across gradients (*analytical layer-II*)

Higher number of overlapping windows were observed between the two methods of signatures of selection analysis in lowland vs highland and in midland vs highland comparisons compared to lowland vs midland (**Figure 3.12**). A large overlap (13.4%) was observed between significant windows ($p < 0.01$) identified by F_{ST} and XP-EHH analyses in the pairwise agroecological comparisons across gradients. There was 14.7% overlap between methods for lowland vs highland, 11.2% overlap between lowland vs midland, and 14.3% overlap between midland vs highland comparison.

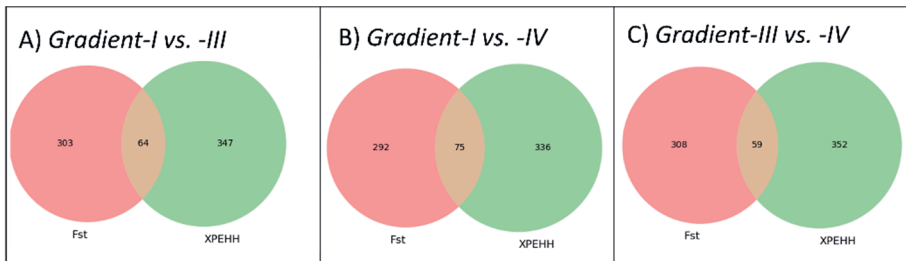**Figure 3.12.** Venn diagrams showing overlaps between significant ($p < 0.01$) F_{ST} and XP-EHH windows per pairwise comparison between agroecologies across gradients (-*I, -III,* and -*IV*). **A)** Lowland vs highland **B)** Lowland vs midland **C)** Midland vs highland

Complete list of genes from overlapping windows jointly identified by F_{ST} and XPEHH in agroecological comparisons in lowland vs highland, lowland vs midland, and midland vs highland respectively across the three gradients (*layer-II*) are presented respectively in **Supplementary Table 3.10A-C.**

3.3.4.3 Overlaps between agroecologies within *gradient-II* (*analytical layer-III*)

The results from within gradient agroecological analysis show that the overlap between F_{ST} and XPEHH (**Figure 3.13**) increased compared to agroecological analysis across gradients (**Figure 3.12**). Selection signatures between agroecologies across gradients were diluted by genetic differentiation associated separate geographies. An overall overlap of 20.9% was observed between significant windows ($p < 0.01$) identified by F_{ST} and XP-EHH analyses in the pairwise agroecological comparisons within *gradient-II*. There was 25.1 % overlap between methods for lowland vs highland, 14.5 % overlap between lowland vs midland, and 23.1% overlap between midland vs highland comparison.

Complete lists of genes from overlapping windows jointly identified by F_{ST} and XPEHH in agroecological comparisons in lowland vs highland, lowland vs midland, and midland vs highland respectively in *gradient-II* (*layer-III*) are presented in **Supplementary Table 3.10D-F.**
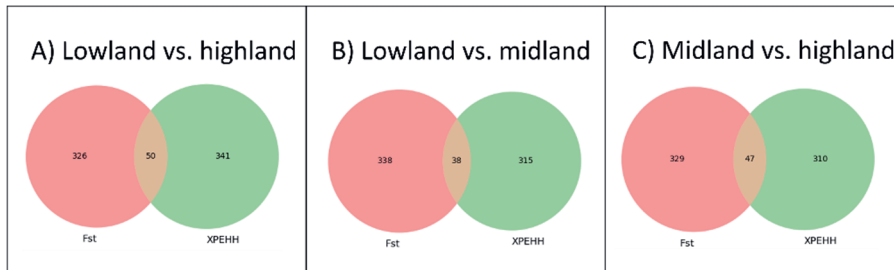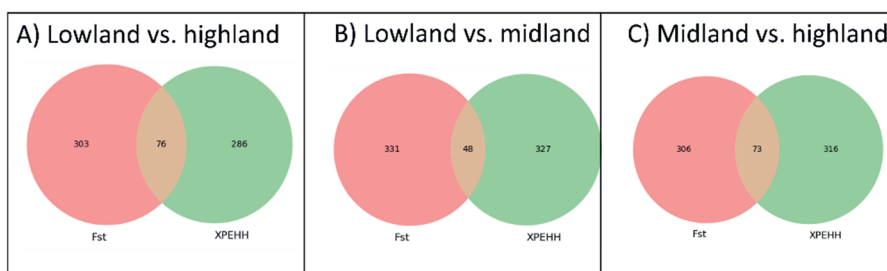

Figure 3.13. Venn diagrams showing overlaps between significant ($p < 0.01$) F_{ST} and XP-EHH windows per pairwise comparison between agroecologies within gradient-*II*. **A)** Lowland vs highland **B)** Lowland vs midland **C)** Midland vs highland

3.3.5 Genotype-environment associations (GEA)

Out of a total of 9 environmental predictors identified through MaxEnt-based species distribution models (SDMs) for their association with habitat suitability of chickens (Kebede et al., 2021) and elevation (added as a tenth predictor), 6 less correlated ($r \leq |0.7|$) predictors were retained for redundancy analysis (RDA) (**Supplementary**

Figure 3.1). These predictors were precipitation of the warmest quarter, precipitation of the coldest quarter, solar radiation of May, elevation, soil clay content and temperature seasonality. We had as many RDA axes as we had predictors (*n*=6) in our model. The first three RDA axes explained more than half (68.1%) of the variance in the environmental predictors (**Supplementary Table 3.11).** The adjusted R^2 considering the number of environmental predictors was 0.02, meaning that our constrained ordination explains about 2% of the variation or that 2% of the SNP variation is associated with the environmental predictors. Based on the magnitude of the arrows in PCA plots based on RDA axes 1 and 2(**Supplementary Figure 3.2**) elevation, precipitation of the warmest quarter, and soil clay content had the highest contributions to genotypic variation, while temperature seasonality and solar radiation had the lowest contributions.

The SNP loadings for environmental predictors on each of the three RDA axes show a relatively normal distribution (**Supplementary Figure 3.3**).The 1,909 SNPs from the two extreme ends of the loading distribution with standard deviation > 3.5 (two-tailed p-value = 0.0005) for each significant axis were taken as outlier SNPs that are associated with environmental variation. The list of candidate SNPs which have significant association (*p* < 0.001) with the six environmental predictors and are considered to be under selection are presented in **Supplementary Table 3.12**.

SNPs associated with the combined set of environmental predictors in *gradients -I,III, and -IV* do not show a clear clustering but are more or less evenly spread across the genome (**Figure 3.14**.).
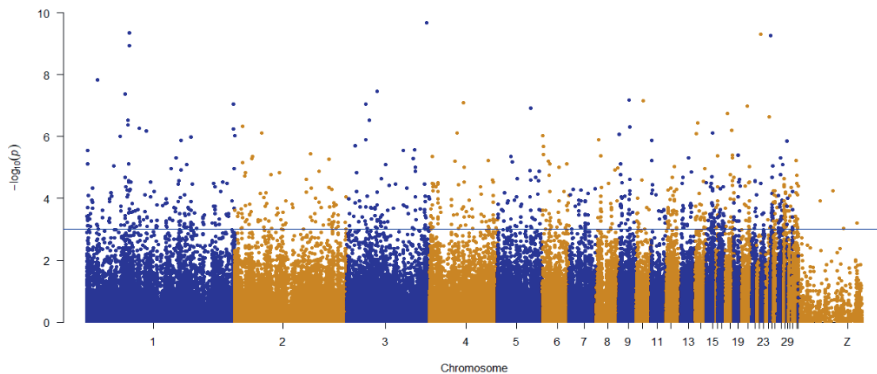
Figure 3.14. Manhattan plot of RDA showing the association of SNPs with the combined set of six environmental predictors in the three gradients (*-I, -III,* and *-IV*) as explanatory variables. The y-axis indicates -log 10 (p-value). Horizontal blue line indicates the significance threshold (p < 0.001).

Some of the highest -log10(p-values) are found on chromosomes 1 and 3 (**Figure 3.14**.). Only the peak on chromosome 1 shows additional significant SNPs near the top SNP. The significant candidate SNPs (n=1,909) that are associated with the combined set of environmental predictors are assigned to individual predictors based on the correlation values estimated by partial RDA analysis (**Figure 3.15**). Most candidate SNPs (942 or 49.3 %) have their highest correlation with elevation. Elevation has also the highest number (*n*=321 or 57.4%) of the moderately to highly associated SNPs (*n*=559) (0.3 < r < 0.6). The second environmental predictor most associated with candidate SNPs is precipitation of the warmest quarter. It has correlation with 410 candidate SNPs (21.47 %) . The other 4 environmental predictors have the highest correlation for a smaller number of SNPs (*n*=557 or 29.17%), but for all predictors considerable number of SNPs (*n*=59) are found with correlations above |0.3| and only two SNPs have correlations above |0.4|.

Figure 3.15. Number of significant candidate SNPs (*p* < 0.001) that are most correlated with each of the six selected environmental predictors grouped by absolute magnitude of their correlation.

3.3.6 Genotype-phenotype association

Out of a total of 8 phenotypic variables identified through MaxEnt-based species distribution models (SDMs) for their utility in phenotypically discriminating study populations (Kebede et al., 2021), five least correlated (|*r*| ≤ 0.72) quantitative traits were selected to be used for RDA **(Supplementary Figure 3.4).** These five traits were mature live body weight, beak length, comb width, wattle width and earlobe width. The correlation between comb width and wattle width was 0.72 which is slightly

higher than the common threshold ($|r| > 0.7$) used to reduce variables as a rule of thumb. But we decided to keep both of these traits because of their adaptive roles documented in literature related with thermoregulation in tropical chickens. The first three RDA axes explained most of the variance (62.1%) in the phenotypic predictors (**Supplementary Table 3.13**). The adjusted R^2 for the partial RDA was 0.002. This shows that only 0.2% of the SNPs variation is associated with quantitative traits.

The SNP loadings for quantitative traits on each of the three RDA axes show a relatively normal distribution (**Supplementary Figure 3.5).** Based on the magnitude of the arrows in the PCA plots based on RDA axes 1 and 2 (**Supplementary Figure 3.6**), comb width, wattle width and body weight were most useful in explaining SNP variation. SNPs associated with the combined set of quantitative traits in *gradients - I, -III,* and *-IV* show strong supportive peaks on chromosomes 1,3, 4, 7,8, 13, 15, and 29 indicating probable regions of quantitative trait loci (QTL) associated with phenotypic variation (**Figure 3.16).**
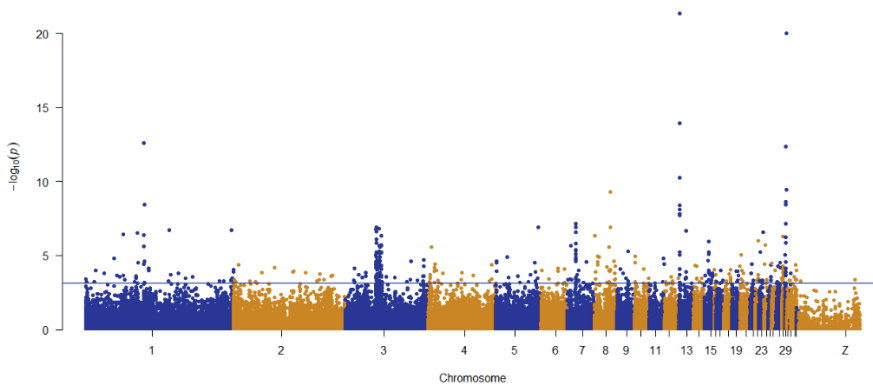
Figure 3.16. Manhattan plot of RDA showing the association of SNPs with phenotypic variation in the five quantitative traits in gradients *-I, -III,* and*-IV*. The y-axis indicates -log 10 (p-value). Horizontal blue line indicates the significance threshold (p < 0.001).

The significant candidate SNPs (p < 0.001) associated with the combined set of quantitative traits are assigned to individual traits based on correlation values estimated by partial RDA analysis (**Figure 3.17).** Partial RDA identified 1340 candidate SNPs that have significant association with the five quantitative traits (Supplementary table 3.14). A total of 19 SNPs have moderate to high correlation

with body weight ($0.3 < r \leq 0.6$). Most candidate SNPs, 39%, were associated with comb width (n=519) and 27% body weight (n=360).

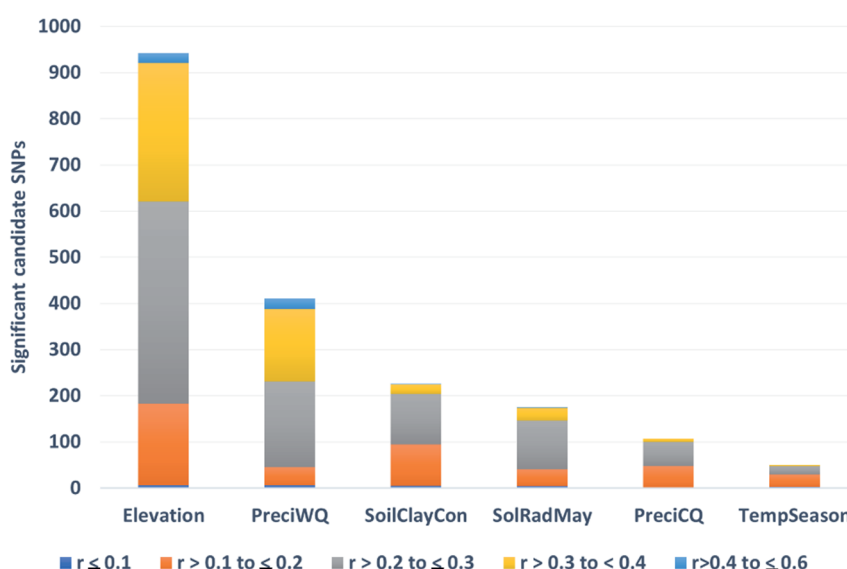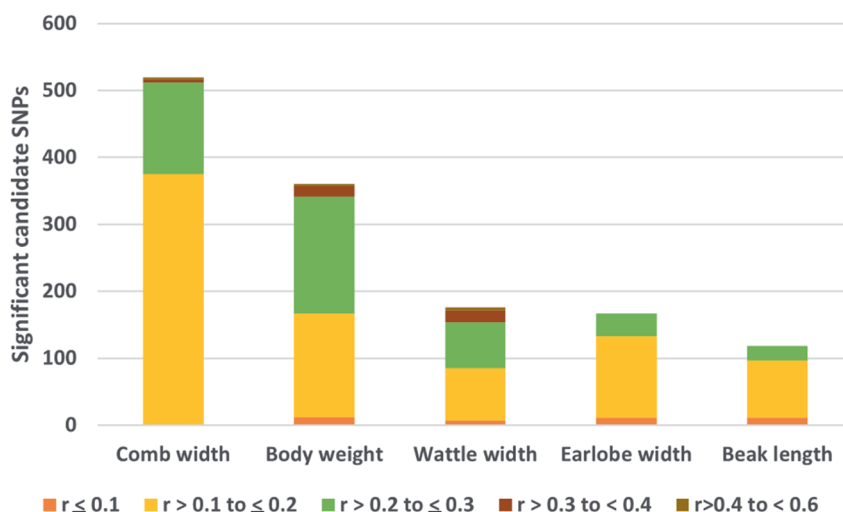**Figure 3.17**. Number of significant candidate SNPs ($p < 0.001$) that are most correlated with each of the 5 quantitative traits, grouped by absolute magnitude of their correlation

3.3.7 Signatures of selection and genotype-environment association

F_{ST} and XP-EHH values were calculated for candidate SNPs identified by genotype-environment association analysis (GEA). The range of F_{ST} and XP-EHH values vary depending on the analytical layer used for comparison (gradient vs agroecology). The F_{ST} values for comparisons between gradients ranged from 0 to 0.18 while the values for agroecological comparison across gradients ranged from 0 to 0.06. The result shows that, based on RDA identified SNPs, populations differentiated more between gradients and less between agroecologies. The list of significant ($p < 0.001$) candidate SNPs identified by RDA and their respective F_{ST} values are presented in **Supplementary Table 3.15.** The XP-EHH values ranged from -4.0 to 3.6 for pairwise comparisons for lowland vs midland; -3.4 to 4.9 for lowland vs highland; and -4.0 to 3.4 for midland vs highland .The XP-EHH values for pairwise comparison between gradients ranged from -3.5 to 4.4 for *gradient-I* vs-*III*; -3.0 to 3.95 for *gradient -I* vs -*IV* ; and -3.7 to 3.55 for *gradient -III* vs -*IV*. The list of significant candidate SNPs identified by RDA and their respective F_{ST} values are presented in **Supplementary Table 3.16**.

3.4. Discussion

Landscape genomics combines environmental, genomic, and phenotypic information to study adaptive variation. We combined different techniques including Species Distribution Models (SDMs), genetic differentiation test (F_{ST}), cross-population Extended Haplotype Homozygosity (XP-EHH), and redundancy analysis (RDA) to study local adaptation in Ethiopian indigenous chicken populations. SDMs were used to identify the most important environmental predictors influencing habitat suitability. Habitat suitability maps produced based on these predictors show the environments in the areas where the chicken populations were sampled from vary in terms of their suitability to chickens, with possible consequences on adaptive genetic and phenotypic variation.

The hybrid sampling strategy that was implemented improved representation of populations from different environments and geographies. The chicken populations were compared at three analytical layers: across gradients (*analytical layer-I*), between agroecologies across gradients in different geographies (*analytical layer-II*), and between agroecologies within *gradient-II* (*analytical layer-III*). The gradient-wise comparisons were considered to detect the influence of different evolutionary processes apart from natural selection. Pairwise comparison of F_{ST} is valuable to detect differentiation of populations in distinct environments (agroecologies, gradients) due to differences in evolutionary history (geography, demography) (Nei, 1986). Chromosomal regions with exceptionally high genetic differentiation (F_{ST} outliers) represent candidate genomic regions contributing to local adaptation (Hohenlohe et al., 2010; Wang et al., 2016). F_{ST} values in gradient-based comparisons (*layer I*) were significant (greater than 0.15) suggesting that, apart from adaptive processes (natural selection), genetic differentiation among Ethiopian indigenous populations is influenced by neutral processes in specific geographies (e.g. gene flow, genetic drift, demographic history). On the other hand, average F_{ST} values were smaller (lower than 0.15) in agroecology-based comparisons across gradients (*layer-II*).

In contrast to the F_{ST} results, strong signals of selection ($p < 0.01$) were detected by XP-EHH in pairwise agroecological comparisons than in gradient-wise comparisons. The XP-EHH results show that selective pressure in Ethiopian chickens are stronger between agroecologies. Classifying sample chicken populations into three agroecologies (lowland, midaltitude, and highland) along each gradient was intended to detect selection signals driven by agroclimatic variation. Positive selection leaves a more conspicuous footprint in the genome that can be detected by signatures of selection analysis (Pavlidis and Alachiotis,

2017; Ronen et al., 2013). XP-EHH is useful to detect differential selection between two populations (Sabeti et al., 2007).

A large overlap was observed between significant windows identified by F_{ST} and XP-EHH analyses, suggesting that both methods identified similar regions in the genome are under selection. The overlap between the two methods was the highest in agroecological comparisons for the three combined gradients. An even higher number of chromosomes with regions under selection was found in the XP-EHH analyses between agroecologies within *gradient-II*. A possible explanation is that selection signals were detected better when the influence from neutral processes (e.g., gene flow, and demographic history) associated with separate gradients was removed. Additionally the distinct demographic history of populations in *gradient-II* may have contributed.

The overlap between F_{ST} and XP-EHH analyses ranged from 13.4% to 20.9% between agroecologies which is considerably higher than the 4.9% overlap reported by Gheyas et al. (2021) between F_{ST} and XP-EHH for Ethiopian chickens. The large overlap between F_{ST} and XP-EHH in the present study might be due to our sampling strategy. Firstly, the sampling design captured a wide range of geographic and environmental variation and helped to survey all possible ecotypes and agroecologies in the country. Secondly, the design minimized confounding between neutral and adaptive processes which could have resulted from mixing of populations that have different demographic histories. By classifying the populations by gradients, we controlled for the effects of population genetic structure associated with specific geographies. For instance, a very high overlap between the F_{ST} and XP-EHH results was found in agroecological comparisons within *gradient-II*. The decision to analyse this gradient on its own was informed by PCA, which clearly separated populations of *gradient-II* from the other three gradients (-*I*, -*II*, and -*III*). *Gradient-II* represents chicken populations from eastern parts of Ethiopia which have a distinct evolutionary history and route of introduction into the country (Lawal et al., 2020; Mwacharo et al., 2013; Mwacharo et al., 2011) in contrast to populations representing the other three gradients. Combining *gradient-II* with the other three was expected to reduce the overlap of F_{ST} and XP-EHH results.

F_{ST} and XP-EHH analyses identify selective pressures, but to investigate how this selection is driving local adaptation we carried out association analyses. Association analyses identified loci associated with specific environmental factors and with quantitative traits. Candidate SNPs associated with the six SDM-identified environmental predictors contributing to habitat suitability were identified by RDA. The RDA found only 2% of the SNP variation to be associated with the six environmental predictors. This is a small value but not unexpected

because most of the SNPs will be neutral and therefore not show a relationship with the environmental predictors. Loadings near 0, at the centre of the distribution, show that the SNPs do not have relationship with the combined variation of the environmental predictors. SNPs that do show association with the environmental predictors are likely to be under selection. This selection can be in response to these selected predictors that were used in the model or some other environmental variable that is correlated with these predictors.

Genetic differentiation (F_{ST}) test was carried out for candidate SNPs associated with environmental predictors. High F_{ST} values were obtained in comparisons between gradients, similar to the highest F_{ST} results shown in **Figure 3.5**. The same candidate SNPs also gave strong signals in the XP-EHH analysis, both between agroecologies and between gradients. The results from the association analyses were supported by the high signatures of selection estimates based on F_{ST} and XP-EHH.

One of the reasons to undertake this research is understanding the genetics of phenotypic variation in indigenous chickens and their local adaptation in response to environmental variation, and to help improve productivity. Growth rates in indigenous chickens are low, making it difficult for smallholder farmers to achieve economic gains. Higher production efficiency can be achieved by developing breeds with wider environmental adaptation and better performance. With RDA, we identified 83 candidate SNPs in regions on chromosomes 1,3, 4, 7,8, 13, 15, and 19 that have a moderate to high correlation (0.3 < r < 0.6) with mature live body weight. Conventional GWAS studies in the past identified body weight associated SNPs and QTLs on chromosomes 1,4, 8, 11, 19 in Chinese, Rwandan, and Ethiopian chicken breeds (Cha et al., 2021; Habimana et al., 2021; Liu et al., 2013; Psifidi et al., 2016; Xie et al., 2012). Our results demonstrate that RDA can be used as an alternative approach to GWAS in random mating, indigenous livestock populations which have sufficiently interacted with the environment.

The presence of distinctive peaks in the Manhattan plot based on genotype-phenotype association analysis (**Figure 3.16**) suggests that phenotypic variation is present for selection to act on. The environmental drivers could increase haplotypes related to adaptive phenotypic plasticity or morphological variation in indigenous chickens. However, candidate SNPs associated with environmental predictors (**Figure 3.14**) were evenly spread across the genome without obvious overlap with the peaks from genotype-phenotype association. While the genotype-phenotype associations showed very distinct peaks, the total amount of SNP variation associated with phenotypic variation was only 0.2%, in contrast with 2% of the SNP variation associated with environmental variation. The underlying mechanisms of

genotype-phenotype associations are well studied and understood in livestock, but this is not the case for genotype-environment associations. Finding 2% of SNP variation related to environment variation is promising for further investigation of the mechanisms leading to these associations.

3.5. Conclusion

We have integrated environmental, genetic, and phenotypic sources of information to understand the genetic basis of adaptive phenotypic and genetic variation. Ethiopian chicken populations were found to have differentiated the most between gradients but selection pressures leading to adaptive variation have been stronger between agroecologies. Higher genetic divergence between gradients suggests that evolutionary processes other than natural selection, such as gene flow and drift, have influenced sample populations in different geographies. Environmental and phenotypic predictors are useful to explain genomic variation in Ethiopian indigenous chickens. The results from RDA were supported by the outputs from signatures of selection analyses (F_{ST} and XP-EHH). Signatures of selection analysis with the two methods can be used complementarily with RDA to shed light on the relationship between genomic, phenotypic, and environmental variation in local adaptation studies in indigenous chickens. In conclusion, our landscape genomic analyses provided compelling evidence for the presence of adaptive phenotypic and genetic variation among Ethiopian indigenous chickens.

3.6 Data availability statement

All supplementary materials referenced in this thesis are available on the Zenodo database (10.5281/zenodo.6583346)

4

Smallholder chicken production agroecologies redefined by distribution models improve model fit in prediction of breed by environment interactions

Fasil Getachew Kebede[1,2], Hans Komen[1], Tadelle Dessie[2], Olivier Hanotte[2,3], Steve Kemp[2], Setegn Worku Alemu[2,4], John W. M. Bastiaansen[1]

[1]Wageningen University & Research, Animal Breeding and Genomics, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands; [2]International Livestock Research Institute (ILRI), P. O. Box 5689, Addis Ababa, Ethiopia; [3]Cells, Organism and Molecular Genetics, School of Life Sciences, University of Nottingham, NG7 2UH, Nottingham, UK; [4]AL Rae Centre for Genetics and Breeding, School of Agriculture, Massey University, Palmerston North, New Zealand.

Abstract

Background: Animal performance is an outcome of genetic effects, environmental influences, and their interaction. Understanding the influences of the environment on performance is important to identify the right breeds for a given environment. Agroecological zonation is commonly used to classify environments and compare performance of breeds before their wider introduction into a new environment. Environmental classes, also referred to as agroecologies, are traditionally defined based on agronomically important environmental predictors. We hypothesised that own classification of agroecologies for livestock at a species level may improve estimations of genotype by environment interactions (GxE). We collected growth performance data on improved chicken breeds distributed to multiple environments in Ethiopia. We applied species distribution models (SDMs) through MaxEnt algorithm to select the most important environmental predictors associated with habitat suitability and phenotypic plasticity of chickens. We then grouped the environments of the performance testing sites into distinct agroecologies based on the selected environmental predictors. Finally, we compared the live body weight predictions of the breeds based on conventional (crop-based) and SDM-defined agroecologies using different models.

Results: Our study shows that predictive ability of GxE models is higher for SDM-defined agroecologies. The study also demonstrates that generalized additive models (GAMs) result in a better model fit when estimating GxE and predicting performance.

Conclusions: Our findings suggest that multi-environment performance evaluations of candidate breeds should be based on agroecologies defined for livestock. Agroecologies defined for a livestock species by SDMs result in better estimation of GxE and lead to a more objective comparison of breed performance. Moreover, GAMs are well-suited to integrating biological (breed and trait) and environmental information in breed performance comparison programmes.

4.1 Introduction

Scavenging family poultry significantly contributes to household nutrition and income in Sub-Saharan Africa. Levels of productivity in these systems are low and can be raised by introducing genetically improved chickens developed elsewhere (Birhanu et al., 2021).

A proper classification of agroecologies considering ecological and biological factors is essential to select productive breeds in specific environments (Dumont et al., 2013). Stratification of environments reduces genotype by environment interactions (Eberhart and Russell, 1966; Soussana et al., 2015). Smallholder environments are traditionally classified into agroecologies based on environmental predictors of agronomic importance such as duration of plant growing period, cropping pattern, and elevations (Dove, 1890; MoA, 1998; Tadesse Mulugeta, 2006). Agronomically defined agroecologies are made for crops and do not adequately consider environmental predictors which have profound association with livestock productivity and local adaptation (Dumont et al., 2014). Definition of agroecologies of farm animals as a framework to evaluate the performance of introduced breeds has received little research attention. Only 5 percent of the indexed studies concerning agroecology deal with livestock (Soussana et al., 2015).

SDMs (also called niche, envelope, or bioclimatic models) associate georeferenced observations of a biotic response variable – typically species occurrence or abundance – with multiple environmental predictors. For this, SDMs use several statistical learning methods to describe a species' niche. These niches describe the environmental conditions that are needed for a species to thrive and are used, for instance, to inform conservation planning (Elith and Franklin, 2013; Elith and Leathwick, 2009; Franklin, 2010; MacKenzie et al., 2017).

SDMs were used in the past to detect environmental factors associated with local adaptation among indigenous chickens (Gheyas et al., 2021; Kebede et al., 2021). Lozano-Jaramillo et al. (2019b) applied SDMs on improved chicken breeds to identify environmental parameters influencing suitability of habitat, to delineate potential habitat range of the breeds for optimum performance, and to rank their suitability for introduction into different administrative regions of Ethiopia. These SDM studies on chickens were implemented using MaxEnt, a machine-learning algorithm developed to model species distributions from presence-only records (Phillips et al., 2006).

SDMs consider a species as a homogenous unit and are constrained to meet some of their underlying assumptions when they are used alone in studying populations or breeds (Hampe, 2004; Wiens et al., 2009). They can be integrated with other modelling approaches such as phenotypic distribution models (PDMs) to predict the

performance of different breeds (Kebede et al., 2021; Lozano-Jaramillo et al., 2019a) and detect population differentiation in response to environmental selective pressures (Kebede et al., 2021). Lozano-Jaramillo et al. (2019a) performed a breed-wise comparison of performance of the introduced chicken breeds based on the existing administrative (political) boundaries of Ethiopia. Predictive ability of the PDMs can be improved if the most important environmental predictors influencing habitat suitability of chickens are used to define agroecologies. More importantly, PDMs such as generalized additive models (GAMs) can be used to evaluate the response of phenotypes to specific environmental predictors.

Generalized additive models (Hastie and Tibshirani, 1986, 1990) combine different types of fixed, random and smooth terms in the linear predictor of a regression model to account for different types of effects. They are useful to detect non-linear relationships, which is a common feature of many ecological datasets (Kebede et al., 2021; Wood, 2017; Zuur et al., 2007).

The objectives of the present study are to 1) apply species distribution models (SDMs) to identify the most important environmental predictors related with habitat suitability of chickens; 2) use SDM-identified environmental predictors to classify environments of the chicken performance testing sites into distinct agroecologies; 3) quantify the contribution of SDM-identified environmental predictors to phenotypic variability of growth traits; and 4) integrate information on agroecology, breed, and environmental predictors through phenotypic distribution models to evaluate growth performance and GxE.

4.2 Methods

4.2.1 Experimental design
Five improved chicken breeds were distributed to 2547 smallholder households across diverse environments and geographies of Ethiopia with the aim of evaluating live body weight at different ages. Performance testing sites were spread over twenty-one districts in five administrative regional states across the country (Oromia, Amhara, Addis Ababa, Southern Nationalities and Peoples' Region, and Tigray). Districts within a region were chosen through cluster sampling, such that they evenly represent three elevational gradients (400-1800; 1800-2400; 2400-3500 m.a.s.l.) as indicators for climatic variation. The selection of performance testing sites within districts, and households within a performance testing site (village), was carried out randomly. The target number of households was 20 for each performance testing site, with 2-3 sites selected per district. Households with less than two years of chicken keeping experience were excluded from the experiment.

In total, 25 unsexed chicks, vaccinated against major diseases (Newcastle disease, Gumboro or infectious bursal disease, fowl pox, and Marek's) and brooded to the end of 42 day-of-age, were distributed to each of the 2547 households. See **Table 4.1** for a summary and **Supplementary Table 4.1** for detailed information on each testing site.

Table 4.1. Summary of sampling sites used for evaluation of growth performance of improved chickens

Geographic region	Number of districts (n=21)	Number of performance testing sites	Number of households receiving a flock of chicken (25 unsexed chicks per breed)
Oromia	5	7	268
Amhara	5	14	774
Tigray	4	7	444
SNNPR	5	13	718
Addis Ababa	2	4	343
Total	21	45	2547

4.2.1.1 Environmental data

All households in this study were georeferenced to 1034 unique coordinates. Values for 34 environmental predictors which were expected to influence chicken productive performance were extracted from online databases (**Supplementary Table 4.2**). These included climatic (*n*=24), soil (*n*=8), and vegetation (*n*=2) predictors. Values for climatic predictors (*n* = 24), related with temperature, precipitation, solar radiation, and water vapour pressure) in different seasons were obtained from WorldClim database (http://www.worldclim.org/; version 2) at a spatial resolution of 30 seconds (~1Km^2) (Fick and Hijmans, 2017) based on the mean values of 30 years (1970-2000). Elevation data was obtained from DIVA-GIS (http://www.diva-gis.org/gdata) (Farr et al., 2007; Hijmans et al., 2001) at a spatial resolution of 30 seconds (~1Km^2). Vegetation data (cropland extent) was obtained from Global Food Security Analysis-Support Data (Xiong et al., 2017) at a spatial resolution of 30 meters. Soil data was obtained from ISRIC database (Hengl et al., 2017; Hengl et al., 2015) based on observations and measurements of African SoilGrids system at 250-meter resolution, with standard numeric soil properties (organic carbon, bulk density, Cation Exchange Capacity (CEC), pH, and soil texture fractions at 15-30 cm depth). Different R software packages: '*sp*' (Pebesma et al., 2012), '*raster*' (Hijmans et al., 2015)*, 'rgdal'* (Bivand et al., 2021a)*, 'maptools'* (Bivand et al., 2021b)*, 'rgeos'*(Bivand et al., 2017) *,* and '*dismo*' (Hijmans et al., 2017) were used to extract, read, and visualize geospatial data.

4.2.1.2 Breeds

The five breeds tested in this study were Improved Horro, Potchefstroom Koekoek, Kuroiler, Sasso, and SRIR (Sasso x Rhode Island Red). Improved Horro is the outcome of a selective breeding program on the local Horro chicken and developed with the aim of improving age at first egg, egg production, body weight and survival (Dana et al., 2010; Dana et al., 2011; Esatu, 2015). The Potchefstroom Koekoek is a composite of the White Leghorn, Black Australorp, and the Barred Plymouth Rock, developed during the 1950s in the Republic of South Africa (Fourie and Grobbelaar, 2003). The Koekoek is very popular among rural farmers in South Africa and neighbouring countries for egg and meat production as well as their ability to hatch their own offspring in medium input production systems (Grobbelaar, 2008)(Grobbelaar et al., 2010). The Kuroiler is a hybrid chicken widely believed to originate from crossing the Rhode Island Red, the White Leghorn, the Barred Plymouth Rock and two Indian indigenous chicken breeds with some introgression of broilers to obtain specific broiler characteristics (Ahuja et al., 2008a, b; Isenberg, 2007). The Kuroiler was developed in India by a commercial firm and was introduced to the market in early 1990s. The Sasso is a dual-purpose commercial hybrid developed by a breeding company in Europe. The SRIR is a hybrid closely related with the Sasso, with some genetic introgression from another dual-purpose genotype.

4.2.1.3 Management of chicks and phenotypic data collection

Farmers participating in the chicken performance evaluation were closely monitored for their adherence to the research protocols. All of them received trainings on basic chicken husbandry practices. Trained enumerators regularly visited the households to ensure they had constructed night shelters for the birds and provided water, and up to 30% of their daily feed requirement as a supplement in addition to scavenging. Every bird was individually identified with a wing-tag and its body weight was measured every two weeks by the enumerators. The average weight at a specific age was interpolated by linear regression from available data points. Live-body-weight-at-90, 120, 150, and 180-days-of-age were considered as traits: (LBW90, LBW120, LBW150, LBW180). In total, performance data was measured on 21,562 female chickens obtained from 2547 households. Out of 25 unsexed brooded chicks of a breed received by a household, roughly 50% were females, and an average of 8 female birds survived per household for phenotypic measurement at the end of the experiment.

4.2.1.4 Species distribution models (SDMs)
Selection of environmental predictors

Complex niche models show low performance in identifying important environmental predictors associated with habitat suitability (Warren and Seifert, 2011; Warren et al., 2014). We performed initial exploration of all geospatial data, independent of the breeds distributed, with Principal Component Analysis (PCA). The R package '*stats*' was used to run PCA and visualize the combined contributions of the predictors to the environmental variance. The relative magnitude and direction of the PCs indicate their usefulness for classification of environments into agroecologies. A final selection of the highest contributing set of uncorrelated environmental predictors to habitat suitability used in SDMs was done by the R package '*MaxentVariableSelection*' (Jueterbock et al., 2016).

Classifying the environments of chicken performance testing sites into agroecologies

Agroecologies in the context of the present study are groupings of locations where environmental conditions are relatively homogenous and within which chickens are expected to be similarly affected. We used *I* statistics to cluster chicken performance testing sites into distinct agroecologies by calculating the difference in suitability score between populations. First, raster files of the highest contributing set of predictors were created and used with ENMTools to calculate similarity statistics (*I*) (Warren et al., 2010)(Phillips and Dudík, 2008). A raster is an imagery from satellite which consists of a matrix of cells (or pixels) organized into rows and columns (or a grid) where each cell contains a value representing information for an environmental variable (Hijmans et al., 2015). Then, at each grid cell, estimates of habitat suitability are calculated using MaxEnt-generated species distribution models so that they sum to 1 over the geographic space being measured. The "*I*" value is an estimate of the probability that the relative ranking of any two patches of habitat is the same for two models (fit for niches occupied by two breeds), irrespective of the quantitative difference in suitability estimates. The '*I*' similarity measure ranges from 0, when species predicted environmental tolerances do not overlap at all (i.e., 0 for all *i*), to 1, when all grid cells are estimated to be equally suitable for both species (i.e., 1 for all *i*). Suitability of 0 is rarely produced by MaxEnt for most real data sets unless a minimum suitability threshold is applied, as MaxEnt SDMs predict a nonzero suitability score at every grid cell. However, values arbitrarily close to 0 are possible. An *I* score of approximately 0.5 for comparisons of two models mean there is a wide band of habitat that is estimated to be of an intermediate suitability for both of them (Warren et al., 2011). Agroecologies classified using SDM are compared with the

conventional AEZ (Dove, 1890; MoA, 1998) in the present study to see their respective utilities in the analysis of genotype by environment interactions.

4.2.1.5 Phenotypic distribution models (PDMs)

PDMs are class of models which describe the response of phenotypic variables (i.e. traits) as a function of environmental parameters (Lozano-Jaramillo et al., 2019a; Michel et al., 2017; Smith et al., 2017). Once the most important environmental parameters (predictors) associated with habitat suitability in chickens have been identified, phenotypic distribution models (PDMs) can be used to estimate their relative contribution to phenotypic variation and evaluate the response of selected quantitative traits to these predictors.

Correlations between repeat records such as body weight at different ages in chickens are expected to be high. A Pearson's correlation analysis was therefore first performed to keep the least related out of the four quantitative traits measured in the present study: live-body-weight-at-90, 120, 150, and 180-days-of-age (LBW90, LBW120, LBW150, LBW180).

Next, we used model based boosting to fit a statistical model while performing parameter selection at the same time (Thomas et al., 2017). We used the R (Team, 2013) packages '*mboost*' (Hothorn et al., 2010, 2012) and '*gam*' to fit boosted generalized linear models (GLMs), (Nelder and Wedderburn, 1972) and boosted generalized additive models (GAMs)(Chambers and Hastie, 1992; Hastie and Tibshirani, 1990; Hothorn et al., 2010, 2012), respectively.

Boosted GLMs

Component-wise gradient boosting aims to model a relationship between y and x : $= (x_1, \dots x_p)$, and   obtain the "optimal" prediction of y given x. A principal difference between the boosted generalized linear model and generalized linear models is that the former can additionally perform variable selection. A boosting algorithm also allows stopping at the optimal iteration to increase the predictive power of GLMs (and GAMs); (Hothorn et al., 2010, 2012) during variable selection. A generalized linear model of the predictors $x = (x_1, \dots x_p)$ has the form

$$g(y) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$$

Where y is the response variable (LBW90, LBW120, LBW180); g is the link function; β_0 is the intercept; and the $x_1, \dots x_p$ are the environmental predictors that are identified by SDMs.

Boosted Generalized Additive Models (Boosted GAMs)
Boosted GAMs were used to estimate the relative contribution of environmental covariates to phenotypic variation in quantitative traits. A GAMs model with the covariates $x = (x_1, \dots x_p)$ has, the form

$$g(E(y)) = \beta_0 + f_1 + \cdots + f_p$$

Where y is the expected response of growth trait (LBW90, LBW120, LBW180); g is the link function; β_0 is the intercept; and the $f_1, \dots f_p$ are the arbitrary function of environmental predictors identified by SDMs. These functions include simple, linear functions as well as smooth, non-linear functions.

Prediction of phenotype
We fitted two linear fixed-effects models and one generalized additive model to predict phenotypic value for each of the three growth traits (LBW90, LBW120 and LBW180).

Linear mixed-effects model (LMEM)
We used the R package '*lmer*' (Bates et al., 2007) to fit two linear mixed-effects models for live body weight at ages LBW90, LBW120, and LBW180 as follows:

$$y_{ijk} = \mu + \beta_i + \gamma_j + \beta\gamma_{ij} + \tau_k + e_{ijk}$$

Where y_{ijk} is the response variable (i.e., average live body weight in grams); μ is the grand mean, β_i is the fixed effect of the ith breed (i= 1, 2, 3, 4, 5); γ_j is the fixed effect of agroecology (j = 1, 2, 3); $\beta\gamma_{ij}$ is the interaction effect of the ith breed with the jth agroecology; τ_k is the random effect of SDM identified environmental predictors (k = 1, 2, ..6); and e_{ijk} is the residual error term. γ_j was either based on conventional agroecological classification with three levels: lowland, midaltitude and highland, or on SDM defined agroecology. The shape of X_k is fully determined by the data structure.
Least square means for the three different traits LBW90, LBW120, and LBW180 were estimated by the R package '*Emmeans*' (Lenth et al., 2018).

Generalized additive models (GAMs)
Phenotypes are expected to include non-linear responses to environmental predictors (Bolker et al., 2013; Oddi et al., 2019; Zuur et al., 2007) violating

assumptions made by classical linear approaches (e.g., linear regression or ANOVA). Generalized additive models (GAMs) can accommodate this non-linear relationship between the response and continuous explanatory variables by making use of non-parametric smoothers (Crawley, 2012; Wiley and Wiley, 2019). The R package '*mgcv*' (Wood and Augustin, 2002) was used to fit GAMs (Hastie and Tibshirani, 1990). The function '*bam*' was used, that is specifically suitable for large data sets, and we invoked a Gaussian smoothing process (bs= '*gp*' with an identity link function in the 'gam' formula.

The interaction effect of agroecology and breed (GxE) was investigated by fitting GAMs with environmental predictors as smoothing parameters, as follows:

$$g(E(y)) = \alpha + \beta_i + \gamma_j + \beta\gamma_{ij} + f_k(X_k),$$

Where $(E(y))$ is the response variable; g is the Gaussian distributed exponential family with identity link function; α is the intercept; β_i is the fixed effect the ith breed (i= 1, 2, 3, 4, 5); γ_j is the fixed effect of agroecology (j = 1, 2, 3); $\beta\gamma_{ij}$ is the linear parameter for the interaction between ith breed and jth agroecology; and f_k are the smoothing term(s) or random effects of selected non-parametric environmental predictor covariate(s) X_k.

4.3 Results

4.3.1 Contribution of environmental parameters to habitat suitability

MaxentVariableSelection retained a final set of six least correlated (|r| < 0.6) environmental predictors related with habitat suitability of chickens (**Figure 4.1**). These included elevation, solar radiation in May, precipitation of the driest month (Bio14), water vapour pressure of May, precipitation of the coldest quarter (Bio19), and precipitation of the wettest month (Bio13). Each of the selected predictors had contributions to habitat suitability of chickens of more than 3%. Elevation has the highest contribution (48%) to habitat suitability. Soil and vegetation types were excluded since they either were highly correlated with the selected predictors or contributed insignificantly to habitat suitability.
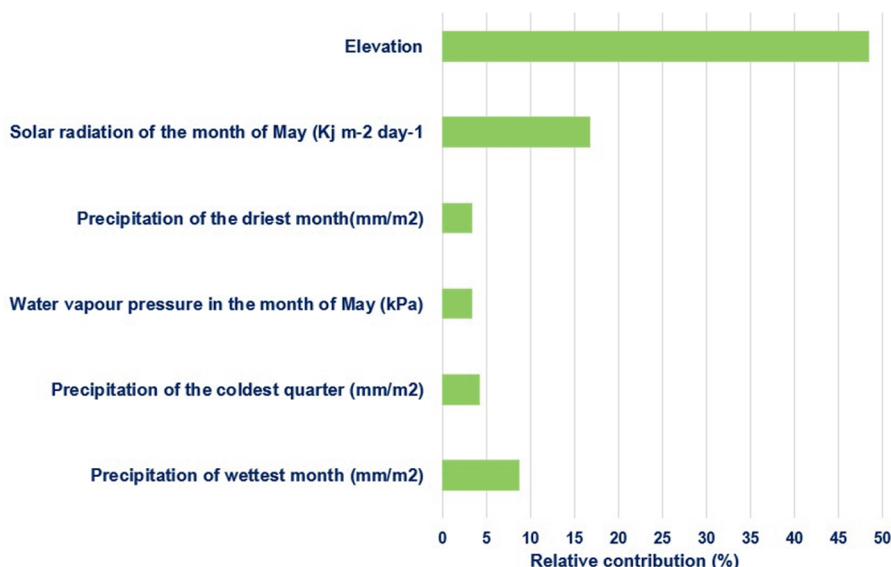
Figure 4. 1. Relative contribution of the six most important environmental predictors selected to habitat suitability

The PCA plot showing the contribution of environmental predictors in the habitat/niche structure of the chicken breeds is presented in **Supplementary Figure 4.1.** Breeds placed close to each other in the plot are similarly affected by those environmental predictors. Altogether, PC1 and PC2 explained 71.3 percent of the variation. Three of the predictors identified by MaxEnt (elevation, solar radiation of May, and water vapour pressure of May) were also identified by the PCA analysis for their higher contributions**.**

Each of the six most contributing climatic predictors selected by PCA and *MaxentVariableSelection* were used to produce environmental maps of Ethiopia (**Figure 4.2**). The maps show the environmental heterogeneity of the country in terms of precipitation, solar radiation, water vapour pressure and elevation, indicating that these predictors deserve consideration in the classification of agroecologies.
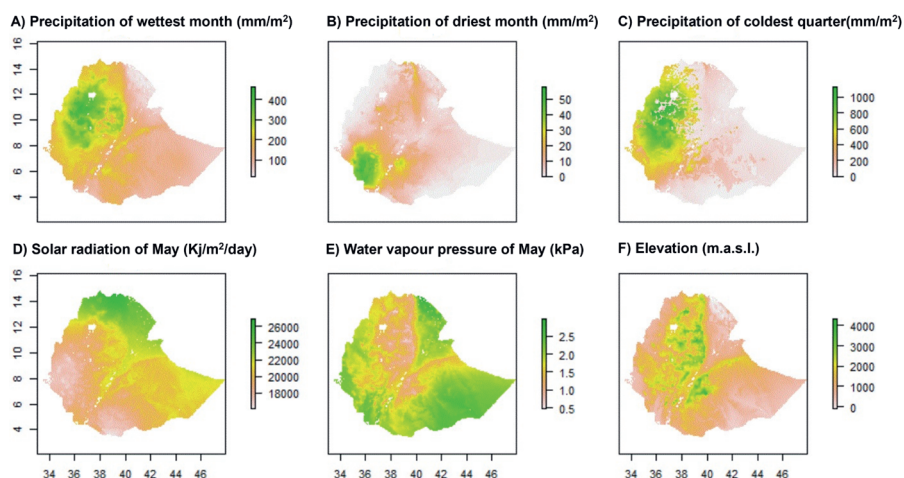
Figure 4.2. Environmental maps of Ethiopia based on the six most important environmental predictors

4.3.2 Agroecologies defined from the most contributing environmental parameters

These six most important environmental predictors associated with habitat suitability of chickens were used to classify the performance testing sites into three distinct agroecologies based on niche overlap statistic (*I*) (**Figure 4.3**).
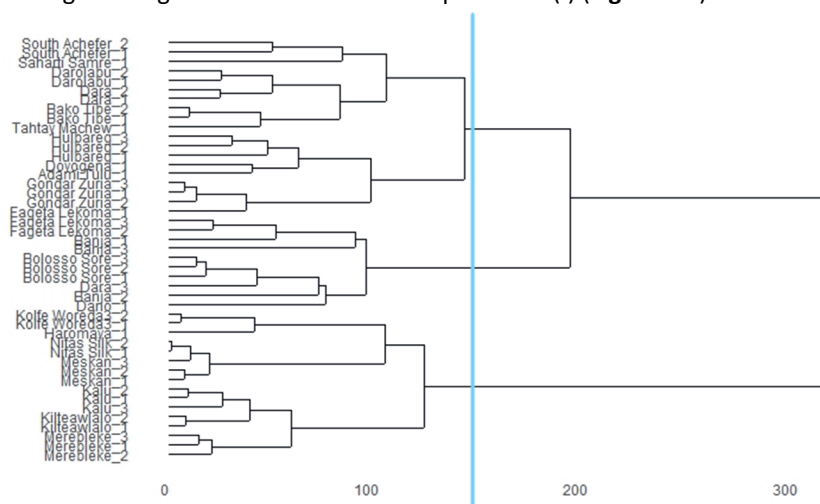
Figure 4.3. Dendrogram of niche overlap statistic (*I*) for testing sites of introduced chicken breeds in Ethiopia. The vertical line indicates the cut-off point to group environments of performance testing sites into three agroecologies.

Out of the forty-five villages where the improved breeds were distributed into, 16 villages were assigned into agroecology-I, 10 were assigned into agroecology-II, and 19 were assigned into agroecology-III (full list in **Supplementary Table 4.3**). Chicken at testing sites within the same agroecology are expected to be affected similarly by the environment.

For each of the three agroecologies, a habitat suitability map based on niche overlap was produced by SDMs (**Figure 4.4**). Most of the suitable areas on the three maps are in the northern and central areas of Ethiopia. Areas far in the west, south, and north-eastern parts of the country have been shown as the least suitable for chicken production. These are characterized by extreme temperatures and high solar radiation, and low availability of scavenging feed resources for chicken (Bayou and Assefa, 1989; CSA, 2017; Gebrechorkos et al., 2019; Getahun, 1978; Mirkena et al., 2018).
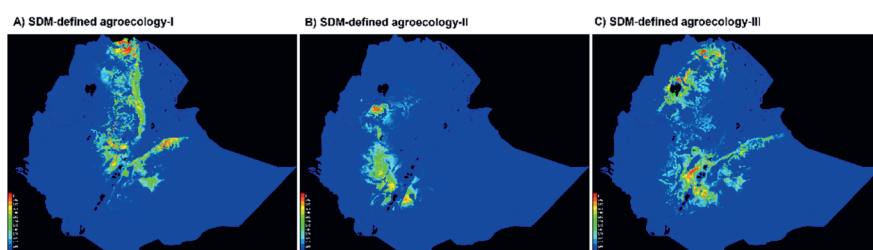
Figure 4.4. Suitability maps of SDM-defined agroecologies for chicken performance testing sites

Colours towards red spectrum in each of the three SDM-defined agroecologies (I, II, & III) show better habitat suitability to chickens

4.3.3 Environmental parameters differ in their contribution to phenotypic variation

Out of the four body weights measured in the present study, LBW150 was highly correlated (r>0.8) with both LBW120 and LBW180 and removed from the analysis (**Supplementary Table 4.4**). A phenotypic correlation ranging between 0.3 to 0.9 was obtained for live body weight of broilers between different ages (weeks 1 through 6) raised in commercial environment (Chu et al., 2020).

The relative contributions of the six most important environmental predictors to female live body weight at the three different ages (LBW90, LBW120, LBW180) were predicted by two types of phenotypic distribution models, namely, boosted GLMs and boosted GAMs (**Table 4.2**).

Table 4.2. Relative contribution (%) of the six most important environmental predictors to female body weight

Trait	Relative contribution (%) estimated by boosted GLMs						Boosted GAMs		
	EL	SRM	Bio13	Bio19	WVP	Bio14	df	R-sq. (adj.)	Deviance (%)
LBW90	0.49 **	0.05 **	0.0 **	0.03	0.39 **	0.04	1441	0.36	37.2
LBW120	0.48 **	0.03 **	0.0 **	0.04 **	0.40 **	0.05 **	1395	0.43	44.8
LBW180	0.47 **	0.04 **	0.0 **	0.05 *	0.41 **	0.03 **	1375	0.48	49.3

Approximate significance of smooth terms estimated by GAMs (*not* by boosted GAMs): 0.001 '**', 0.01 '*'. EL=elevation; SRM=solar Radiation of May; BIO13= precipitation of the wettest month; BI019= precipitation of the coldest quarter; WVP=water vapour pressure of May; BIO14= precipitation of the driest month.

Four of the six environmental predictors identified by MaxEnt had also a significant effect on the prediction of body weight. Elevation, solar radiation of May, precipitation of the wettest month (Bio13) and water vapour pressure of May had significant but varying contributions to phenotypic differentiation on all the traits. Elevation and water vapour pressure in May were most useful for predicting female live body weight and were highly significant (p <0.001) at the three different ages (at 90, 120, and 180-days-of-age). Precipitation of the coldest quarter (Bio19) and precipitation of the driest month (Bio14) had no significant contribution to LBW90. The selected environmental predictors together explained only about 25 to 40 percent of the variation in phenotype, the rest of the deviance being attributed to other unknown factors.

4.3.4 Prediction of phenotype with additive models gives a better fit

Comparison of model efficiency of linear mixed-effects models (LMEM) and General Additive Models (GAMs) in predicting live body weight at 90-days-of-age (LBW90) and GxE is presented in **Table 4.3**. Conventionally defined agroecology, fitted as fixed effect, had no significant effect (*p* =0.84) on LBW90 when fitted with LMEM. AIC and adjusted R-square values show GAMs had the highest model fit in predicting LBW90. GAMs (M3) explained 36 percent of the variation in this trait while the other two models (M1 and M2) explained 19 percent. The interaction between breed and agroecology was highly significant (*p* < 0.001) in all the three models (M1,M2, and M3).

Table 4.3. Model comparison for prediction of live body weight at 90-days-of-age and GxE in female chickens

AE Class	Model	Linear predictor			[1]Random effects	Model fit		
		AEs (n=3)	Breed (n=5)	AE x Breed (GxE)	Climatic Predictors	df	*AIC	R-sq. (adj.)
[§] Conventional	LMM (M1)	p=0.84	***	***	BIO13, BIO14, BIO19, SRM, WVP, EL	1346	96	0.19
SDM-based	LMM (M2)	***	***	**	BIO13, BIO14, BIO19, SRM, WVP, EL	1377	92	0.19
	GAM (M3)	***	***	***	BIO13, BIO14, BIO19, SRM, WVP, EL	1441	0	0.36

[1]Akaike's information criterion (AIC) and likelihood value AIC were set to zero as reference for the best model; AIC = 2× # parameters – 2 × log-likelihood; thus lower values indicate a better model. Significant codes: '***' 0.001; '**' 0.01
[1]BIO13= precipitation of the wettest month; BIO14= precipitation of the driest month; BIO19= precipitation of the coldest quarter; SRM=solar Radiation of the month of May; WVP=water vapour pressure of the month of May; EL=elevation. M1=Model 1; M2=Model 2; M3=Model 3. R-square: is adjusted for linear mixed-effects models (M1 and M2); conditional R-square calculated for M3.[1]The six environmental smoothers had significant effect on LBW180 (p<0.001). [§]Conventional or traditional AEZ with three classes: I=lowlands (400-1800m.a.s.l.); II=1800-2400m.s.a.l.; III=2400-3500m.a.s.l. (Dove, 1890) (MoA, 1998).

The effect of conventionally defined agroecology (AE) was insignificant (*p* = 0.65) on LBW120 in the model fitted by LMEM in **Table 4.4**. The model fit statistics (AIC and adjusted R-square values) show superiority of the GAMs compared to LMEM fitted with SDM-defined AEs in predicting LBW120. GAMs (M6) explained 43 per cent of the variation in LBW120. Breed and agroecology had highly significant interactions (p < 0.001).

Table 4.4. Model comparison for prediction of live body weight at 120-days-of-age and GxE in female chickens

AE Class	Model	Linear predictor			Random effects	Model fit		
		AEs (n=3)	Breed (n=5)	AE x Breed (GxE)	Climatic predictors	df	* AIC	R-sq. (adj.)
§ Conventional	LMM (M4)	p=0.65	***	***	[1]BIO1, BIO14, BIO19, SRM, WVP, EL	1275	28	0.16
SDM-based	LMM (M5)	***	***	***	[1]BIO13, BIO 14, BIO19, SRM, WVP, EL	1286	35	0.20
	GAM (M6)	***	***	***	[1]BIO13, BIO14, BIO19, SRM, WVP, EL	1395	30	0.43

[1]Akaike's information criterion (AIC) and likelihood value AIC were set to zero as reference for the best model; AIC = 2× # parameters – 2 × log-likelihood; thus lower values indicate a better model. Significant codes: 0.001; '***' [1]BIO13= precipitation of the wettest month; BIO14= precipitation of the driest month; BIO19= precipitation of the coldest quarter; SRM=solar Radiation of the month of May; WVP=water vapour pressure of the month of May; EL=elevation. M1=Model 1; M2=Model 2; M3=Model 3. R-square: is adjusted for models (M1 and M2); conditional R-square calculated for M3.[1]The six environmental smoothers had significant effect on LBW180 (p<0.001). §Conventional or traditional AEZ with three classes: I=lowlands (400-1800m.a.s.l.); II=1800-2400m.s.a.l.; III=2400-3500m.a.s.l. (Dove, 1890) (MoA, 1998).

We also performed predictions of performance and GxE for LBW180 using mixed models and GAMs (**Table 4.5**). All three models (M7, M8, M9) detected breed by environment interactions. Breed and agroecology had highly significant interactions ($p < 0.001$) in all three models. Based on the AIC values, LMEM fitted with the agroecologies AEI, AEII and AEIII defined by SDMs (M8) was better than LMEM fitted with conventionally defined AEs (M7). The best prediction of LBW180 was obtained by the model fitted by GAMs (M9). Conventionally defined AE had no significant effect (p = 0.2) on LBW180. The explained variance (deviance explained) with GAMs

was 48.0%. The corresponding R-square value for LMEM fitted with SDM-based AEs was only 24%.

Table 4.5. Model comparison for prediction of live body weight at 180-days-of-age and GxE in female chickens

AE Class	Model	Linear predictor			Random effects	Model fit		
		AEs (n=3)	Breed (n=5)	AE x Breed (GxE)	Climatic predictors	df	* AIC	R-sq. (adj.)
[§] **Conventional**	LMM (M7)	p=0.2	***	***	[1]BIO13, BIO14, BIO19, SRM, WVP, EL	1363	45	0.24
SDM-based	LMM (M8)	***	***	***	[1]BIO13, BIO14, BIO19, SRM, WVP, EL	1361	44	0.24
	GAM (M9)	***	***	***	[1]BIO13, BIO14, BIO19, SRM, WVP, EL	1375	0	0.48

[1]Akaike's information criterion (AIC) and likelihood value AIC were set to zero as reference for the best model; AIC = 2× # parameters – 2 × log-likelihood; thus lower values indicate a better model. Significant codes: '***' 0.001. [1]BIO13= precipitation of the wettest month; BIO14= precipitation of the driest month; BI019= precipitation of the coldest quarter; SRM=solar Radiation of the month of May; WVP=water vapour pressure of the month of May; EL=elevation. M1=Model 1; M2=Model 2; M3=Model 3. R-square: is adjusted for models (M1 and M2); conditional R-square calculated for M3.[1]The six environmental smoothers had significant effect on LBW180 (p<0.001). [§]Conventional or traditional AEZ with three classes: I=lowlands (400-1800m.a.s.l.); II=1800-2400m.s.a.l.; III=2400-3500m.a.s.l. (Dove, 1890) (MoA, 1998).

The effect of environmental covariates on live body weight of female improved chickens at 180-days-of-age (LBW180) estimated by GAMs are shown by partial dependence plots (PDPs) (**Figure 4.5**). GAMs revealed non-linear relationships

between environmental predictors and the response trait that would not have been explained by the linear fixed-effects models. The highest body weight is predicted for solar radiation between 22500 and 25000 (Kj/m^2/day), elevation below 2400 m.a.s.l., precipitation in the coldest quarter of about 600 mm/m^2, and water vapour pressure below 1.3 kPa in May. Precipitation of less than 240 mm/m^2 in the wettest month and precipitation of less that 5mm/m^2 in the driest month affected LBW180 negatively.
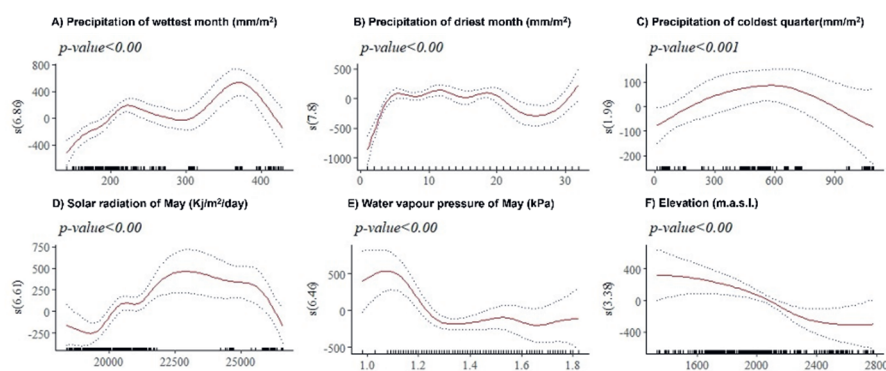
Figure 4.5. Generalized additive model partial dependence plots for live body weight in female chickens

Each plot shows a covariate and their partial dependence on LBW180 in the context of the model. The y axis shows the mean of observed change in live body weight and the x axis the covariate interval. The blue line represents the 95% confidence interval. Redline = mean of observed live body weight; blue line=standard error; *s* = smoothed variable; and *()* = effective degrees of freedom.

Genotype by agroecology interactions (GxE)

Figure 4.6 shows least square means of live body weight (*g*) with standard deviations of female chickens at different ages for improved breeds distributed into three different SDM-defined agroecologies in Ethiopia. The Kuroiler and the Sasso breeds had the highest live body weight (LBW90, LBW120, and LBW180) across all three agroecologies while the locally improved Horro had the smallest. Out of the five breeds, the Kuroiler and the SRIR showed an average performance across the three environments.

Colours indicate each of the five chicken breeds with their live body weight (along y-axis) at different ages a). LBW90; b) LBW120; c). LBW180. The three agroecologies

are displayed for each of the plots along the x-axis (AEI = agroecology-I; AEII = agroecology-II; AEIII = agroecology-III).
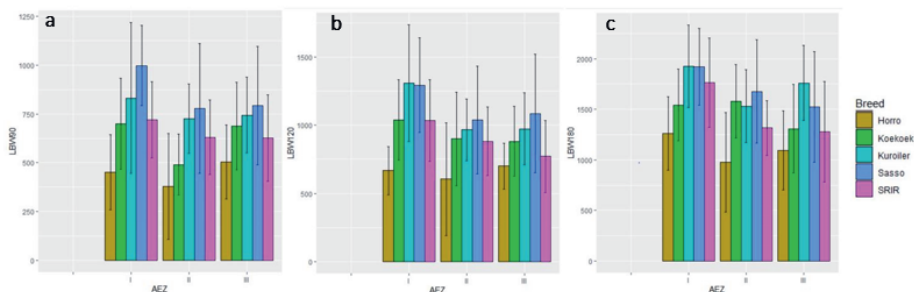
Figure 4.6. Bar plots depicting means live body weight with standard deviations of female chickens

Reaction norm plots of live body weight (*g*) at the three different ages (LBW90, LBW120, and LBW180) based on least square means show GxE interactions (**Figure 4.7**). Changes in rank or magnitude are observed across agroecologies for the five breeds. SRIR performed better than the Koekoek in AEI but not in AEII and AEIII. A change in rank was also noted at the same age between the Kuroiler and the Sasso between AEII and AEIII. Except for the Horro, which had higher live body weight in agroecology-III, for BW90 and LBW120, all the breeds performed the best in agroecology-I than in agroecology-II and III, suggesting its better suitability for the breeds. Thus, the results showed a clear genotype by environment interaction among breeds assigned to the three SDM-defined agroecologies.

Figure 4.7. Reaction norm plots for female live body weight at different ages

Colours indicate each of the five chicken breeds with their live body weight (along y-axis) at different ages a). LBW90; b) LBW120; c). LBW180. The three agroecologies are displayed for each of the plots along the x-axis (AEI = agroecology-I; AEII = agroecology-II; AEIII = agroecology-III).

4.4 Discussion

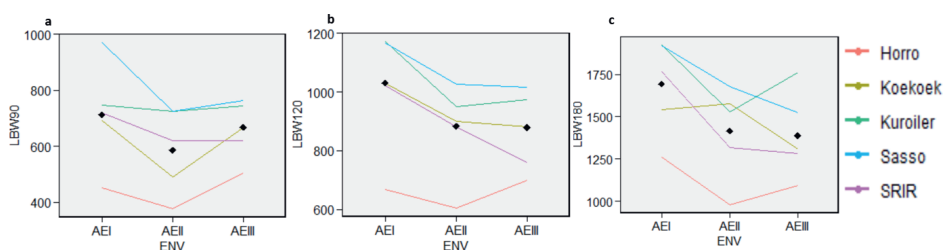An important reason for the failure of improved chicken introduction schemes designed in the past seventy years to improve smallholder poultry productivity in Africa is the lack of proper matching between environement and genotype. Input and service delivery systems, in terms of availability of vaccines, chicks (fertile eggs), market outlets, and feeds were also not so well developed in these areas to the sustain adoption of improved chicken breeds (Magothe et al., 2011; Safalaoh, 2001; Tadelle et al., 2000). The poultry production landscape has been positively changing in the last few decades to allow the introduction of improved dual-purpose chickens as a feasible genetic improvement strategy to raise productivity in a reasonable amount of time (Birhanu et al., 2021). Chicken breeds introduced into smallholder systems in the present study were managed under reasonably uniform conditions (i.e. received vaccines, supplementary feeds, and night shelters).

Multi-environment performance evaluation of candidate breeds developed elsewhere requires proper definition of livestock production agroecologies to improve GxE estimations. Knowledge on GxE helps to select breeds adapted to, and productive under specific or wider environmental conditions (Birhanu et al., 2021; de Kinderen et al., 2020; Lozano-Jaramillo et al., 2019a; Lozano-Jaramillo et al., 2019b). Existing agroecological zone definitions in Ethiopia (Deressa et al., 2010; Dove, 1890; MoA, 1998; Tadesse Mulugeta, 2006) are based on environmental predictors that are primarily of agronomic importance. Classifications primarily based on predictors such as cropping pattern, land use type, and elevation are not sufficient to compare performance of livestock breeds and recommend the best performing ones for different environments.

An analytical framework for classification of poultry production agroecologies and predicting GxE among improved chicken breeds distributed to smallholder farmers is introduced in the present study. We applied species distribution models (SDMs) to identify the most important environmental predictors and used the information on these predictors to group environments in chicken performance testing sites into distinct agroecologies. These SDM-classified agroecologies were then incorporated into phenotypic distribution Models (PDMs) for two purposes. First, Boosted GLMs and boosted GAMs were used to predict the relative contribution of selected environmental predictors to phenotypic variation of three live body weight traits (LBW90, LBW120 and LBW180). Secondly, the PDMs were used to predict phenotypic values and GxE for the three live body weight traits (LBW90, LBW120 and LBW180).

Results from Boosted GLMs and boosted GAMs have shown that elevation and water vapour pressure of the month of May had the highest contribution to phenotypic

differentiation of all three traits. This might be because of the correlation between elevation and other environmental predictors. A tight relation is expected between elevation and climatic elements (e.g. precipitation, temperature, radiation) (Dinka, 2019; Fazzini et al., 2015; Gamachu, 1988) in Ethiopia. Effects of elevation, solar radiation, and precipitation on performance of traits such as feed consumption, growth, meat and egg productivity and quality are well established (Howlider and Rose, 1987; Huang et al., 2017; Lara and Rostagno, 2013; Lin et al., 2006; Lozano-Jaramillo et al., 2019a; Marsden and Morris, 1987; Shane, 1988; Shlomo, 2000). Precipitation of the coldest quarter is a quarterly index which approximates the total precipitation that prevails during the three months of the year. Kebede et al. (2021) suggested precipitation of the coldest quarter above 700mm/m^2 might be related with less availability of scavenging feed resources and more prevalence of diseases and parasites, having adverse effects on mature body weight in indigenous chickens.

4.5 Conclusions

To the best of our knowledge this is the first attempt to classify livestock agroecologies based on environmental predictors identified by species distribution models (SDMs) and to compare performance on SDM-defined agroecologies. We have demonstrated that better defined agroecology, considering environmental predictors associated with habitat suitability for a species, improves the prediction of GxE in multi-environment livestock performance evaluations. The use of phenotypic distribution models, such as GAMs, is recommended as a powerful method to integrate biological (breed and trait) and environmental information in breed performance evaluation and comparison programmes.

Future research can investigate improvements in prediction of GxE for other important traits such as egg productivity, robustness, and yield stability for breeds introduced into smallholder production systems of diverse agroecologies. We anticipate accuracies of predictions by species and phenotypic distribution models will increase if the models incorporate additional information (e.g., genomic data).

4.6 Declarations

Ethics approval and consent to participate

The animal study was reviewed and approved by the Institutional Research Ethics Committee (IREC) of International Livestock Research Institute (ILRI).

Availability of data and materials

The datasets analysed during this study and supporting its conclusions are available in the data repositories of International Livestock Research Institute (ILRI) **https://data.ilri.org/portal/dataset/acggonfarmeth**

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

FGK, TD, JB, HK, SWA, and SK, OH conceived the ideas and designed the study. FGK trained enumerators, monitored data collection, performed data analysis, and drafted the manuscript. JB and SWA supported in the GAMs analysis. JB, HK, and SWA critically reviewed the manuscript, provided useful comments, and helped improve the data analysis and the draft manuscript. TD, SK and OH secured funding. All authors have read and approved the final manuscript for publication.

Author's information

[1]Wageningen University & Research, Animal Breeding and Genomics, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands; [2]International Livestock Research Institute (ILRI), P. O. Box 5689, Addis Ababa, Ethiopia.

5

Multi-environment performance analysis (MEPA) to evaluate productivity and yield stability in chicken

Fasil Getachew Kebede[1,2], Hans Komen[1], Tadelle Dessie[2], Olivier Hanotte[2,3], Steve Kemp[2], Setegn Worku Alemu[2,4], John W. M. Bastiaansen[1]

[1]Wageningen University & Research, Animal Breeding and Genomics, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands; [2]International Livestock Research Institute (ILRI), P. O. Box 5689, Addis Ababa, Ethiopia; [3]Cells, Organism and Molecular Genetics, School of Life Sciences, University of Nottingham, NG7 2UH, Nottingham, UK; [4]AL Rae Centre for Genetics and Breeding, School of Agriculture, Massey University, Palmerston North, New Zealand.

Abstract

Background: The performance of livestock breeds currently held in smallholder systems of the tropics is too low to the meet growing demands for animal source proteins. One breeding approach to enhance current levels of productivity and address food security is to evaluate candidate improved breeds developed for medium-input systems for their productivity and yield stability in smallholder systems and introduce them to farmers at scale. The biophysical environmental factors and management practices of smallholder livestock farmers are heterogenous and require experimental designs and analytical methods which can identify breeds that perform optimally in multiple environments. However, there is a lack of analytical frameworks for multi-environment livestock breed performance evaluation in smallholder (extensive) systems. Experimental designs and methods used in plant breeding to predict GxE and test yield stability are applicable to evaluate livestock breed performance in different agroecologies. We fitted two of such statistical models, namely, additive main effects multiplicative interaction model (AMMI) and linear mixed-effects models (LMM) on 5 improved chicken breeds to evaluate growth performance and stability until days 90, 120, and 180 (W90, W120, and W180). A total of 21, 562 animals were evaluated in 2547 smallholder households, distributed across 45 performance testing sites, in three agroecologies defined by Species Distribution Models (SDMs) in Ethiopia.

Results: Our results show that LMM had the best model fit in comparing breeds on productivity and yield stability. In both methods of MEPA applied in the present study, Sasso and Kuroiler were the most productive breeds for the three traits (W90, W120, and W180). Koekoek and SRIR had comparable but average growth performance. The ranking of breeds for yield stability varied according to the method of MEPA and stability indexes used. However, the most advanced indexes based on LMM revealed those breeds with the highest productive performance were the also most stable.

Conclusion: our findings show that agroecologies defined by Species Distribution Model (SDMs) are useful to undertake Multi-environment performance analysis (MEPA) in livestock. Our results also demonstrate that existing methods of MEPA are applicable to livestock breed performance comparisons. LMM-based productivity and stability indexes combine high performance with stability and are considered superior for MEPA of livestock in smallholder systems.

5.1 Introduction

Most of the attempts to genetically improve livestock productivity in sub-Saharan Africa did not bring the desired results. Interventions aiming at introducing high-producing breeds developed for intensive temperate livestock systems were not able to transform extensive tropical systems primarily because they did not match the right genetics with the right environment (Birhanu et al., 2021; Magothe et al., 2012; Safalaoh, 2001; Tadelle et al., 2000).

Sustainability of livestock genetic improvement schemes, particularly in smallholder systems, depends on reducing breed by environment interaction (GxE) and identifying productive breeds adapted to a wide range of environments and management practices. Evidence for differential performance of livestock breeds in response to environmental factors, such as extremes of temperature, solar radiation, relative humidity and wind speed have already been documented, in poultry (Kebede et al., 2021; Kebede et al., 2022; Lan et al., 2016; Lara and Rostagno, 2013; Lozano-Jaramillo et al., 2019a; Mazzi et al., 2003), cattle (Bagath et al., 2019; Brown-Brandl, 2013), swine (Mayorga et al., 2019; Ross et al., 2015).

Apart from their level of performance for the trait of interest, candidate breeds need to be evaluated on their yield stability before they are introduced at scale to smallholder farmers. The concept of 'yield stability' or 'stability' is closely related with terms that are more frequently used by breeders such as 'phenotypic plasticity', 'developmental plasticity' and genotype by environment interactions. The subtle differences among these terms need be acknowledged to exploit the genetic and phenotypic variability among livestock populations and to bring practical improvements in performance in smallholder systems.

Individuals in a population or populations of organisms are called 'phenotypically plastic' or 'plastic' when they show higher phentoypic variability (e.g., morphological, behavioural, physiological) across environments (agroecologies) for a trait compared to other individuals or populations. The term 'developmental plasticity' is alternatively used in some literature to describe the permanent behavioral, anatomical, or physiological changes in the developmental trajectory adopted by an organism during the life span, influenced by external environmental factors or other internal physiological factors, and that occurs through gene–environment interactions (Lafuente and Beldade, 2019). Developmental plasticity allows an organism to sense environmental cues in early stages of life and produce

phenotypes better adapted to environments encountered later in life (Xue and Leibler, 2018).

Plasticity is thought to be under genetic control (Diouf et al., 2020). Some genes control both plasticity and mean phenotype while many others are associated only with plasticity (Lafuente and Beldade, 2019; Sieriebriennikov et al., 2018). Stability is the ability of a genotype to be less 'sensitive' to environmental influences (Becker and Leon, 1988; Eberhart and Russell, 1966; Finlay and Wilkinson, 1963; Lin et al., 1986; Shukla, 1972). Stable genotypes or breeds are populations that show less plasticity across environments and are also referred to as 'robust' (De Jong and Bijma, 2002).

The outcome of phenotypic plasticity for a quantitative trait shows up during genotype by environment interaction (GxE) or breed by environment interaction analysis with the help of a reaction norm (Costa, 2021; De Jong and Bijma, 2002; Woltereck, 1909). GxE comprises the extent to which genotypes differ in their plastic responses to environmental changes for a given trait (Sultan, 2021). Several studies have applied reaction norms to describe the response of breeds to environmental variation (Chen et al., 2021; Cheruiyot et al., 2020; Oliveira et al., 2018).

Heterogenous environments (e.g., with diverse climate, management levels) require robust experimental designs and analytical methods which can identify breeds that perform optimally in multiple environments. Quantifying productivity and yield stability across environments (agroecologies) is taken as a strategy in the present study to select and recommend improved chicken breeds which are productive and widely adapted. Analytical frameworks are not readily available for smallholder livestock breed evaluations. On-farm experimental designs and statistical methods used in plant breeding trials to predict GxE and to test yield stability across environments could be adabted to evaluate breed performances across diverse agroecologies.

Livestock species specific definition of agroecologies based on species distribution models (SDMs) were proposed by Kebede et al. (2022). SDM-defined agroecologies are delineated based on environmental predictors influencing habitat and phenotypic differentiation of a livestock species and improve the accuracy of GxE predictions compared to conventional agroecologies classified on environmental predictors influencing plant growth (e.g., length of growing season). SDM-defined

agroecologies have not been tested for their utility in multi-environment performance analysis (MEPA) of livestock breeds.

Two methods of MEPA are commonly used in plant breeding. These include additive main effects and multiplicative interaction models (AMMI) and linear mixed-effects models (LMM). The pros and cons of using the different models in crop trials are well documented. When a standard Additive Main effects and Multiplicative Interaction (AMMI) (Gauch Jr, 2013) model provides more accurate estimates compared to traditional ANOVA (Van Eeuwijk et al., 2016). AMMI is a family of robust multi-environment analytical techniques which is widely used to study GxE interactions (Barhdadi and Dubé, 2010; Culman et al., 2009; Mukherjee et al., 2012; Rincent et al., 2019; Rodrigues et al., 2016). LMMs through REML/BLUP are predictively more accurate than AMMI (Piepho, 1994; Van Eeuwijk et al., 2016). Combining different analytical approaches helps identify stable and productive breeds across environments (Olivoto et al., 2019a; Olivoto et al., 2019b). The different MEPA tools also vary in terms of their graphical outputs which limits the applicability of some methods to prioritize breeds based on productivity and stability.

The relative efficiency of different MEPA models in analysing productivity levels and yield stability among livestock breeds has not been investigated. The objectives of the present study are to 1) evaluate two commonly used methods of MEPA for their applicability in livestock breed performance comparisons in smallholder systems based on agroecologies defined by species distribution models (SDMs); and 2) apply MEPA to compare productivity and yield stability of improved chicken breeds across SDM-defined agroecologies and recommend superior ones for wider use by smallholder farmers.

5.2 Material and methods

5.2.1 Phenotypic data
Five improved chicken breeds (Improved Horro or Horro, Kuroiler, Potchefstroom Koekoek or Koekoek, Sasso, and SRIR) were evaluated for their growth performance. Three traits were measured, namely, live-body-weight-at-90, 120, and 180-days (W90, W120, W180). Additional information on the breed of chickens and the dataset used in the present study is available in Kebede et al. (2022).
Every bird was individually identified with a wing-tag and its live body weight was measured every two weeks with digital weighing scale (10g accuracy). The scale was calibrated to "0.0" after hanging a plastic bucket and before each chicken was placed inside for measurement. Once all chickens in a household were measured, the

average weight of the flock at that specific datapoint was reported in near real time to a central computer server via ODK collect, an Android based mobile application (Hartung et al., 2010).

Chicks were vaccinated against major diseases (Newcastle disease, Gumboro, fowl typhoid, fowl pox, and Marek's) and brooded to the end of 42-days-of-age before they were distributed to households. Participant farmers received training on basic chicken husbandry practices. Enumerators monitored households to ensure they had constructed night shelters for their birds and provided 30% of their daily feed requirement as a supplement, on top of scavenging.

5.2.2 Environmental data

Individual households at performance testing sites were georeferenced and important environmental variables were extracted for their locations (**Supplementary Tables 5.1** and 5.**2**). The classification of chicken performance testing sites into three agroecologies (AEI, AEII, and AEIII) thorough Species Distribution Models (SDMs) by Kebede et al. (2022) was adopted in the present study.

5.2.3 Experimental design

The study was designed in a Randomized Complete Block. The chicken breeds (*n*=5) were assigned as treatments (**GEN**) into each of the three agroecologies (**ENV**). Performance testing sites within each ENV were considered as environmental replicates (**REP**) or blocks. The administrative concepts of village and districts were not relevant in the design of the present study and were ignored.

The assignment of breeds into households within a performance testing site was completely at random. A household received 25 brooded chicks of one of the five breeds. Performance data was collected on a total of 21,562 female chickens distributed to 2547 households across 45 performance testing sites (**Supplementary Table 5.3**).

Data preparation for MEPA

Supplementary Figure 1 shows that the performance data meets the assumptions of parametric analysis. Data inspection identified 8 possible outliers in the dataset for W90 and 9 possible outliers for W120 based on interquartile range (IQR). IQR is not affected by extreme values and has an advantage over standard deviation as a

measure of dispersion. Low outliers (below Q1 - 1.5*IQR) and high outliers (above Q1 + 1.5*IQR) of live body weight at a specific age were removed from further analysis. No outlier was detected for W180. The focus of the study was on comparison of breed performance within and between SDM-defined agroecologies and we used flock average of each household rather than individual weights from each household in our analysis.

5.2.4 Modelling of multi-environment performance data

The term 'stability' is used in the present study to describe the level of fluctuation in growth performance until specific time points of breeds across production environments (i.e., three SDM-defined agroecologies; AEI, AEII, and AEIII). Level of performance and yield stability of the breeds for three different traits (W90, W120, and W180) was evaluated through two different approaches with the R package *metan* (Olivoto and Lúcio, 2020).

5.2.4.1 AMMI-based productivity and stability analysis

Additive Main effects and Multiplicative Interaction (AMMI) model works under a fixed-model framework and is fitted at two stages. First, the main effects of the model (i.e., the effects of breeds and environments) are estimated using the additive two-way analysis of variance (ANOVA) least squares. Then, PCA is applied to the residuals of the ANOVA, which includes the interaction, to obtain the multiplicative terms of the AMMI model (Gauch Jr, 1992; Rodrigues et al., 2014; Zobel et al., 1988). The estimate of the response variable in a completely randomized design, can be obtained by the following model:

$$Y_{i,j} = \mu + \beta_i + \tau_j + \sum_{k=1}^{n} \lambda_k \gamma_{i,k} \delta_{j,k} + \rho_{i,j} + \epsilon_{i,j,k},$$

where $Y_{i,j}$ is the response variable (i.e., average live body weight in grams) of the ith breed (i= 1, 2, 3, 4, 5) in the jth environment (j = 1, 2, 3); μ is the grand mean; β_i are the breed deviations from μ; τ_j are the environment deviations from μ; λ_k is the singular value for the kth interaction principal component (IPC) axis; $\gamma_{i,k}$ is the ith element of the kth eigenvector; $\delta_{j,k}$ is the jth element of the kth eigenvector; $\rho_{i,j}$ is the residual, containing all multiplicative terms not included in the model; $\epsilon_{i,j,k}$ is the experimental error; and n is the number of principal components retained in the model.

We have chosen the AMMI stability value (ASV) and Yield and Stability Index (YSI) to analyse our data out of several AMMI based stability measures available in the plant breeding literature. The ASV is a relatively simple estimate based on IPCA1 and IPCA2 scores and can be more clearly explained in terms of environmental and / or biological factors. Moreover, according to Purchase et al. (2000), ASV is shown to be highly corelated with other stability measures such as those of Eberhart and Russel (Eberhart and Russell, 1966), Wricke (Wricke, 1964) and Shukla (Shukla, 1972). AMMI stability indexes were computed using the R package *agricolae* (de Mendiburu, 2021). The AMMI stability value (ASV) was computed from the first and the second interaction principal components (IPCA1 and IPCA2) of the AMMI model in line with Purchase et al. (2000).

$$\text{ASV} = \sqrt{\left[\frac{IPCA1_{Sum\,of\,squares}}{IPCA2_{Sum\,of\,squares}}(IPCA1_{score})\right]^2 + (IPCA2_{score})^2} \ .$$

The yield and stability index (YSI) was calculated by adding together the AMMI stability value (ASV) with the rank of mean growth yield of breeds (RY) across environments (i.e., $YSI = \text{rASV} + \text{RY}$). The advantage of YSI is that it incorporates both mean yield and stability into a single criterion. Low values of both parameters (i.e., ASV and YSI show desirable breeds with high stability).

Graphic nominal AMMI yield plot (Gauch Jr and Zobel, 1997) as a function of the environmental IPCA1 scores was used to visualize yield stability of the five improved chicken breeds across agroecologies for live body at the three different ages (WW90, W120, and W180). The winner breed in a given environment has the highest nominal yield in that environment.

5.2.4.2 LMM-based productivity and stability analysis

We have analysed our productivity data with four linear mixed-effects models (LMMs). In **model 1**, we fitted environment (ENV) as a fixed effect and included breed (GEN) and breed by environment interactions (GXE) as random effects. In **model 2**, we fitted breed (GEN) as a fixed effect and included replication (ENV), environment (ENV) and breed by environment interactions (GXE) as random effects. In **model 3**, we fitted breed (GEN), replication (ENV), environment (ENV) and breed by environment interactions (GXE) as random effects. In **model 4**, we fitted breed (GEN), replication (ENV), environment (ENV) and breed by environment interactions (GXE) as random effects. The best LMM model will be selected based on Akaike information criterion (AIC).

LMM for **model 1,** for instance, was fitted for the three live body weight traits (W90, W120, and W180) as follows:

$$y_{i,j,k} = \mu + \beta_i + \tau_j + (\beta\tau)_{ij} + \gamma_{jk} + \epsilon_{i,j,k},$$

where $Y_{i,j,k}$ is the response variable (i.e., average live body weight in grams at a household) observed in the kth site of the ith breed in the jth environment (i= 1, 2, 3, 4, 5; j = 1, 2, 3; k = 1, 2, ..11); μ is the grand mean; β_i is the random effect of the ith breed; τ_j is the fixed main effect of the jth environment; $(\beta\tau)_{ij}$ is the random interaction effect of the ith breed with the jth environment; γ_{jk} is the fixed effect of the kth block (REP) within the jth environment; and $\epsilon_{i,j,k}$, is the random normally distributed error.

We computed the harmonic mean of the relative performance of genotypic values (HMRPGV) of the five improved chicken breeds for three traits (W90, W120, and W180). HMRGV proposed by Resende (2007) was used efficiently as LMM-based stability index to select genotypes with high adaptability and stability (Azevedo Peixoto et al., 2018; Colombari Filho et al., 2013; Dias et al., 2018).

WAASB biplots (WAASBY stability index biplots) are generated based on LMM predictions (Olivoto et al., 2019a). WAASBY stability index biplots allow weighing and visualizing between stability (y-axis) and performance (x-axis) of the different chicken breeds across three agroecologies.

5.3 Results

5.3.1 AMMI-based productivity and stability analysis

A visual inspection of performance W90 (**Figure 5.1a,b,c**) shows that the productivity of breeds at specific ages varies across agroecologies (AEI, AEII, and AEIII).
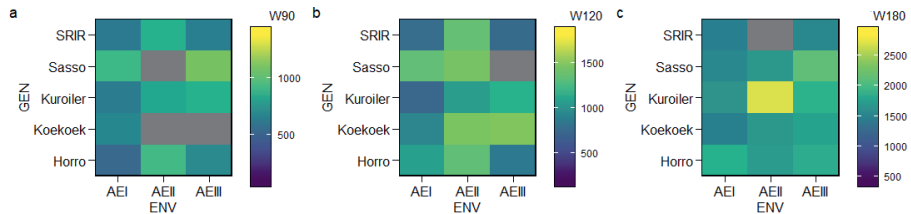
Figure 5.1. Visual inspection of performance data to see a change in the productivity of breeds (GEN) across environments, characterizing a cross-over interaction

The AMMI analysis of variance indicated highly significant (p <0.001) effects of genotype, environment, and their interaction for all traits (**Table 5.1**).

Table 5.1. Results of AMMI models for three traits (W90, W120 and W180) of improved chicken breeds

Trait	Source	DF	SS	MS	p-value	§Percent
W90	Environment (ENV)	2	2,992399	1496199	*	15.35
	Replicate/ Environment (REP)	20	12,768165	638408	*	
	Breed (GEN)	4	14,212622	3553155	*	72.92
	Interaction (GXE)	8	2,286701	285838	*	11.73
	PC1	5	237797	47559	*	65.90
	PC2	3	123029	41010	*	34.10
	Residuals	1376	52,503613	38157		
	Total	1418	85,124326	60031		
W120	Environment (ENV)	2	7,410000	3706919	*	26.46
	Replicate/ Environment (REP)	20	27,100000	1356594	*	
	Breed (GEN)	4	18,600000	4655698	*	66.42
	Interaction (GXE)	8	1,990000	249274	*	7.10
	PC1	5	270000	54032	*	86.20
	PC2	3	43300	14431	*	13.80
	Residuals	1324	81,400000	61508		
	Total	1366	137,000000	100229		
W180	Environment (ENV)	2	21,800000	11835319	*	47.29
	Replicate/ Environment (REP)	20	57,900000	3602218	*	
	Breed (GEN)	4	18,600000	14636059	*	40.35
	Interaction (GXE)	8	5,700000	1344379	*	12.36
	PC1	5	961000	191910	*	68.40
	PC2	3	94600	147610	*	31.60
	Residuals	1340	153000000	129221		
	Total	1382	258,000000	245711		

W90, W120, and W180 denote live body weight of female chickens at different days of age (at 90, 120 and 180). *Significant (p < 0.001) value. §Values obtained from AMMI analysis.

The treatment the sum of squares (SS) for growth performance is partitioned into three sources: the genotype main effect (GEN), the environment main effect (ENV), and the interaction (GXE)(**Table 5.1**). Genotype sum of squares (GEN SS) accounted for larger proportion of the treatment sum of squares (SS) for W90 (72.9%) and W120 (66.4%) compared to ENV SS (15.4% and 26.5%, respectively for the two traits) indicating the higher influence of breed on productivity. Environment (ENV SS) had higher contribution to productivity of W180 (47.3%) compared to breed (40.4%). This suggests that comparison of growth performance of breeds for W180 might be difficult based on AMMI analysis because of too much confounding between ENV and GEN. The AMMI analysis has identified only two principal components for having effects on productivity. Both the first and the second principal component (PCs) or Interaction Principal Component Axes (IPCAs) had significant effect ($p < 0.001$) on productivity of the three traits (W90, W120, and W180). PC1 or IPCA1 explained most of the GxE interaction for the three traits i.e., 65.9% for W90, 86.2 % for W120, and 68.4% for W180. Higher values of principal component (PC1) indicate GxE interaction (GXE) is important in explaining phenotypic variation in the dataset. The contributions of Interaction Principal Component axis 1 (IPCA1) and Interaction Principal Component axis 2 (IPCA2) to GXE sum of squares (SS) are used in the AMMI stability analysis of the breeds.

Table 5.2 shows AMMI stability indexes computed for W90. Breeds of average productivity (SRIR and Koekoek) are the most stable breeds based on rankings by AMMI stability value (rASV) while Sasso and Horro are the least stable. This is unlike the rankings by yield and stability index (rYSI) which identifies Sasso as the most stable breed followed by Kuroiler. The results suggest that the yield and stability index (YSI) is more influenced by productivity of the genotypes, giving higher stability values to higher performing breeds.

Table 5.2. AMMI stability indexes to analyse yield stability for W90 in improved chickens

Breed	Trait	Stability index			
	W90 (g)	ASV	YSI	rASV	rYSI
Horro	454.3	8.2	9	4	5
Koekoek	621.0	7.2	6	2	4
Kuroiler	735.1	7.8	5	3	2
Sasso	785.5	10.0	6	5	1
SRIR	654.2	5.8	4	1	3

rASV = rank (AMMI stability value) and rYSI = rank (yield and stability index).

The AMMI stability indexes calculated for W120 and presented in **Table 5.3** show the same rankings by the two indexes. Breeds with the highest productivity (Sasso and Kuroiler) are ranked as the most stable both by rASV and rYSI while breeds with the lowest productivity are ranked the least stable (Horro and SRIR).

Table 5.3. AMMI stability indexes to analyse yield stability for W120 in improved chickens

Breed	Trait		Stability index		
	W120 (g)	ASV	YSI	rASV	rYSI
Horro	693.2	162.2	10	5	5
Koekoek	936.9	25.9	6	3	3
Kuroiler	1000.2	4.9	4	2	2
Sasso	1043.7	0.9	2	1	1
SRIR	892.7	138.3	8	4	4

rASV = rank (AMMI stability value) and rYSI = rank (yield and stability index).

Table 5.4 shows the AMMI stability indexes (rASV and rYSI) computed for W180. Sasso, the breed with the highest performance was identified as the most stable breed by both indexes. The rankings given by the two indexes for this trait were less correlated for the other breeds.

Table 5.4. AMMI stability indexes to analyse yield stability for W180 in improved chickens

Breed	Trait		Stability index		
	W180 (g)	ASV	YSI	rASV	rYSI
Horro	1132.3	13.1	8	3	5
Koekoek	1477.5	22.9	8	5	3
Kuroiler	1692.6	15.5	6	4	2
Sasso	1697.7	9.5	2	1	1
SRIR	1455.2	10.7	6	2	4

rASV = rank (AMMI stability value) and rYSI = rank (yield and stability index).

AMMI nominal yield plots presented in **Figure 5.2** show the stability of the five improved chicken breeds for live body weight at three different ages (W90, W120, and W180) in three different agroecologies (AEI, AEII, and AEIII). The nominal yield plots have a complementary advantage over AMMI stability indexes because they integrate and visualize information on productivity and stability by agroecology. Koekoek was the most stable for W90. Sasso and Kuroiler are stable both across favourable (AEI) and unfavourable environment (AEIII) for W120. SRIR was the most stable for W180 across the three environments but not the most productive one.
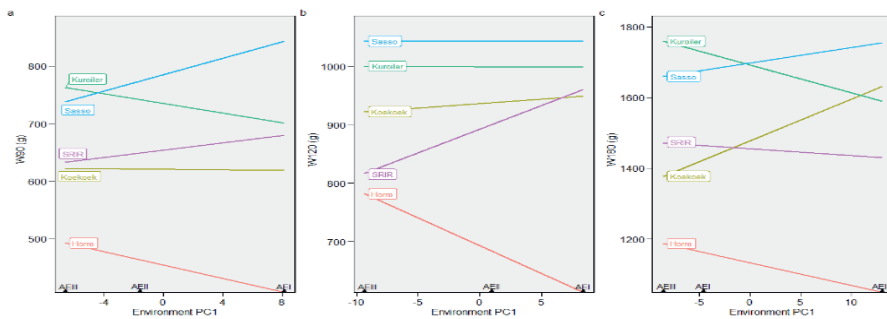
Figure 5.2. AMMI nominal yield plots showing mean growth performance and stability of five improved chicken breeds introduced into three different agroecologies. a)W90 b)W120 and c)W180

5.3.2 LMM-based productivity and stability analysis

The linear mixed-effects model (LMM) comparisons for three live body weight traits (W90, W120, and W180) of improved chicken breeds using AIC are presented in **Table 5.5**.

Table 5.5. Comparison of models fitted with linear mixed-effects model (REML) for three traits (W90, W120, and W180) of improved chicken breeds using AIC[1]

Model	Fixed effect	Random effect	AIC values for different traits		
			W90	W120	W180
1	ENV + REP (ENV)	GEN	60	0	1226
		GXE	69	20	1263
			*	*	*
2	GEN	REP(ENV)	468	496	1829
		ENV	237	197	1448
		GXE	258	224	1491
			*	*	*
3		GEN	295	254	1506
		REP(ENV)	515	543	1880
		ENV	283	244	1498
		GXE	303	272	1542
			*	*	*
4	ENV+GEN+GXE		486	523	1853
			*	*	*

[1]Akaike's information criterion (AIC) and likelihood value AIC were set to zero as reference for the best model; AIC = 2× # parameters – 2 × log-likelihood; thus lower values indicate a better model. *Significant (p < 0.001) value.

The model with the best fit (lowest AIC value) had environment (ENV) and environmental replicates (REP) as a fixed effect, and breed (GEN) and breed by environment interaction (GXE) as random effects (**model 1**)(**Table 5.5**).

Table 5.6. Harmonic mean of the relative performance of genotypic values (HMRPGV) stability index based on linear mixed-effect model

Breed	Trait		
	W90	W120	W180
Horro	0.70	0.77	0.76
Koekoek	0.95	1.03	0.99
Kuroiler	1.13	1.09	1.13
Sasso	1.20	1.14	1.14
SRIR	1.00	0.97	0.97

Table 5.6 shows harmonic mean of the relative performance of genotypic values (HMRPGV) stability index based on the best fit linear mixed-effect model (**model 1**). In the HMRPGV method for stability analysis, the breeds are simultaneously sorted by values for yield and stability using the harmonic means of the LMM so that the smaller the standard deviation of breed performance among the agroecologies, the greater the Harmonic Mean of Genotypic Values. The breeds with the highest stability (highest HMRPGV) in all the three traits (W90, W120, and W180) were Sasso and Kuroiler.
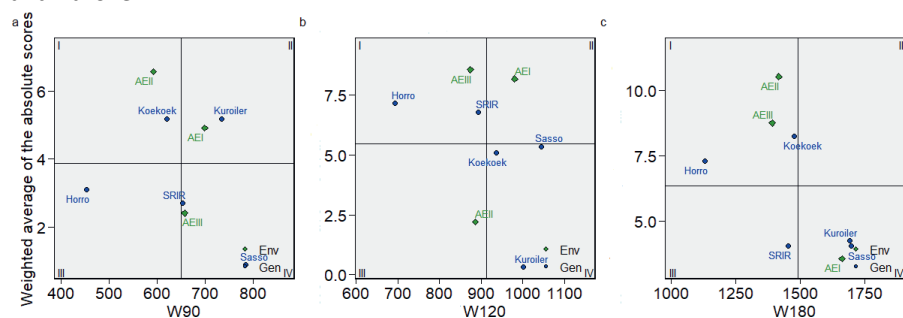
Figure 5.3. WAASB biplots generated based on LMM analysis showing mean growth performance (x-axis) and stability (four-quadrants) of five improved chicken breeds introduced into three different agroecologies. a)W90 b)W120 and c)W180.

The results of WAASB stability analysis based on based on the best fit LMM (**model 1**) are displayed in **Figure 5.3**. The four quadrants in each of the three biplots represent classifications regarding the joint interpretation of mean performance and stability. The breeds or agroecologies included in quadrant I are considered unstable. Besides, breeds or agroecologies in this quadrant represent high discriminatory ability and having productivity level below the grand mean. Unstable breeds with productivity above the grand mean are included in quadrant II. The only breed which fell into this category is Kuroiler (for W90). Breeds withing quadrant III (Horro for

W90; SRIR for W180) have low productivity but can be considered stable due to the lower value of WAASB. The lower the WAASB value, the more stable the breed can be considered. The agroecology included in the third quadrant are considered as unfavourable and having low discrimination ability. The breeds within quadrant IV (Sasso and Kuroiler) are considered highly productive and broadly adapted for W180.

5.4 Discussion

Breeds in tropical smallholder systems are locally adapted to challenging environments (e.g., to low availability of feeds, diseases, climatic extremes), but they are not productive enough to achieve economic efficiency. One approach to enhance the existing levels of performance in smallholder systems is introducing more productive breeds developed for similar management systems (Alabi et al., 2020; Birhanu et al., 2021). Candidate breeds need to be evaluated for their performances before they are introduced at scale to farmers.

Analytical frameworks (experimental designs and evaluation methods) are not available to compare candidate breeds and identify those with optimal productive performance and adaptability for wider use by smallholder farmers. Previous researches that focused on estimating GxE in smallholder poultry systems (Alemu et al., 2021; de Kinderen et al., 2020; Lozano-Jaramillo et al., 2019a; Lozano-Jaramillo, 2019) did not consider evaluation of yield stability in livestock performance analyses. Yield stability refers to the ability of a breed to be less affected by environmental influences and to provide reasonable yield. Evaluating breeds for yield stability apart from productivity levels helps to recommend those populations of animals having both of these qualities. Plant breeders have long exploited the concept of yield stability (phenotypic plasticity) in their desire to select and recommend productive and adaptive genotypes. The most commonly used procedures of stability analysis include Additive Main Effects and Multiplicative Interaction Methods (AMMI; (Gauch Jr, 2013)) and Linear Mixed-effect Models (LMM; (Piepho, 1994; Piepho and Möhring, 2005)).

In the present study we have used chicken agroecologies defined by species distribution models (SDMs;(Kebede et al., 2021; Kebede et al., 2022)) to group environments of smallholder farmers based on environmental predictors influencing habitat suitability and phenotypic differentiation of the species. Then we adapted two methods (AMMI and LMM) for multi-environment performance analysis to evaluate productivity and yield stability of five improved chicken breeds distributed to smallholder farmers in Ethiopia.

In both methods of MEPA applied in the present study, breed by environment interaction (GxE) was detected and Sasso and Kuroiler were identified as the most productive breeds for the three growth traits (W90, W120, and W180). Koekoek and SRIR had comparable but average productive performance for these traits. The locally improved Horro had the lowest productive performance. This is not surprising given that Improved Horro is a breed under development and has only been selected for ten generations from the population of indigenous Horro chickens when it was included in the performance comparison trial.

The ranking of breeds for yield stability varied according to the method of MEPA and stability indexes used and was not consistent. However, the most advanced indexes based on LMM favoured the breeds with superior productive performance (Sasso and Kuroiler). Studies comparing the different methods (AMMI versus BLUP) are documented for other species (Balestre et al., 2009; Sa'diyah and Hadi, 2016). LMM models allow fitting of random effects to improve model efficiency. After comparing Eberhart-Russel, AMMI, and mixed-models, Ferraudo and Perecin (2014) concluded that the three methods detected GxE differently but effectively, with mixed models showing the highest sensitivity. Some of the indexes (e.g. WAASB) can even integrate the features of AMMI and LMM techniques and display their results on performance using biplots (Olivoto et al., 2019a). Biplot-based interpretation of MEPA on LMM results helps easily identify 'which breed wins where' and facilitates recommendation of specific breeds for specific agroecologies. LMMs such as the best linear unbiased prediction (BLUP) and restricted maximum likelihood estimation (REML) are more predictively accurate than AMMI and the main advantages of these methods can be combined to recommend highly productive and stable individuals or populations (Olivoto et al., 2019a; Van Eeuwijk et al., 2016). Kuroiler (W120 and W180) and Sasso (W180) were identified as breeds combining high growth performance with high stability based on WAASB biplots of LMM. Interpretations regarding performance (stability and productivity) of the breeds shall better be decided based on W180 considering that that the chickens have stayed long enough in the respective agroecologies to experience the challenges and express their genetic potential compared to the other two traits (W90 and W120). AMMI (rASV and rYSI) and LMM (HMRPGV) stability indexes have converged at W180 and jointly identified Sasso as the most stable of the five chicken breeds evaluated in this study for live body weight.

The importance of reducing possible confounding between management practices of individual households and the biophysical environmental factors (e.g., climate,

geography) should be underscored in multi-environment performance evaluation studies of livestock. In the present study, we tried to standardize the management practices of the farmers by taking precautionary measures. The amount of feed supplemented to the chickens on daily basis on top of scavenging was regulated, animal health practices were reasonably kept uniform, and night enclosures were put in place in all the households. It is also essential to include sufficiently large sample size in such on-farm performance evaluations to offset possible variations in the management of flocks between households. Putting experimental units (households) that are similar as possible (i.e., have access to similar feed resources, comparable level of health services, and share common culture) and assigning all types of breeds into each performance testing site (block) can further help reduce experimental error.

There was no pedigree information and genomic data available in the present study to fit in mixed models and estimate important genetic parameters (e.g., variance components, heritability) that would have improved the predictive ability of MEPA. Future research may consider incorporation of additional sources of information (e.g., egg productivity, survival, economic efficiency, farmers' trait and breed preferences) which contribute to better identification/recommendation of suitable breeds and their sustained adoption by smallholders.

5.5 Conclusion

The present study has demonstrated that agroecologies defined by Species Distribution Models (SDMs) are useful to undertake multi-environment performance analysis (MEPA) in livestock. Our results also show that existing methods of MEPA from plant breeding are equally applicable to livestock breed comparison in smallholder systems. Based on the results of the present study and the existing literature, LMM-based productivity and stability indexes/biplots combine information on productivity with yield stability and are strongly recommended for wider use in livestock breed performance comparisons across agroecologies.

5.6 Data availability statement

The datasets analysed during this study and supporting its conclusions are available in the data repositories of International Livestock Research Institute (ILRI) https://data.ilri.org/portal/dataset/acggonfarmeth

6

General discussion

6.1 Introduction

Smallholder chicken production is an integral part of the farming system in tropical regions of the world. Chickens kept in smallholder systems account for 80% of the poultry flock in sub-Saharan Africa and significantly contribute to income generation, food security, ecosystem services, and cultural values (Birhanu et al., 2022; FAO, 2020). Chicken meat and eggs are a relatively affordable source of animal protein to millions of people who live in sub-Saharan Africa and Southeast Asia.

The demand for animal source foods will increase substantially in these regions in the coming decades (Desiere et al., 2018). Asian countries increase consumption because of rapid economic development while the increased demand in African countries is driven mainly by human population growth (ILRI, 2019). The demand for meat (beef, sheep, goat, poultry, and pork combined) could reach around 221 million tonnes (MTs) in Asia and 58 million MTs in Africa by 2050 (Robinson et al., 2015). The projected increase in egg consumption between 2005 and 2015 was 26 percent in these regions, compared with only 2.4 percent in the most developed countries (Windhorst, 2008).

Smallholder chicken production system in Africa is characterized by low input-low output production system and is unable to meet the growing demands. Chickens kept by smallholder have small body size (low nutritional maintenance requirement) and low egg production performance (18 to 100 eggs per year; with small egg size of 34 to 48 grams). Sexual maturity delays up to 36 weeks of age and clutch size varies from 2 to 12 eggs, with a predominant inclination to broodiness. Local males reach a live weight of 1.5kg at 6-months-of-age while females weigh 30 percent less at a similar age (Sonaiya and Swan, 2007). The reproductive cycle takes about 74 days (16 days for egg laying and clutch formation, 21 days for hatching, and 37 days for chick rearing) in total (Dessie, 2003b; Getachew et al., 2016).

With appropriate genetic interventions and development strategies, smallholder poultry enterprises have the potential to transform the livelihoods of smallholder farmers (Birhanu et al., 2021; Birhanu et al., 2022). Attempts at enhancing chicken productivity in many parts of Africa have relied on the introduction of best performing breeds. Most of these breeds were developed and perform well in temperate production systems (i.e., in areas with less environmental stressors) and are not expected to exhibit similar performance in smallholder systems of the tropics. The lack of proper matching between the right genetics and the right environment is a major factor for low productivity (de Kinderen et al., 2020; Lozano-

Jaramillo et al., 2019b). Testing candidate breeds and accounting for genotype by environment (GxE) interactions is essential to identify the most productive and adaptive breeds for a new environment. However, current approaches of evaluating GxE for chicken breeds introduced into smallholder systems have significant drawbacks to achieve this goal. Firstly, breed comparison studies are based on agroecologies defined on suitability of the environment for major crops. GxE estimates based on these agroecologies may not give accurate results for levels of GxE. Conventional definition of agroecology in Ethiopia considers agronomic parameters such as the length of crop growing period, soil types, and the types of major crops cultivated (Hurni, 1998; MoA, 1998, 2000). Secondly, current agroecological definitions consider an agroecology as a homogenous class and treat its effects as a whole. There is little effort to identify and evaluate the contributions of individual environmental predictors within that environmental class. Model fit and accuracy of GxE predictions for a livestock species might improve if the agroecological classes are defined and performance is predicted considering the most important environmental predictors influencing habitat suitability and phenotypic variation for that species. The influence of certain climatic factors such as extremes of temperature and solar radiation on performance, physiology and behaviour of chickens has been documented (Bettridge et al., 2018; Gicheha, 2021; Lara and Rostagno, 2013). Thirdly, the existing agroecological classes assume all livestock species have similar environmental requirements and hence breeds of any species can be compared for performance and adaptability based on the existing classes. However, livestock species vary in their environmental requirements (e.g., chicken vs camels) and livestock-species-specific agroecologies may allow a more accurate comparison of performance among breeds within a species. Fourthly, current approaches of studying GxE in smallholder livestock systems that consider the environment as a continuous gradient often assume a linear relationship between breed performance for a trait and an environmental class. To explain such relationships, they use reaction norm plots (e.g. the change of body weight of breeds 1 and 2 in response to a range of values for an environmental parameter). Analytical methods developed in ecology, geographic information system, and machine learning can be applied to overcome limitations of conventional GxE prediction methods in the context of smallholder livestock systems.

In part-I of my thesis, I studied indigenous chicken populations kept by smallholder farmers. Indigenous chickens are managed for generations with minimum external input, in semi-feral status (i.e., they mate at random, are nondescript, they fulfil their daily nutritional requirements through scavenging or semi-scavenging, spend the

night in simple enclosures) and have direct contacts with the physical environment. As such, they are suitable to investigate the effects of evolutionary selective forces on adaptive phenotypic and genetic variation. I applied species distribution models (SDMs) to identify the most important environmental predictors contributing to habitat suitability of chickens.

Once I identified the most important environmental predictors with SDMs, I used them to redefine chicken production agroecologies. Along with the new agroecological classes, the most contributing climatic predictors were individually considered to improve model fit in GxE predictions through phenotypic distribution models (PDMs). Instead of reaction norm plots, I used partial dependence plots to explain the relationship between phenotypic values for a trait and a specific environmental predictor. PDMs such as generalized additive models, allow incorporation of the effects of specific predictors on top of agroecological classes for each species and estimating phenotypic values.

PDMs are useful to overcome certain shortcomings linked with SDMs. SDMs were initially developed for wild populations and lack the concept of breed or sub-populations, which are important in animal breeding. SDMs (Phillips et al., 2006) implement a correlative approach to establish relationships between species occurrence and habitat suitability and are also not suitable to declare cause-effect relationships when used alone. By use of PDMs, we can combine suitability information derived from SDMs (on agroecology and most important environmental predictors) with breed (ecotype) information to predict performance across the environmental landscape of present and future production conditions for livestock.

In a separate chapter in part-I, I continued to use the most important environmental predictors identified through SDMs to identify genes that are putatively associated with local adaptation. I wanted to test whether phenotypic differentiation among indigenous chicken populations along the environment has a genetic basis. I followed a robust sampling strategy and study design which allowed inclusion of sample chicken populations from different environments and geographies. Based on higher matching percentage between the presence of distinct phenotypes and specific environmental niches, I grouped the Ethiopian indigenous chicken populations into three ecotypes. I did signatures of selection analyses with two different methods (F_{ST} and XP-EHH) at three analytical layers (between gradients, between agroecologies across gradients, and between agroecologies within gradient) and looked for overlap between their respective results. I performed redundancy

analysis (RDA) to detect the association between environmental predictors and genomic variation. I also performed RDA to explain variation in the chicken genome associated with phenotypic variability. In doing so, I demonstrated that RDA is applicable in indigenous livestock as an alternative method to undertake genome-wide association analysis (GWAS). I show that the results from RDA are supported by the outputs from signatures of selection analyses (F_{ST} and XP-EHH). I demonstrate that signatures of selection analysis with the two methods (F_{ST} and XP-EHH) can be used complementarily with RDA to shed light on the relationship between genomic, phenotypic, and environmental variation in local adaptation studies in indigenous chickens.

In <u>part-II</u> of my thesis, I wanted to validate if SDMs can be applied to identify environmental predictors influencing habitat suitability for improved chicken breeds introduced into smallholder systems. I also wanted to improve model fit in GxE predictions and to compare performance among breeds. Improved chicken breeds have been artificially selected for specific traits (e.g., growth, egg production or both) and are useful to fit models quantifying breed by environment interactions. I followed a different approach to establish my sampling frame and to analyse environmental and phenotypic data through SDMs and PDMs. This included taking large sample size for phenotypic and environmental data, clustering environments to represent contrasting habitats where chickens can be potentially kept, and using a large number of background points for environmental analysis. Background or pseudo-absence data were drawn at random from the entire country, while occurrence locations are spatially biased toward chicken performance testing areas. By including sufficient number of background points, we can improve the accuracy of our species distribution models (Phillips et al., 2009). I used SDM-identified environmental predictors for chicken performance testing sites to redefine agroecologies, and showed that this improves model fit in GxE predictions. I also adapted two commonly used methods in plant breeding to perform multi-environment performance analysis (MEPA) for chicken breeds. Apart from growth performance, I was interested in comparing breeds on their phenotypic plasticity in different SDM-defined agroecologies. This approach helped to identify those chicken breeds with the best productive performance and with wider adaptability in smallholder systems.

6.2 Distribution models identify environmental variables linked with habitat suitability and phenotypic variation

Understanding the relations between phenotypic variation and environmental variation is useful to predict populations' response to environmental change and expected levels of animal productivity. Evidence on presence of phenotypic differentiation in response to climatic and other environmental variables among Ethiopian indigenous chicken populations was not available.

In chapter 2 of this thesis, I sampled Ethiopian chickens from different locations and analysed their environmental and phenotypic data. I used environmental-gradation approach to survey populations across all possible agroclimatic clines in the country. This approach was not followed by previous studies conducted on Ethiopian chickens (Adebabay, 2018; Gheyas et al., 2021; Vallejo-Trujillo et al., 2021). I employed different analytical techniques (e.g., distribution modelling, cluster analysis) and combined environmental and phenotypic data to detect population differentiation in the environmental space. I found that Ethiopian chickens have differentiated phenotypically most probably in response to environmental variation (chapter 2). I fitted Species distribution modelling (SDMs) to produce habitat suitability maps for indigenous chicken breeds. The habitat suitability maps were produced based on a final set of nine most contributing and least correlated environmental predictors identified through variable selection (Jueterbock et al., 2016) out of an initial set of 34 predictors. Habitat suitability in this thesis refers to a set of environmental conditions favouring better reproductive (egg number) and productive (growth rate) performance for the species. Eight, out of the nine SDM-identified predictors also had significant effects on phenotypic differentiation of quantitative traits. Several of these SDM-identified environmental predictors, such as temperature, precipitation, and solar radiation were reported in earlier studies for their influences on availability of feed and prevalence of diseases and parasites (Bettridge et al., 2018; Lozano-Jaramillo et al., 2019b). On the other hand, starting from the same set of 34 environmental predictors, only six were identified for their relationship with habitat suitability of improved chickens.

Out of the six, three of the predictors (solar radiation in May, water vapour pressure in May, and precipitation of the coldest quarter) were commonly identified by SDMs for both indigenous and improved chickens indicating their strong influences on habitat suitability at species level. Differences in the sets of environmental predictors identified for the two types of populations suggests that predictor selection by SDMs is probably influenced by sampling strategy. Attention should be given to the choice

of species occurrence locations, background locations, and sample size. Be it for indigenous or improved breeds, the sampling frame for SDMs should be developed considering environmental variability, geography, and the presence of chicken as a species as opposed to considering the distribution of introduced breeds. Distributing the breeds and then using the occurrence points of the chicken performance testing sites for habitat suitability mapping may lead to bias. In our case, while we followed a hybrid sampling strategy (considering environmental and geographic variability) for both types of chicken populations, the number of locations used for variable selection step in SDMs for indigenous chickens (n = 260) was lower than those used for introduced breeds (n = 2547) probably having consequences on the number and type of predictors selected and eventually on the habitat suitability maps produced. Apart from the environmental values of each of the most influential predictors identified by SDMs, the inclusion of genomic information from the animals may improve predictive ability of phenotypic distribution models (PDMs). More studies need to be conducted to evaluate the effects of sampling strategy and inclusion of additional sources of information on predictive ability of PDMs in livestock.

Out of 19 traits measured on adult birds, eight traits were most useful in explaining the influence of environmental predictors on phenotypic variation in indigenous chickens. These included body length, wingspan, comb length, comb width, live body weight, earlobe width, wattle width, and beak length. Adaptive roles of quantitative phenotypic traits included in the present study are documented in literature. For instance, comb shape is linked with adaption to cold stress (Wright et al., 2009). The pea-comb, a dominant mutation in chickens, drastically reduces the size of the comb and wattle, decreasing heat loss and making the chicken less susceptible to frost lesions (Wright et al., 2009). Histological analysis of dermal papillary layer has revealed that red earlobes have many more blood vessels and were associated with thinner skin than that of white earlobes (Luo et al., 2018) indicating the role of earlobes in thermoregulation.

SDM identified environmental predictors can be incorporated into phenotypic distribution models (PDMs) to improve model fit in GxE predictions. PDMs based on Generalized additive models (GAMs) were applied to integrate environmental and biological information (e.g., type of breed or ecotype) and predict phenotypic values for specific traits across the environmental landscape in chapter 4. The study demonstrated that GAMs achieve acceptable goodness of fit when investigating complex (e.g., non-linear) relationships between phenotypes and environmental

predictors. I investigated the effects of SDM-defined-agroecology, chicken breed, and their interactions on three growth traits.

6.3 Distribution-model-defined-agroecologies improve model fit in GxE

Agroecological zonation refers to the spatial classification of the geographic and environmental landscape into units, each of them with similar agricultural and ecological characteristics within. Such classifications into roughly similar units has been used in agronomy and forestry to compare agroclimatic conditions and manage land resource conditions considering soil, water, vegetation, topography parameters (Hurni, 1998). There were some attempts to use agroecological zonation in animal production as an operational framework to improve productivity (Dumont et al., 2014; Soussana et al., 2015). However, proper methodologies for definition of species-specific livestock agroecologies for use in animal breeding programmes were not available.

The SDM-identified environmental predictors from chapter 4 of this thesis were used to group the locations of chicken performance testing sites into agroecologies. The newly defined agroecologies were used in phenotypic distribution models (PDMs) to predict phenotypic values of three production traits (female live body weight at 90-, 120- and 180-days-of-age (i.e., W90, W120, and W180). I hypothesized that model fit in GxE predictions could be improved if I redefine agroecologies for chicken production based on environmental predictors influencing habitat suitability for chickens. This is unlike previous GxE studies conducted on introduced chickens (de Kinderen et al., 2020; Lozano-Jaramillo et al., 2019a) which relied on conventional agroecological classifications to evaluate performance. My findings suggest that this grouping of environments into environmentally homogenous units based on the most contributing environmental predictors is useful to improve model fit in GxE estimations. The AIC values for PDMs fitted on SDM-classified agroecologies were lower. The study also demonstrates that PDMs like boosted GAMs and boosted GLMs are useful in animal breeding to integrate environmental and phenotypic data and predict phenotypic values.

6.4 Improved chicken breeds introduced into smallholder systems show different levels of productivity

Three main genetic improvement strategies could be considered to improve the productivity of chickens in smallholder systems: selection, cross-breeding, and the introduction of improved breeds. Introducing productive and adaptive breeds to

smallholder systems can potentially enhance productivity in a relatively short period. It helps to improve household nutrition, increase farm income, create employment opportunities, and contribute to gender empowerment.

In underline chapters 4 and underline chapters 5 I evaluated the growth performance of improved dual-purpose chicken breeds (i.e., developed both for meat and eggs) introduced to Ethiopian smallholder farmers in different agroecologies. The five chicken breeds can be grouped into three types based on their genetic backgrounds: composite breed (Potchefstroom Koekoek), commercial hybrids (Kuroiler, Sasso, and SRIR), and locally improved chicken (Improved Horro). Koekoek is a composite breed and is described in more detail in (Fourie and Grobbelaar, 2003; Grobbelaar, 2008; Grobbelaar et al., 2010; Wondmeneh et al., 2011). The Kuroiler is a commercial hybrid which originated from the crossing of the Rhode Island Red, the White Leghorn, the Barred Plymouth Rock and two Indian indigenous chicken breeds. Introgression of a broiler genotype was also made to incorporate broiler characteristics into the breed (Ahuja et al., 2008b; Isenberg, 2007). The Sasso is a dual-purpose commercial hybrid developed by a poultry breeding company in Europe (IPP, 2018). The SRIR is a hybrid closely related with the Sasso, with some genetic introgression from another dual-purpose chicken genotype. Improved Horro originates from the indigenous, non-descript, Horro chickens kept by smallholder farmers in Horro district of Western Ethiopia, in Oromia region (Esatu, 2015; Wondmeneh et al., 2011).

The analyses on the growth performance of the five chicken breeds until selected time points show that two of the breeds (Sasso and Kuroiler) were the most productive and two had near average performance (SRIR and Koekoek). The performance of the locally improved breed (Horro) was not as good as the other four strains in the test. This is not surprising given that Horro had only been selected for 10 generations at the time of the on-farm performance evaluation. The performance of improved Horro was reported in other studies to be significantly higher than unimproved local Horro chickens (Esatu, 2015; Wondmeneh et al., 2011). The economic efficiency of the five breeds has not been investigated in the present study. Improved Horro might still be an appropriate choice for introduction in areas with lower access to poultry inputs and services (e.g., formulated feeds, market infrastructure) to sustain commercial hybrids and sell their products.

6.5 MEPA combines productivity and yield stability to identify the best breeds

Statistical frameworks for multi-environment performance analysis (MEPA) are common in plant breeding to compare different genotypes in different agroecologies. Such frameworks can benefit animal breeding programmes that are aiming at testing candidate genotypes (breeds/strains/lines) for introduction at scale into new environments. Especially when the environment of introduction is not uniform, a simple experiment that does not consider the environmental heterogeneity and specific breed responses is not sufficient to make performance evaluations.

In chapter 5, I evaluated different MEPA models for their use in comparison of chicken breeds for growth performance and yield stability. Yield stability means there is less phenotypic plasticity for the yield traits. Three live body weight traits in female chickens, namely body weight at 90-, 120-, and 180-days-of-age were evaluated. I used the SDM-based agroecological classification that was developed in chapter 4 of this thesis to evaluate the performance and yield stability of Sasso, SRIR, Kuroiler, Horro and Koekoek breeds in smallholder systems of Ethiopia. Two MEPA methods were applied: additive main effects and multiplicative interaction (AMMI) models, and linear mixed-effects models (LMM). While their approaches of estimating GxE varied, both multi-environment breed performance comparison methods were found applicable to identify more productive and stable breeds before wider introduction to smallholder systems. LMMs, more specifically, best linear unbiased prediction (BLUP) and restricted maximum likelihood estimation (REML) give more accurate predictions than AMMI (chapter 5). Combining the features of different models of multi-environment performance analysis such as through the use of more advanced indexes like WAASB helps overcome limitations associated with the use of a single method (e.g. in terms of visualization of results) and provides wider insights on stable and productive genotypes (Olivoto et al., 2019a). I used the WAASB biplot to show productive yet stable chicken breeds in different SDM-defined agroecologies.

Ranking of breed productivity and yield stability could be summarized based on indexes developed for such purposes. Based on the findings of the present study and documented literature (Balestre et al., 2009; Sa'diyah and Hadi, 2016), LMM-based performance indexes are powerful to rank genotypes by taking into account high productive performance with stability. Hence, they can help identify and recommend based on information of 'which breed wins where'. The chicken breeds

which were evaluated in the present study were phenotypically plastic (showed varying yield or were less stable) for growth probably due to their different genetic backgrounds (e.g., selection history, levels of inbreeding, or presence of different sets of responsible genes for plasticity). Inbreeding may directly affect plasticity by altering phenotypic expression in one environment or by altering the organism's ability to detect or respond appropriately to different environmental conditions (Auld and Relyea, 2010; Smith, 1956). The consequences of inbreeding on phenotypic and behavioural plasticity varies with the organism being studied (Murren and Dudash, 2012; Schlichting and Levin, 1986; Schou et al., 2015).

6.6 Candidate genes linked with local adaptation and phenotypic variation

Landscape genomics combines environmental, genomic, and phenotypic information in studying adaptive variation. The presence of several climatic conditions and rich domestic animal genetic diversity in Ethiopia makes it an ideal place for landscape genomic studies. Altitude ranges from 125 m below sea level in the Danakil Depression to as high as 4620 m above sea level in the Semien Mountains, creating a striking geographic contrast in the country.

Evidence on population differentiation of Ethiopian chickens (chapter 2) alone was not enough to claim that phenotypic variation has a genetic basis. Hence, genomic analysis of sample populations (chapter 3) was undertaken. I implemented a hybrid sampling strategy that captured adaptive and neutral genomic variation of indigenous chickens across the environmental space in Ethiopia. With a landscape genomics approach, I synthesized genetic, phenotypic, and environmental information to get evidence of local adaptation.

Confounding between neutral and adaptive variation becomes a serious problem in detection of selection signals when populations are characterized by high genetic structure. I clustered the environments in chicken sampling sites into geographic (latitude and longitudes) and agroecological (based on elevational classes influencing climate) classes and considered these sources of variations in my analytical models. I estimated F_{ST} values to quantify genetic differentiation between environmental classes (gradients and agroecologies). A different approach can be followed to detect neutral variation i.e., by measuring additional population genetic parameters such as effective population size, inbreeding, and linkage disequilibrium).

The results from structure analysis showed the presence of two major ancestral gene pools in Ethiopian chickens. Populations of *gradient-II* clearly separated from populations of the other three gradients (*-I, -III,* and *-IV*) during principal component (PCA) analysis suggesting that evolutionary forces like geographic distance (drift) and migration history might have played a role in genetic differentiation of Ethiopian chickens in addition to adaptive processes in different agroecologies (natural selection driven by environmental variation). Neutral variation in this chapter needs to be analyzed using additional population genetics parameters (e.g., on gene flow, effective population size, inbreeding, and linkage disequilibrium).

Out of the populations studied in chapters 2 and 3, morphologically unique phenotypes were seen at higher frequencies in populations derived from environmental *gradient-III* (in Metekel zone of Benishangul-Gumuz region of Ethiopia), particularly in populations sampled from lowlands between 500 to 1000 meters above sea level (**Figure 6.1).**

Figure 6.1. Unique morphological phenotypes sampled from sample populations in *gradient-III*. **(A)** Frizzle feathered cockerel from Bengo; **(B)** Feathered shank(feathered foot) cockerel from Bengo; **(C)** Frizzle chicks from Bengo; **(D)** Naked-neck hen from Tumi; **(E)** Stunted tail (tailless) cockerel from Zigh; **(E)** Dwarf hen from Zigh.

Phenotypes such as naked-neck, frizzle feather, feathered shank (feathered feet) have roles in thermoregulation and are considered beneficial to environmental adaptation to hot climates. These unique phenotypes (with underlying mutations) are present in higher frequency in lowland populations of the district. The maximum temperature in Guba district from where these populations were sampled can reach as high as 48^{0}C. These areas have also less access to other parts of the country because of their geographic distance, and are shown to be biodiversity hot spots for chickens. High phenotypic variation has also been reported in this district for other livestock species (Getachew et al., 2016), cattle (Fasil Getachew, 2020), and donkeys (Mustefa et al., 2020).

Figure 6.1 A and **C** show frizzle feather phenotypes. Frizzle feather chicken are highly valued in Africa and the trait is speculated to confer adaptability to tropical climates (Adomako et al., 2014). Frizzle plumage may cause the acceleration of basal metabolism due to the loss of body heat (Ng et al., 2012). The frizzle phenotype is due to an alpha-keratin (*KRT75*) mutation (Ng et al., 2012). Preliminary analysis of structural chromosomal variants in my sequence dataset revealed a novel mutation of 16 kb deletion at the *KRT75* instead of 84bp deletion reported by Ng et al. (2012). The presence of novel mutations in Ethiopian indigenous chickens indicates phenotypic variability at individual and population level may have a unique genetic basis which needs to be explored further. Phenotypic variability that warrants further investigation into its genetic control includes sample chickens showing feet feathering (ptilopody) **(Figure 6.1 B).** Mutations causing feathered shank (foot feathering) in chickens and pigeon was investigated in Bortoluzzi et al. (2020). They found a parallel genetic origin of foot feathering in the two species. I would like to find out if the same genes affect molecular pathways leading to foot feathering in indigenous chickens **Figure 6.1 D** shows a hen with naked neck phenotype. The naked neck mutation was described as an insertion of ~180bp in *GDF7* protein coding gene and is suggested to contribute to heat tolerance and robustness through efficient dissipation of heat from the body (Desta, 2021). Some chickens have stunted tails (**Figure 6.1 E**) compared to the normal phenotype with long tails. Stunted tail is also called rumpless when this abnormal phenotype is caused by a reduction in the number of coccygeal vertebrae in some breeds of chickens (Dunn, 1925). Stunted tail chickens may possess the normal number of coccygeal vertebrae but have abnormal feather development because of a pseudogene *LOC431648* located on the Z chromosome (Wang et al., 2017).

A hen with a mature live body weight of only 750 grams (**Figure 6.1 C**) was the lightest of all the 380 adult females measured in entire dataset. The average weight for female chickens for all samples collected in the country was 1000g and this small hen was suspected to be carrying a mutation causing dwarfism. A novel nonsense mutation in the *TMEM263* gene is reported as associated with dwarfism in Dutch bantam chickens (Wu et al., 2021).

Some of these major phenotypic differences explained above could be adaptive. In cases when adaptive variations cannot be recognized phenotypically, cross-population differences in allelic and haplotype frequencies should be detected. I applied two of such techniques (F_{ST} and XPEHH) in chapter 3. F_{ST} measures population genetic differentiation based on the variance of allele frequencies between populations while XP-EHH measures reduction in haplotype diversity between populations by computing the probability that two extended haplotypes around a given locus are the same. The presence of considerably large overlap indicates that both methods capture regions under selection. The overlap between the two methods was also large (19.2%) during pairwise gradient comparisons that did not consider the effects of agroecological variations. This suggests that populations in different geographies may have been exposed to different environmental selective pressures.

A relatively low overlap (13.4%) between the two methods in pairwise agroecological comparisons across gradients and high overlap (20.9%) for pairwise agroecological comparisons within a gradient shows that the two methods are better at detecting signatures of selection when the confounding between neutral processes (e.g., genetic drift and gene flow associated with specific geography) and adaptive processes (i.e., natural selection associated with environmental variations between agroecologies) is reduced. That is achieved by undertaking the analysis within the same gradient.

The selection signals seem to have been diluted when comparisons between agroecologies are made without considering their gradients. Populations from *gradient-II* have a different demographic history (constitute a different genetic structure) and that may have also contributed to higher overlap between the two methods within this gradient. I hypothesise that stronger signals by F_{ST} over XP-EHH shows the influence of neutral processes is higher while stronger signal by XP-EHH over F_{ST} shows the influence of selective processes is higher. The overlap between F_{ST} and XP-EHH was generally larger than what was previously reported by Gheyas

et al. (2021) on Ethiopian indigenous chickens (4.9%). In their study the sampling strategy that did not consider environmental gradation.

In underline chapter 3 I also performed Redundancy Analysis (RDA). The use of RDA in genome-wide association studies (GWAS) in animal breeding is not common. However, RDA is a powerful approach that can be utilized in GWAS studies in livestock (Kess and Boulding, 2019; Torrado et al., 2020; Valette et al., 2020). Phenotype association analysis was performed with RDA to explain variation in the genome associated with five quantitative traits (mature live body weight, beak length, comb width, wattle width and earlobe width).

Information on genomic regions associated with quantitative traits, such as body weight, can be useful to develop breeds with wider environmental adaptation and better performance. Significant candidate SNPs ($p < 0.01$) associated live body weight of chickens ($r > |0.3|$) were identified by RDA. Body weight associated genomic regions were also identified in some African and Asian indigenous chicken populations through conventional genome-wide association studies (Cha et al., 2021; Habimana et al., 2021). Chapter 3 demonstrates that RDA was able to identify SNPs that are associated five quantitative traits, selected for their usefulness in discriminating (classifying) sample chicken populations on phenotype. These SNPs may have roles in adaptive phenotypic variation among indigenous chickens in response to environmental variation. More studies, comparing the outputs of RDA analysis with conventional GWAS need to be conducted to promote its use in animal breeding. In theory, variation in RDA identified SNPs can be incorporated into genomic selection programmes to improve prediction of performance and stability. The contributions of candidate SNPs should also be supported by additional studies (e.g., functional annotation, gene expression).

6. 7 Recommendations and future research areas

6.7.1 Multi-environment genetic evaluations should be based on agroecologies defined for the targeted species

Classifications of environments into agroecologies provides an analytical framework for breeders to make performance comparisons among different genotypes for wider introduction into smallholder systems. A common practice for livestock breeders aiming at identifying best breeds for specific environments has been to adopt agroecologies classified for crops. For instance, all agroecologies in Ethiopia were defined considering environmental parameters influencing the suitability of a

given agricultural area for major crops, such as the length of growing the period. Classification of agricultural production areas into agroecologies based on predictors such as cropping pattern and land use type is less informative to evaluate the performance of livestock breeds.

The analytical framework used for the classification of poultry production agroecologies in the present study was based on Species Distribution Models (SDMs). SDM-defined agroecologies could be developed for specific livestock species (e.g., cattle, small ruminants, swine) and utilized for breeding and production purposes in low- and medium-input systems. To define agroecology, we need to identify the most important environmental predictors associated with suitability of habitat. This can be done using machine learning algorithms like MaxEnt. Because of differences in occurrence locations for the different species (e.g., chickens versus camels in extreme case), the predictors that will be identified by the models and resultant suitability maps will be species-specific. The values extracted for each of the most contributing environmental predictors from global climatic/environmental databases will be used as inputs to calculate habitat suitability scores (niche suitability scores) (Warren et al., 2011; Warren et al., 2010). These scores are then used to cluster the species occurrence locations into distinct agroecologies with SDM models (chapters 2 and 3).

These SDM-classified agroecologies can be incorporated into phenotypic distribution Models (PDMs) via other machine learning based models like Boosted GLMs and boosted GAMs to predict the relative contribution of selected environmental predictors to phenotypic variation of traits of interest (e.g., live body weight, egg production, survival). The use PDMs is advantageous in multi-environment genetic evaluations because of their ability to combine biological information (e.g., breed, age, sex, genetic variability) with environmental information (e.g., agroecological class, important climatic predictors influencing performance). PDMs can also show complex relationship between a trait and environmental predictor which are difficult to grasp by conventional linear models. Such relationships can be visualized by partial dependence plots.

Our study demonstrates that the model fit of GxE predictions in PDMs improves when agroecologies are fitted based on SDM-identified environmental predictors compared to those fitted on conventional agroecologies (i.e., defined on predictors influencing suitability of the environment for crops). Theoretically, predictions of livestock performance with distribution models can be done both for present and

future production scenarios. Future model predictions of performance require the use of climatic projections from the coming decades which are readily available on global databases. Such predictions are useful to forecast how much the performance of livestock in smallholder systems will be affected by climate change and to take mitigation measures in time. Depending on geography, undesirable consequences of climate change might include reduced feed and water availability, high disease and parasitic prevalence, and extremes of climate (e.g., high solar radiation and high temperature) having negative impacts on productivity.

6.7.2 Yield stability is an important parameter to be considered in breed performance evaluation for smallholders

Apart from evaluation of the levels of productivity for traits of interest, comparison of breed performance for wider introduction into smallholder systems should consider yield stability. Smallholder livestock keepers in extensive (low- and medium-input) systems of Africa are prone to be affected by changes in production conditions and traits such as yield stability and resilience are of paramount importance.

Combining productivity with yield stability improves economic efficiency and contributes to sustainable adoption of genetic technologies (improved breeds). Phenotypic plasticity or trait plasticity is the ability of an organism to alter its physiology, morphology, behaviour, or development in response to changes in its environment (Debat and David, 2001). A GxE refers to a change in relative performance of two or more breeds measured in two or more environments (Bowman, 1972). Animals that combine high production potential with resilience to external stressors, allowing for unproblematic expression of high production potential and less phenotypic plasticity in a wide variety of environmental conditions are regarded as 'robust' (Knap and Su, 2008). Their finding implies that robust animals do not need to be plastic to cope up with environmental variations. Plasticity has a genetic basis and needs to be exploited by animal breeders. Discovery of loci that influence plasticity implies that animals vary in their plastic responses to environmental challenges. While some genes control both plasticity and mean phenotype, many others are associated only with plasticity (Sieriebriennikov et al., 2018) (Lafuente and Beldade, 2019).

Quantifying phenotypic and adaptive genetic variation among indigenous chickens informs long-term genetic improvement programmes. The genetics of phenotypic plasticity (yield stability) supported by evidence of population differentiation along

environmental gradients can be incorporated in breeding programs to develop animals suitable for varying agroecologies, or that are adapted to both current and future environmental conditions.

Studies have not been carried out on yield stability under smallholder management conditions in livestock. There is a need to conduct more studies within and between breeds, and this brings about a need for developing appropriate analytical frameworks to make statistical comparisons. Integrating stability in breeding programs for smallholder management systems might be problematic because of the confounding effects of environment and genotype. Under such conditions, careful considerations of on-farm experimental design such as sample size, randomization, blocking, and replication are vital to ensure sound statistical inferences. In chapter 4, we carried out our measurements on 21,562 female chickens of five chicken breeds in randomized block design (RBD) and paid attention to these important statistical aspects (sample size, randomization, heterogeneity of households in terms of their management of animals, and replication).

There were some limitations regarding the study we carried out on yield stability in this thesis that could be addressed to get better results in the future. There was a lack of performance data on other economically traits besides body weight (e.g., egg productivity, survival, resilience). In addition, there was no pedigree and genomic information on the animals to fit different breeding models and select individuals or populations on stability. Models applicable to plasticity in evolutionary biology and animal breeding are reviewed by De Jong and Bijma (2002) and can be used in selective breeding to improve plasticity within breed. Developing new methodologies of multi-environment breed evaluation should consider additional sources of information besides environmental, phenotypic, and genetic data. Evidence obtained through socio-economic analyses on consumers' preferences, farmers' trait and breed preferences (likability), benefits of improved breeds to nutrition of vulnerable groups in the household (children and lactating women), and contributions to household economic gain will lead to more informed policy decisions regarding the choice of suitable of breeds and species across geographies and environments.

Commercial breeding companies will also benefit from developing productive and stable breeds which can provide optimal yield across agroecologies. Changing livestock production conditions (e.g., climate change, emerging diseases, management decisions, animal welfare standards) force commercial farms to revise

their breeding goals and incorporate additional parameters besides productivity (e.g., resilience, disease tolerance). Resilient animals are minimally affected by disturbances in the environment and return to their state before the disturbance (Berghof et al., 2019; Colditz and Hine, 2016). The link between yield stability and resilience has not been clearly established in livestock.

6.7.3 Improved tropical breeds enhance livestock productivity for smallholders

It is necessary to take advantage of steady improvements in input and service delivery in smallholder systems to increase chicken productivity, meet the growing market demands for meat and eggs, improve livelihoods, and boost economic profits of farmers and other actors in the value chain . Sustained adoption of tropically adapted improved breeds by smallholders, enhances productivity, and leads to better nutritional outcomes in farming households (Birhanu et al., 2022; Passarelli et al., 2020).

Better purchasing power of the rising middle class and rapid urbanization are contributing to the emergence of new business opportunities for poultry industries in Africa. Recent trends show multi-national poultry breeding companies are expanding their markets to this region (Carmody, 2017). These multi-national companies already have the highest capacity to deliver specialized chicken breeds (broilers and layers) suitable for intensive commercial farms. Specialized chickens are not suitable to the needs of smallhoder farmers in Africa considering their high management demands. The breeding companies need to invest in their research and development programmes to meet the growing demands for adaptive and productive dual-purpose chickens suitable for smallholder systems in the tropics. The number of such breeds developed so far is too few to provide choices to farmers and consumers in different geographies and environments.

Breeding programmes in large poultry companies conventionally put most emphasis on productivity traits such as growth rate and egg laying performance and less on resilience and yield stability (Neeteson-van Nieuwenhoven et al., 2013). Attention should also be paid to exploit existing genetic variation in indigenous chickens for adaptive traits because indigenous chickens serve as a reservoir of genes linked with local adaptation, product quality and consumer preference (e.g., coat colour type and pattern, taste of meat and eggs, meat, and egg quality). These adaptive and preference genes can be utilized through tailored breeding programmes or biotechnology.

Tropically adapted and productive chickens targeting the needs of tropical smallholder farmers can be of different genetic backgrounds: composites, commercial hybrids, or improved pure-bred populations. The advantages and disadvantages of each of these types of chickens for smallholders has not been comprehensively documented to inform policy decisions. Composites, hybrids, or pure-bred populations do differ in their utility for tropical smallholder farmers. Composite constitutes a population made up of two or more component breeds, designed to retain heterosis (hybrid vigour) in future generations without the need for crossbreeding, and is maintained like a purebred. A composite breed combining relative strengths of different breeds can be created to meet the needs of specific natural and production environments. Composite breeds like Koekoek produce eggs that can be hatched by local hens to sustain themselves. This self-sustaining quality of flocks is important at least in areas where farmers do not have access to artificial incubators, or delivery schemes for day-old-chicks and fertile eggs.

Commercial hybrids are developed by crossing (e.g., through two-way, three-way, or four-way crosses) of different chicken breeds or lines to produce chickens specialized for fast growth (broilers), high egg production (layers), or optimal yields of growth and eggs (dual-purpose). More information on commercial hybrids potentially suitable to African smallholder systems (e.g., Shika-Brown, Kuroiler, Sasso, Noiler) can be found in the following sources (Adeyinka, 1998; Ajayi et al., 2020; Isenberg, 2007; Kabir and Muhammad, 2012). Designing breeding programmes aiming at developing high producing but adapted hybrid chickens is feasible to enhance productivity and economic gains in areas with better access to markets, inputs, and services. Traits which are well regarded by smallholder farmers but are not included in conventional breeding programmes should not be overlooked (e.g., resilience, yield stability; ability to escape from predators because of non-uniform coat colour serving as a camouflage or alertness; scavenging ability in the free range; and tolerance to diseases and parasites).

The Improved Horro is the only representative of an improved local chicken breed, developed through a selective breeding programme in Ethiopia. The Horro selective breeding program was started in Ethiopia in the year 2004 to improve survival and productivity (growth and eggs) using mass selection. Both sexes were selected for growth-at-16-weeks-of-age while egg-number-at-45-weeks was the criterion for females. Breeding goals were identified in consultation with local farmers and

comprised age-at-first-egg, egg production, body weight, and survival (Dana et al., 2011; Esatu, 2015).

Selective breeding on indigenous chickens requires more resources than breed introduction, but once genetic progress is made it holds many advantages for smallholder farmers compared to commercial hybrids. To mention but a few 1) the chickens can be used by farmers without intellectual property right (IP) restrictions; 2) the chickens of such breeds can be used to hatch their own eggs in areas with no access to electricity to run artificial incubators; 3) they have well established consumer and producer preference or niche market for their appearances, product quality and sociocultural values; 4) genetic gains are permanent and such breeds can be used in cross-breeding programmes with introduced or other local breeds; 5) they have certain adaptive qualities making them suitable for smallholder systems (e.g., alertness, scavenging ability, and adaptability to harsh environmental conditions); and 6) the development of such breeding programs promotes sustainable utilization and conservation of local genetic resources.

Considering the dire need to address nutritional security and to unleash the potentials of the poultry sector to spur economic growth, different types of breeding strategies or programmes (e.g., use of selected pure-breeds, composites, hybrids) can be implemented in African countries in tandem but strategically. The choice of the genetic intervention depends on the socio-economic and bio-physical context. For instance, farmers close to urban and peri-urban areas have better access to markets, inputs and services and will benefit if they keep hybrids. Pure-breeds and composites are more suitable to be used in areas far from major towns.

Genetic improvement schemes on local chickens, like the Horro breeding programme, are promising and lead to the development of productive yet adaptive breeds suitable to the needs of smallholder farmers in Ethiopia. But breeding programmes have a long gestation period to come to fruition. It is also essential to learn from success stories and failures of chicken breeding programmes in other tropical countries of Africa and Asia (Lwelamira and Kifaro, 2010; Niknafs et al., 2012; Nwagu et al., 2007; Okeno et al., 2013; Singh, 2002; Venkatramaiah et al., 1986).

6.7.4 Genomics to inform the development of adaptive breeds for smallholders
There is a strong need to harness local adaptive variation to improve food security, sustain economic growth, and strengthen resilience of production systems in developing regions of the world. Indigenous livestock have been exposed to selective

pressures for many generations and harbour useful genes conferring local adaptation (e.g., disease resistance, tolerance to extreme climate, feed utilization). On the other hand, improved breeds are selected for specific traits and excel in productive performance under optimal environmental conditions. Advances in genomics, breeding, machine learning, ecological niche modelling, and gene editing have created opportunities to combine the merits in different populations (breeds) and produce genotypes with optimum productivity and resilience. Lack of genomic, phenotypic and environmental information and unavailability of suitable analytical methods had so far prevented integrative approaches that might lead to development of such breeds in a shorter time.

Both commercial and smallholder systems benefit from the discovery and utilization of genes related with local adaptation (e.g., climate change, management levels, diseases and parasites, market preferences). Several studies show that beneficial alleles/variants in indigenous chickens can be introduced into commercial chickens through breeding programmes (e.g., cross-breeding, composite development) (Horst, 1989; Sheng et al., 2013) or genome editing (Khwatenge and Nahashon, 2021; Oishi et al., 2016) (Ballantyne et al., 2021).

In chapter 3, I focused on the use of single nucleotide polymorphisms (SNPs) in my analyses. However, the variation between genomes can cover bigger areas, from a kilobase, or even megabase sized structural variants (SVs), such as deletions, insertions, inversions, and more complex rearrangements. Future studies may consider investigating these differences between populations to further improve our understanding of adaptive phenotypic and genetic variation (e.g., comb shape, feathering pattern, muscle development, immunity, disease, and resilience traits). The results from SV analysis of short-read WGS data (from Illumina platform) can also be complemented by outputs of long-read sequencing (Oxford Nanopore Technology-ONT or Pacific-Bioscience-PacBio). Considering its relatively affordable cost, long-read sequencing is promising for the comprehensive discovery of SVs at population scale (Beyter et al., 2021; Jiang et al., 2020).

Screening of genes associated with environmental and phenotypic variation and their inclusion in genomic selection is useful to improve prediction of estimated breeding values including those traits with low heritability and develop breeds for specific or wider agroecologies (Brito et al., 2021; Ruane and Sonnino, 2007). There are a few studies demonstrating the direct application of genomic information and phenotypic variability to enhance resilience and likability under tropical poultry

production systems. The introduction of the frizzle feather phenotypes into the light Sussex temperate chicken breed and removal of the dominant white phenotype from white leghorn chicken breed to develop socially acceptable breeds by certain communities (Ballantyne et al., 2021) are two examples of the direct application of genomic and phenotypic information to enhance resilience and likability under tropical poultry production systems.

Harnessing other ~omics technologies (e.g. to measure gene expression, epigenetic variation, and proteins) advances our understanding of phenotypic and genetic variation and provides more power to detect genomic regions associated with evolutionary processes such as natural selection, and gene flow (Rajora, 2019). For instance, spatial variation in protein expression among stonefly species (*Plecoptera*) sampled from four geographic regions along a latitudinal gradient was highly positively correlated with water temperature suggesting that regulation of protein expression in different environments relates to local adaptation (Gamboa et al., 2017).

Transcriptomic and metatranscriptomic data such as the one generated from metagenomics (e.g., microbes from chicken gut, faeces) collected from different environmental gradients across space and time may also add to our understanding of adaptive phenotypic and genetic variation in indigenous chickens. Productivity traits (e.g., growth, feed conversion efficiency), nutrient absorption, physiology and fitness (immunity, adaptability) in chickens are influenced by gut microbiota (Choi et al., 2015; Diaz Carrasco et al., 2019; Pedroso and Lee, 2015) and the environment. For instance, Kumar et al. (2020) and Banos et al. (2020) identified antibiotic and disease resistance genes in the microbiomes of Ethiopian chicken populations.

Combined analysis of multiple but independent data types (e.g., geography, environment, behaviour, ecology, physiology, transcriptomics, and genomics) maximizes the power and reliability of delineating adaptively differentiated populations (Allendorf and Luikart, 2009). Improvements in predictive ability of distribution models is achieved when environmental information is used along with genomic information (Gotelli and Stanton-Geddes, 2015; Razgour, 2015).

The different studies described in this section, including marker assisted selection, structural variation, metagenomics, and epigenetics, demonstrate how genomic information can be combined with phenotypic and environmental information to

improve our understanding of adaptive genetic and phenotypic variation and develop breeds suitable for smallholders.

6.7.5 Conservation and sustainable utilization of genetic resources shall never be neglected

Genetic diversity is the raw material for breeders to work with and offers the solutions which are needed to respond to future changes in production conditions. Without conservation and sustainable utilization of domestic animal biodiversity, it would be difficult to mitigate the undesirable effects of climate change, continuously improve genetics, respond to shifts in consumer demands, overcome challenges from emerging diseases and parasites, and alleviate other unforeseen problems.

Most of the indigenous chicken populations in Ethiopia and other tropical developing countries face a risk of extinction from rapid introduction and replacement by exotic chicken strains (composites, commercial hybrids) (Scherf and Pilling, 2015). The urgency of food insecurity in sub-Saharan African countries and the arguments behind introducing more productive strains are plausible. However, the necessity for *in situ* conservation (in their natural habitats) and/or *ex situ* conservation (outside their natural habitats either *in vivo* or *in vitro*) of the indigenous chicken genetic resources should also not be overlooked.

Designing long-term breeding programmes is an avenue for sustainable utilization of chicken genetic resources. To this end, chickens should be sourced from biodiversity hot spot areas of the country such as Guba district in Benishangul-Gumuz region, Horro, Jarso and Girawa districts in Oromia region, Karat district in Southern region, and Guagusa Shekudad in Amhara region (Dessie, 2003a; Desta, 2015; Getachew et al., 2016). Information on neutral and adaptive variation of individuals and populations from landscape genomic studies is applicable to making conservation priority decisions at national level. Policy makers, commercial breeders, researchers, and development practitioners should also consider taking different conservation management steps such as restricting the introduction of exotic breeds into selected biodiversity hot spot areas, improving the awareness of producers and consumers on the virtues of local genetic resources, and developing niche markets for products from indigenous chickens at premium price. Apart from this, establishment of experimental flocks at poultry research facilities can be used to, in the short-run, conserve populations (individuals) with unique phenotypes *ex-situ* and serve as experimental flocks for in depth studies (e.g., molecular, physiological, immunological, behavioural, and phenotypic).

Chicken genetic resources conserved *in situ* as live populations are at risk of loss due to disease outbreaks, genetic disorders, inbreeding, natural and human-made disasters. Hence *ex-situ* conservation through cryobanking of semen and primordial germ cells (PGCs) should be considered. Conservation of PGCs has an added advantage because it allows long-term storage of precursor cells for the gametes of both sexes. The technologies for avian PGC manipulation and cryoconservation are constantly being refined and seem highly promising (Mucksová et al., 2019; Nakamura, 2016; Nandi et al., 2016).

6.7.6 Concluding remarks

Smallholder chicken production can serve as a way out of poverty for millions of farmers in developing regions of the world. However, improving levels of productivity and production require more than availing the best performing breeds to smallholder farmers. Increasing the production efficiency of the smallholder poultry farmers in tropical developing countries requires careful examination of other interrelated factors besides delivering improved genetics. This entails a comprehensive analysis of the poultry-value-chain and the production system (e.g., reproduction strategy, germplasm sourcing, germplasm dissemination, marketing, health, feeds, management decisions, agroecology/environment, quality grading of poultry products and inputs, policy matters, and legal aspects governing intellectual property rights and product development). Economic, social, technical, and ecological viability; and access to efficient input and output markets are vital for sustained adoption of agricultural technologies such improved chicken breeds by smallholder farmers (Birhanu et al., 2022).

Tropical poultry production systems are complex. Phenotypic and genetic variation exhibited by indigenous chicken populations in these systems are the result of an interplay between genetic and environmental factors (e.g., climate, geography, management decisions). Phenotypic and genetic variation observed in indigenous and improved chicken populations reared in low- and medium-input systems offer opportunities to improve the economic efficiency of smallholder and commercial systems. Changes in production scenarios (e.g., climate, diseases and parasites, consumer preferences) require development of breeds that combine superior productive performance with traits such as robustness, resilience, and likability.

In this PhD thesis, I studied adaptive phenotypic and genetic variation among indigenous and improved breeds of chickens raised in smallholder systems. The rationales for undertaking this PhD research were the unavailability of analytical

approaches that integrate genomic, phenotypic, and environmental information to investigate adaptive phenotypic and genetic variation. Proper methodologies need to be developed to match the right breeds with the right environment and to enhance productivity. I applied advances in ecology, statistical genetics, and landscape genomics to identify the most important climatic predictors influencing chicken productive performance and adaptability. I used this information to redefine agroecologies for the species, to improve model fit in GxE predictions for improved breeds, and compare breeds on growth performance and yield stability. I also used the most informative environmental predictors and quantitative morphological traits to find associated variability in the chicken genome. The approaches applied in this thesis can be used to study adaptive phenotypic and genetic variation in other livestock species. Exciting opportunities lie ahead to exploit adaptive phenotypic and genetic variation in chickens to enhance productivity and resilience and improve economic efficiency in smallholder systems.

References

Adebabay, K. (2018). Whole genome based characterization of indigenous chicken populations in Ethiopia. Dissertation, Addis Ababa University, Addis Ababa.

Adeyinka, I. (1998). Short term response to selection in layer type chickens, Ph. D thesis submitted to the department of Animal Science, Ahmadu Bello ….

Adomako, K., Olympio, O., Hagan, J., and Hamidu, J. (2014). Effect of the frizzle gene (F) on egg production and egg quality of laying hens kept in tropical villages. *British poultry science* **55**, 709-714.

Ahlgren, J., Yang, X., Hansson, L.-A., and Brönmark, C. (2013). Camouflaged or tanned: plasticity in freshwater snail pigmentation. *Biology letters* **9**, 20130464.

Ahuja, V., Dhawan, M., Punjabi, M., and Maarse, L. (2008a). Chickening out of poverty? Story of 'kuroiler'from India. *In* "International Conference".

Ahuja, V., Dhawan, M., Punjabi, M., and Maarse, L. (2008b). Poultry based livelihoods of rural poor: Case of Kuroiler in West Bengal. *South Asia Pro Poor Livestock Policy Programme, initiative of NDDB and FAO: Document* **12**.

Ajayi, F. O., Bamidele, O., Hassan, W. A., Ogundu, U., Yakubu, A., Alabi, O. O., Akinsola, O. M., Sonaiya, E. B., and Adebambo, O. A. (2020). Production performance and survivability of six dual-purpose breeds of chicken under smallholder farmers' management practices in Nigeria. *Archives Animal Breeding* **63**, 387-408.

Alabi, O., Ajayi, F., Bamidele, O., Yakubu, A., Ogundu, E., Sonaiya, E., Ojo, M., Hassan, W., and Adebambo, O. (2020). Impact assessment of improved chicken genetics on livelihoods and food security of smallholder poultry farmers in Nigeria. *Livestock Research for Rural Development* **32**, 77.

Alemayehu, K., and Getu, A. (2016). Impacts of climate variability on livestock population dynamics and breed distribution patterns in selected districts of Western Amhara, Ethiopia. *Animal Genetic Resources/Resources génétiques animales/Recursos genéticos animales* **59**, 113-121.

Alemu, S. W., Hanotte, O., Kebede, F. G., Esatu, W., Abegaz, S., Bruno, J. E., Abrar, B., Alemayehu, T., Mrode, R., and Dessie, T. (2021). Evaluation of live-body weight and the number of eggs produced for introduced and local chickens in Ethiopia. *Acta Agriculturae Scandinavica, Section A—Animal Science* **70**, 71-77.

Allendorf, F. W., and Luikart, G. (2009). "Conservation and the genetics of populations," John Wiley & Sons.

Apley, D. W., and Zhu, J. (2020). Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **82**, 1059-1086.

Atkins, K., and Travis, J. (2010). Local adaptation and the evolution of species' ranges under climate change. *Journal of Theoretical Biology* **266**, 449-457.

Auld, J. R., and Relyea, R. A. (2010). Inbreeding depression in adaptive plasticity under predation risk in a freshwater snail. *Biology Letters* **6**, 222-224.

Austin, M. (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological modelling* **157**, 101-118.

Authority, E. M. (1988). National atlas of Ethiopia.

Azevedo Peixoto, L. d., Teodoro, P. E., Silva, L. A., Rodrigues, E. V., Laviola, B. G., and Bhering, L. L. (2018). Jatropha half-sib family selection with high adaptability and genotypic stability. *PloS one* **13**, e0199880.

Bagath, M., Krishnan, G., Devaraj, C., Rashamol, V., Pragna, P., Lees, A., and Sejian, V. (2019). The impact of heat stress on the immune system in dairy cattle: a review. *Research in veterinary science* **126**, 94-102.

Balestre, M., Von Pinho, R., Souza, J., and Oliveira, R. (2009). Genotypic stability and adaptability in tropical maize based on AMMI and GGE biplot analysis. *Genetics and Molecular Research* **8**, 1311-1322.

Balkenhol, N., Cushman, S. A., Waits, L. P., and Storfer, A. (2016). Current status, future opportunities, and remaining challenges in landscape genetics [Chapter 14]. *In: Balkenhol, Niko; Cushman, Samuel A.; Storfer, Andrew T.; Waits, Lisette P., eds. Landscape Genetics: Concepts, Methods, Applications, First Edition. John Wiley and Sons Ltd. p. 247-255.*, 247-255.

Ballantyne, M., Woodcock, M., Doddamani, D., Hu, T., Taylor, L., Hawken, R. J., and McGrew, M. J. (2021). Direct allele introgression into pure chicken breeds using Sire Dam Surrogate (SDS) mating. *Nature communications* **12**, 1-10.

Banos, G., Lindsay, V., Desta, T. T., Bettridge, J., Sanchez-Molano, E., Vallejo-Trujillo, A., Matika, O., Dessie, T., Wigley, P., and Christley, R. M. (2020). Integrating genetic and genomic analyses of combined health data across ecotypes to improve disease resistance in indigenous African chickens. *Frontiers in genetics* **11**.

Barhdadi, A., and Dubé, M. P. (2010). Testing for gene-gene interaction with AMMI models. *Statistical applications in genetics and molecular biology* **9**, Article 2.

Bates, D., Sarkar, D., Bates, M. D., and Matrix, L. (2007). The lme4 package. *R package version* **2**, 74.

Bayou, T., and Assefa, A. (1989). Solar radiation maps fot Ethiopia. *Zede Journal* **8**, 7-15.

Becker, H. (1981). Correlations among some statistical measures of phenotypic stability. *Euphytica* **30**, 835-840.

Becker, H., and Leon, J. (1988). Stability analysis in plant breeding. *Plant breeding* **101**, 1-23.

Berghof, T. V., Poppe, M., and Mulder, H. A. (2019). Opportunities to improve resilience in animal breeding programs. *Frontiers in genetics* **9**, 692.

Bergland, A., Behrman, E., O'Brien, K., Schmidt, P., and Petrov, D. (2014). Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in. Drosophila.

Bergmann, C. (1848). "Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse."

Bettridge, J. M., Psifidi, A., Terfa, Z. G., Desta, T. T., Lozano-Jaramillo, M., Dessie, T., Kaiser, P., Wigley, P., Hanotte, O., and Christley, R. M. (2018). The role of local adaptation in sustainable production of village chickens. *Nature Sustainability* **1**, 574-582.

Beyter, D., Ingimundardottir, H., Oddsson, A., Eggertsson, H. P., Bjornsson, E., Jonsson, H., Atlason, B. A., Kristmundsdottir, S., Mehringer, S., and Hardarson, M. T. (2021). Long-read sequencing of 3,622 Icelanders provides insight into the role of structural variants in human diseases and other traits. *Nature Genetics* **53**, 779-786.

Billi, P. (2015). "Landscapes and landforms of Ethiopia," Springer.

Birhanu, M. Y., Alemayehu, T., Bruno, J. E., Kebede, F. G., Sonaiya, E. B., Goromela, E. H., Bamidele, O., and Dessie, T. (2021). Technical Efficiency of Traditional Village Chicken Production in Africa: Entry Points for Sustainable Transformation and Improved Livelihood. *Sustainability* **13**, 8539.

Birhanu, M. Y., Bruno, J. E., Alemayehu, T., Esatu, W., Geremew, K., Yemane, T., Kebede, F. G., and Dessie, T. (2022). Beyond diffusion to sustained adoption of innovation: A case of smallholder poultry development in sub-Saharan Africa. *International Journal of Agricultural Sustainability*, 1-19.

Bivand, R., Keitt, T., Rowlingson, B., Pebesma, E., Sumner, M., Hijmans, R., Rouault, E., and Bivand, M. R. (2021a). Package 'rgdal'. *In* "Bindings for the Geospatial Data Abstraction Library".

Bivand, R., Lewin-Koh, N., Pebesma, E., Archer, E., Baddeley, A., Bearman, N., Bibiko, H.-J., Brey, S., Callahan, J., and Carrillo, G. (2021b). Package 'maptools'.

Bivand, R., Rundel, C., Pebesma, E., Stuetz, R., Hufthammer, K. O., and Bivand, M. R. (2017). Package 'rgeos'. *The Comprehensive R Archive Network (CRAN)*.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120.

Bolker, B. M., Gardner, B., Maunder, M., Berg, C. W., Brooks, M., Comita, L., Crone, E., Cubaynes, S., Davies, T., and de Valpine, P. (2013). Strategies for fitting nonlinear ecological models in R, AD M odel B uilder, and BUGS. *Methods in Ecology and Evolution* **4**, 501-512.

Bolnick, D. I., Svanbäck, R., Fordyce, J. A., Yang, L. H., Davis, J. M., Hulsey, C. D., and Forister, M. L. (2003). The ecology of individuals: incidence and implications of individual specialization. *The American Naturalist* **161**, 1-28.

Borcard, D., Gillet, F., and Legendre, P. (2011). "Numerical ecology with R," Springer.

Bortoluzzi, C., Megens, H.-J., Bosse, M., Derks, M. F., Dibbits, B., Laport, K., Weigend, S., Groenen, M. A., and Crooijmans, R. P. (2020). Parallel genetic origin of foot feathering in birds. *Molecular biology and evolution* **37**, 2465-2476.

Bowman, J. (1972). Genotype× environment interactions. *In* "Annales de génétique et de sélection animale", Vol. 4, pp. 1-7. BioMed Central.

Bradshaw, A. D. (1965). Evolutionary significance of phenotypic plasticity in plants. *In* "Advances in genetics", Vol. 13, pp. 115-155. Elsevier.

Brito, L., Bédère, N., Douhard, F., Oliveira, H., Arnal, M., Peñagaricano, F., Schinckel, A., Baes, C. F., and Miglior, F. (2021). Genetic selection of high-yielding dairy cattle toward sustainable farming systems in a rapidly changing world. *animal*, 100292.

Brown-Brandl, T. (2013). Managing thermal stress in feedlot cattle: environment, animal susceptibility and management options from a US perspective. *Livestock housing: Modern management to ensure optimal health and welfare of farm animals*, 189.

Brun, P., Thuiller, W., Chauvier, Y., Pellissier, L., Wüest, R. O., Wang, Z., and Zimmermann, N. E. (2019). Model complexity affects species distribution projections under climate change. *Journal of Biogeography*.

Cain, A. J., and Sheppard, P. M. (1954). Natural selection in Cepaea. *Genetics* **39**, 89.

Capblancq, T., Luu, K., Blum, M. G., and Bazin, E. (2018). Evaluation of redundancy analysis to identify signatures of local adaptation. *Molecular Ecology Resources* **18**, 1223-1233.

Carmody, P. (2017). "The new scramble for Africa," John Wiley & Sons.

Carvalho, C. S., Forester, B. R., Mitre, S. K., Alves, R., Imperatriz-Fonseca, V. L., Ramos, S. J., Resende-Moreira, L. C., Siqueira, J. O., Trevelin, L. C., and Caldeira, C. F. (2021). Combining genotype, phenotype, and environmental data to delineate site-adjusted provenance strategies for ecological restoration. *Molecular Ecology Resources* **21**, 44-58.

Cesconeto, R. J., Joost, S. p., McManus, C. M., Paiva, S. R., Cobuci, J. A., and Braccini, J. (2017). Landscape genomic approach to detect selection signatures in locally adapted Brazilian swine genetic groups. *Ecology and Evolution* **7**, 9544-9556.

Cha, J., Choo, H., Srikanth, K., Lee, S.-H., Son, J.-W., Park, M.-R., Kim, N., Jang, G. W., and Park, J.-E. (2021). Genome-Wide Association Study Identifies 12 Loci Associated with Body Weight at Age 8 Weeks in Korean Native Chickens. *Genes* **12**, 1170.

Chambers, J., and Hastie, T. (1992). Linear models. Chapter 4 of statistical models in S. *Wadsworth & Brooks/Cole*.

Chen, S.-Y., Freitas, P. H., Oliveira, H. R., Lázaro, S. F., Huang, Y. J., Howard, J. T., Gu, Y., Schinckel, A. P., and Brito, L. F. (2021). Genotype-by-environment

interactions for reproduction, body composition, and growth traits in maternal-line pigs based on single-step genomic reaction norms. *Genetics Selection Evolution* **53**, 1-18.

Cheruiyot, E., Nguyen, T., Haile-Mariam, M., Cocks, B., Abdelsayed, M., and Pryce, J. (2020). Genotype-by-environment (temperature-humidity) interaction of milk production traits in Australian Holstein cattle. *Journal of Dairy Science* **103**, 2460-2476.

Choi, K. Y., Lee, T. K., and Sul, W. J. (2015). Metagenomic analysis of chicken gut microbiota for improving metabolism and health of chickens—a review. *Asian-Australasian journal of animal sciences* **28**, 1217.

Chu, T. T., Madsen, P., Norberg, E., Wang, L., Marois, D., Henshall, J., and Jensen, J. (2020). Genetic analysis on body weight at different ages in broiler chicken raised in commercial environment. *J Anim Breed Genet* **137**, 245-259.

Clausen, J., Keck, D. D., and Hiesey, W. M. (1940). Experimental studies on the nature of species. I. Effect of varied environments on western North American plants. *Experimental studies on the nature of species. I. Effect of varied environments on western North American plants.*

Colautti, R. I., Alexander, J. M., Dlugosch, K. M., Keller, S. R., and Sultan, S. E. (2017). Invasions and extinctions through the looking glass of evolutionary ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**, 20160031.

Colditz, I. G., and Hine, B. C. (2016). Resilience in farm animals: biology, management, breeding and implications for animal welfare. *Animal Production Science* **56**, 1961-1983.

Colombari Filho, J. M., de Resende, M. D. V., de Morais, O. P., de Castro, A. P., Guimaraes, E. P., Pereira, J. A., Utumi, M. M., and Breseghello, F. (2013). Upland rice breeding in Brazil: a simultaneous genotypic evaluation of stability, adaptability and grain yield. *Euphytica* **192**, 117-129.

Conover, D. O., Duffy, T. A., and Hice, L. A. (2009). The covariance between genetic and environmental influences across ecological gradients. *Annals of the New York Academy of Sciences* **1168**, 100-129.

Costa, J. T. (2021). There Is Hardly Any Question in Biology of More Importance: Charles Darwin and the Nature of Variation. *In* "Phenotypic Plasticity & Evolution", pp. 25-54. CRC Press.

Crawley, M. J. (2012). "The R book," John Wiley & Sons.

Crispo, E. (2008). Modifying effects of phenotypic plasticity on interactions among natural selection, adaptation and gene flow. *Journal of evolutionary biology* **21**, 1460-1469.

CSA (2017). "Livestock Characteristics, Agricultural sample Survey. Addis Ababa, Ethiopia." Central Staistical Authority, Addis Ababa.

CSA (2020). Agricultural Sample Survey 2019/20 [2012 E.C.]. Volume II report on livestock and livestock

characteristics (private peasant holdings). (C. S. A. (CSA), ed.), Addis Ababa, Ethiopia.

Culman, S. W., Bukowski, R., Gauch, H. G., Cadillo-Quiroz, H., and Buckley, D. H. (2009). T-REX: software for the processing and analysis of T-RFLP data. *BMC bioinformatics* **10**, 171.

Dana, N. (2011). Breeding programs for indigenous chicken in Ethiopia analysis of diversity in production systems and chicken populations, Wageningen University.

Dana, N., Van der Waaij, L. H., Dessie, T., and van Arendonk, J. A. (2010). 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* **42**, 1519-1529.

Dana, N., Vander Waaij, E., and van Arendonk, J. A. (2011). Genetic and phenotypic parameter estimates for body weights and egg production in Horro chicken of Ethiopia. *Tropical animal health and production* **43**, 21-28.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., and Sherry, S. T. (2011). The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158.

Darwin, C., Wallace, A. R., Lyell, S. C., and Hooker, J. D. (1858). On the tendency of species to form varieties: and on the perpetuation of varieties and species by natural means of selection. Linnean Society of London.

De Jong, G., and Bijma, P. (2002). Selection and phenotypic plasticity in evolutionary biology and animal breeding. *Livestock production science* **78**, 195-214.

de Kinderen, M. A., Sölkner, J., Mészáros, G., Alemu, S. W., Esatu, W., Bastiaansen, J. W., Komen, H., and Dessie, T. (2020). Genotype by environment interaction for egg number and egg weight of five dual-purpose chicken breeds in different zones of Oromia region in Ethiopia. *Acta Fytotechnica et Zootechnica* **23**.

de Mendiburu, F. (2021). agricolae tutorial (Version 1.3-5). *Universidad Nacional Agraria: La Molina, Peru*.

De Mita, S., Thuillet, A. C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., and Vigouroux, Y. (2013). Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular ecology* **22**, 1383-1399.

Debat, V., and David, P. (2001). Mapping phenotypes: canalization, plasticity and developmental stability. *Trends in ecology & evolution* **16**, 555-561.

Decker, B. L. (1986). World geodetic system 1984.

Delacour, J. (1977). The Pheasant of the World Spurr Publication. *Saiga Publising Co Ltd Surr GU* **26**.

Deneke Negassa, A. M., and Sandip Banerje (2014). <Phenotypic characterization of indigenous chicken

populations in Southeastern Oromia Regional State

of Ethiopia.pdf>. *Animal Genetic Resources* **55**, 101-113.

Deressa, T. T., Ringler, C., and Hassan, R. M. (2010). Factors affecting the choices of coping strategies for climate extremes. *The case of farmers in the Nile Basin of Ethiopia IFPRI Discussion Paper* **1032**.

Desiere, S., Hung, Y., Verbeke, W., and D'Haese, M. (2018). Assessing current and future meat and fish consumption in Sub-Sahara Africa: Learnings from FAO Food Balance Sheets and LSMS household survey data. *Global food security* **16**, 116-126.

Dessie, T. (2003a). Phenotypic and genetic characterization of local chicken ecotypes in Ethiopia, Humboldt-Universität zu Berlin.

Dessie, T. (2003b). Phenotypic and genetic characterization of local chicken ecotypes in Ethiopia (Doctoral dissertation), Humboldt-Universität zu Berlin.

Desta, T. T. (2015). Phenomic and genomic landscape of Ethiopian village chickens, University of Nottingham.

Desta, T. T. (2021). The genetic basis and robustness of naked neck mutation in chicken. *Tropical Animal Health and Production* **53**, 1-13.

Dias, P. C., Xavier, A., Resende, M. D. V. d., Barbosa, M. H. P., Biernaski, F. A., and Estopa, R. A. (2018). Genetic evaluation of Pinus taeda clones from somatic embryogenesis and their genotype x environment interaction. *Crop Breeding and Applied Biotechnology* **18**, 55-64.

Diaz Carrasco, J. M., Casanova, N. A., and Fernández Miyakawa, M. E. (2019). Microbiota, gut health and chicken productivity: what is the connection? *Microorganisms* **7**, 374.

Dinka, M. O. (2019). Development and Application of Conceptual Rainfall-Altitude Regression Model: The Case of Matahara Area (Ethiopia). *In* "Topics in Hydrometerology". IntechOpen.

Diouf, I., Derivot, L., Koussevitzky, S., Carretero, Y., Bitton, F., Moreau, L., and Causse, M. (2020). Genetic basis of phenotypic plasticity and genotype× environment interactions in a multi-parental tomato population. *Journal of experimental botany* **71**, 5365-5376.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., and Leitão, P. J. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 27-46.

Dove, K. (1890). "Kulturzonen von Nord-Abessinien," J. Perthes.

Duijvesteijn, N., and Perrault, L. (2019). 180 How dual-purpose chickens can help African smallholder farmers. *Journal of Animal Science* **97**, 182.

Dumont, B., Fortun-Lamothe, L., Jouven, M., Thomas, M., and Tichit, M. (2013). Prospects from agroecology and industrial ecology for animal production in the 21st century. *Animal* **7**, 1028-1043.

Dumont, B., González-García, E., Thomas, M., Fortun-Lamothe, L., Ducrot, C., Dourmad, J.-Y., and Tichit, M. (2014). Forty research issues for the redesign of animal production systems in the 21st century. *Animal* **8**, 1382-1393.

Dunn, L. (1925). The inheritance of rumplessness in the domestic fowl. *Journal of Heredity* **16**, 127-134.

Eberhart, S. t., and Russell, W. (1966). Stability parameters for comparing varieties 1. *Crop sci* **6**, 36-40.

Elbeltagy, A. R., Bertolini, F., Fleming, D. S., Van Goor, A., Ashwell, C. M., Schmidt, C. J., Kugonza, D. R., Lamont, S. J., and Rothschild, M. F. (2019). Natural Selection Footprints among African Chicken Breeds and Village Ecotypes. *Frontiers in Genetics* **10**, 376.

Elith, J., and Franklin, J. (2013). Species distribution modeling. *In* "Encyclopedia of Biodiversity: Second Edition", pp. 692-705. Elsevier Inc.

Elith, J., and Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual review of ecology, evolution, and systematics* **40**, 677-697.

Elith, J., Leathwick, J. R., and Hastie, T. (2008). A working guide to boosted regression trees. *Journal of animal ecology* **77**, 802-813.

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., and Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and distributions* **17**, 43-57.

ENTAG (2020). "Invest in the Ethiopian Poultry Sector,

Business Opportunity Report."

Esatu, W. (2015). Genetic improvement in indigenous chicken of Ethiopia, Wageningen University.

Etterson, J. R., and Shaw, R. G. (2001). Constraint to adaptive evolution in response to global warming. *Science* **294**, 151-154.

Falconer, D. S. (1996). "Introduction to quantitative genetics," Pearson Education India.

FAO (2012). "Phenotypic characterization of animal genetic resources," Rome.

FAO (2019a). "Poultry Sector Ethiopia," Rome.

FAO (2019b). "Poultry Sector in Ethiopia," Rome.

FAO, F. (2020). Gateway to poultry production and products. *2020. Food an d agriculture organisation of the united nations*.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., and Roth, L. (2007). The shuttle radar topography mission. *Rev Geophys* **45**.

Fasil Getachew, A. A., Tesfaye Getachew, Solomon Abegaz Kebede, Abebe Hailu, Manaye Mesganaw, Yibrehu Emishaw, Misikire Tessema (2020). On-farm phenotypic characterization of Begaria cattle population and their production system in Guba district, North Western Ethiopia *Eth. J. Anim. Prod.* **20**, 1-17.

Fasiolo, M., Nedellec, R., Goude, Y., and Wood, S. N. (2020). Scalable visualization methods for modern generalized additive models. *Journal of computational and Graphical Statistics* **29**, 78-86.

Fazzini, M., Bisci, C., and Billi, P. (2015). The climate of Ethiopia. *In* "Landscapes and landforms of Ethiopia", pp. 65-87. Springer.

Ferraudo, G. M., and Perecin, D. (2014). Mixed model, AMMI and Eberhart-Russel comparison via simulation on genotype× environment interaction study in sugarcane. *Applied Mathematics* **2014**.

Fick, S. E., and Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* **37**, 4302-4315.

Finlay, K., and Wilkinson, G. (1963). The analysis of adaptation in a plant-breeding programme. *Australian journal of agricultural research* **14**, 742-754.

Fitzpatrick, M. C., and Keller, S. R. (2015). Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. *Ecology letters* **18**, 1-16.

Fleming, D. S., Weigend, S., Simianer, H., Weigend, A., Rothschild, M., Schmidt, C., Ashwell, C., Persia, M., Reecy, J., and Lamont, S. J. (2017). Genomic comparison of indigenous African and Northern European chickens reveals putative mechanisms of stress tolerance related to environmental selection pressure. *G3: Genes, Genomes, Genetics* **7**, 1525-1537.

Forester, B. (2019). Detecting multilocus adaptation using Redundancy Analysis (RDA). *Population Genetics in R. Retrieved from https://popgen. nesce nt. org/2018-03-27_RDA_GEA. html*.

Forester, B. R., Lasky, J. R., Wagner, H. H., and Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Molecular Ecology* **27**, 2215-2233.

Foucault, Q., Wieser, A., Waldvogel, A. M., Feldmeyer, B., and Pfenninger, M. (2018). Rapid adaptation to high temperatures in Chironomus riparius. *Ecology and evolution* **8**, 12780-12789.

Fourie, C., and Grobbelaar, J. (2003). Indigenous poultry breed spp. Krugersdorp. Germany, Wing Nut Publications.

Frankham, R., Ballou, S. E. J. D., Briscoe, D. A., and Ballou, J. D. (2002). "Introduction to conservation genetics," Cambridge university press.

Franklin, J. (2010). "Mapping species distributions: spatial inference and prediction," Cambridge University Press.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.

Gamachu, D. (1988). Some Patterns of Altitudinal Variation of Climatic Elements in the Mountainous Regions of Ethiopia. *Mountain Research and Development* **8**, 131-138.

Gamboa, M., Tsuchiya, M. C., Matsumoto, S., Iwata, H., and Watanabe, K. (2017). Differences in protein expression among five species of stream stonefly

(Plecoptera) along a latitudinal gradient in Japan. *Archives of insect biochemistry and physiology* **96**, e21422.

Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:1207.3907*.

Gauch Jr, H. (1992). "Statistical analysis of regional yield trials: AMMI analysis of factorial designs," Elsevier Science Publishers.

Gauch Jr, H. G. (2013). A simple protocol for AMMI analysis of yield trials. *Crop Science* **53**, 1860-1869.

Gauch Jr, H. G., and Zobel, R. W. (1997). Identifying mega-environments and targeting genotypes. *Crop science* **37**, 311-326.

Gebrechorkos, S. H., Hülsmann, S., and Bernhofer, C. (2019). Long-term trends in rainfall and temperature using high-resolution climate datasets in East Africa. *Scientific reports* **9**, 1-9.

Getachew, F., Abegaz, S., Assefa, A., Misganaw, M., Emshaw, Y., Hailu, A., Tessema, M., and Okore, C. (2016). Multivariate analyses of morphological traits in indigenous chicken populations of Metekel zone, Northwestern Ethiopia. *Animal Genetic Resources/Resources génétiques animales/Recursos genéticos animales* **59**, 15-25.

Getahun, A. (1978). Agricultural systems in Ethiopia. *Agric Syst* **3**, 281-293.

Gheyas, A. A., Vallejo-Trujillo, A., Kebede, A., Lozano-Jaramillo, M., Dessie, T., Smith, J., and Hanotte, O. (2021). Integrated Environmental and Genomic Analysis Reveals the Drivers of Local Adaptation in African Indigenous Chickens. *Molecular Biology and Evolution*.

Gicheha, M. G. (2021). The Effects of Heat Stress on Production, Reproduction, Health in Chicken and Its Dietary Amelioration. *In* "Advances in Poultry Nutrition Research". IntechOpen.

Gotelli, N. J., and Stanton-Geddes, J. (2015). Climate change, genetic markers and species distribution modelling. *Journal of Biogeography* **42**, 1577-1585.

Gouveia, J. J. d. S., Silva, M. V. G. B. d., Paiva, S. R., and Oliveira, S. M. P. d. (2014). Identification of selection signatures in livestock species. *Genetics and molecular biology* **37**, 330-342.

Grimes, S. E. (2002). A basic Laboratory Manual for the small-scale production and testing of 1-2 Newcastle disease vaccine.

Grobbelaar, J. (2008). Egg production potentials of four indigenous chicken breeds in South Africa. Pretoria, Tshwane University of Technology. M. Tech, thesis.

Grobbelaar, J. A. N., Sutherland, B., and Molalakgotla, N. (2010). Egg production potentials of certain indigenous chicken breeds from South Africa.

Anim Genet Resour **46**, 25-32.

Guisan, A., Thuiller, W., and Zimmermann, N. E. (2017). "Habitat suitability and distribution models: with applications in R," Cambridge University Press.

Habimana, R., Ngeno, K., Okeno, T. O., Hirwa, C., Keambou Tiambo, C., and Yao, N. K. (2021). Genome-Wide Association Study of Growth Performance and

Immune Response to Newcastle Disease Virus of Indigenous Chicken in Rwanda. *Frontiers in Genetics*, 1438.

Hackinger, S., Kraaijenbrink, T., Xue, Y., Mezzavilla, M., van Driem, G., Jobling, M. A., de Knijff, P., Tyler-Smith, C., and Ayub, Q. (2016). Wide distribution and altitude correlation of an archaic high-altitude-adaptive EPAS1 haplotype in the Himalayas. *Human genetics* **135**, 393-402.

Hällfors, M. H., Liao, J., Dzurisin, J., Grundel, R., Hyvärinen, M., Towle, K., Wu, G. C., and Hellmann, J. J. (2016). Addressing potential local adaptation in species distribution models: implications for conservation under climate change. *Ecological Applications* **26**, 1154-1169.

Hampe, A. (2004). Bioclimate envelope models: what they detect and what they hide. *Glob Ecol Biogeogr* **13**, 469-471.

Harrisson, K. A., Amish, S. J., Pavlova, A., Narum, S. R., Telonis-Scott, M., Rourke, M. L., Lyon, J., Tonkin, Z., Gilligan, D. M., and Ingram, B. A. (2017). Signatures of polygenic adaptation associated with climate across the range of a threatened fish species with high genetic connectivity. *Molecular ecology* **26**, 6253-6269.

Hartung, C., Lerer, A., Anokwa, Y., Tseng, C., Brunette, W., and Borriello, G. (2010). Open data kit: tools to build information services for developing regions. *In* "Proceedings of the 4th ACM/IEEE international conference on information and communication technologies and development", pp. 1-12.

Hassen, H., Neser, F., De Kock, A., and van Marle-Köster, E. (2009). Study on the genetic diversity of native chickens in northwest Ethiopia using microsatellite markers. *African journal of biotechnology* **8**.

Hastie, T., and Tibshirani, R. (1986). Generalized additive models (with discussion), Statistical Science.

Hastie, T., and Tibshirani, R. (1990). "Generalized Additive Models " Taylor & Francis.

Hendry, A. P. (2016). Key questions on the role of phenotypic plasticity in eco-evolutionary dynamics. *Journal of Heredity* **107**, 25-41.

Hengl, T., de Jesus, J. M., Heuvelink, G. B., Gonzalez, M. R., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., and Bauer-Marschallinger, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS one* **12**.

Hengl, T., Heuvelink, G. B., Kempen, B., Leenaars, J. G., Walsh, M. G., Shepherd, K. D., Sila, A., MacMillan, R. A., de Jesus, J. M., and Tamene, L. (2015). Mapping soil properties of Africa at 250 m resolution: Random forests significantly improve current predictions. *PloS one* **10**.

Hijmans, R. J., Guarino, L., Cruz, M., and Rojas, E. (2001). Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genet Resour Newsl*, 15-19.

Hijmans, R. J., Phillips, S., Leathwick, J., Elith, J., and Hijmans, M. R. J. (2017). Package 'dismo'. *Circles* **9**, 1-68.

Hijmans, R. J., Van Etten, J., Cheng, J., Mattiuzzi, M., Sumner, M., Greenberg, J. A., Lamigueiro, O. P., Bevan, A., Racine, E. B., and Shortridge, A. (2015). Package 'raster'. *R package* **734**.

Hohenlohe, P. A., Phillips, P. C., and Cresko, W. A. (2010). Using population genomics to detect selection in natural populations: key concepts and methodological considerations. *International journal of plant sciences* **171**, 1059-1071.

Horst, P. (1989). Native fowl as reservoir for genomes and major genes with direct and indirect effects on the adaptability and their potential for tropically orientated breeding plans. *Archiv fuer Gefluegelkunde (Germany, FR)*.

Hothorn, T., Bühlmann, P., Kneib, T., Schmid, M., and Hofner, B. (2010). Model-based boosting 2.0. *Journal of Machine Learning Research* **11**, 2109-2113.

Hothorn, T., Bühlmann, P., Kneib, T., Schmid, M., and Hofner, B. (2012). mboost: Model-based boosting. *R package version*, 2.1-2.

Howlider, M., and Rose, S. (1987). Temperature and the growth of broilers. *World's Poultry Science Journal* **43**, 228-237.

Huang, S., Zhang, L., Rehman, M. U., Iqbal, M. K., Lan, Y., Mehmood, K., Zhang, H., Qiu, G., Nabi, F., and Yao, W. (2017). High altitude hypoxia as a factor that promotes tibial growth plate development in broiler chickens. *PloS one* **12**, e0173698.

Hurni, H. (1998). Agroecological belts of Ethiopia. *Explanatory notes on three maps at a scale of* **1**.

Hutchinson, G. E. (1957). Concluding remarks. *In* "Cold Spring Harbor symposia on quantitative biology", Vol. 22, pp. 415-427. Cold Spring Harbor Laboratory Press.

ILRI (2019). "Options for the livestock sector in developing and emerging economies to 2030 and beyond." Meat: the future series. World Economic Forum Geneva, Switzerland.

IPP (2018). *In* "International Poultry Production ", Vol. 26. Positive Action Publications Ltd., South Plainfield, USA.

Isenberg, D. (2007). Keggfarms (India)-Which came first, the Kuroiler or the Kegg. *Harvard Business School Case Study*, 9-807.

Jiang, T., Liu, Y., Jiang, Y., Li, J., Gao, Y., Cui, Z., Liu, Y., Liu, B., and Wang, Y. (2020). Long-read-based human genomic structural variation detection with cuteSV. *Genome biology* **21**, 1-24.

Jombart, T., Pontier, D., and Dufour, A.-B. (2009). Genetic markers in the playground of multivariate analysis. *Heredity* **102**, 330-341.

Joost, S., Bonin, A., Bruford, M. W., DespréS, L., Conord, C., Erhardt, G., and Taberlet, P. (2007). A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology* **16**, 3955-3969.

Jueterbock, A., Smolina, I., Coyer, J. A., and Hoarau, G. (2016). The fate of the Arctic seaweed Fucus distichus under climate change: an ecological niche modeling approach. *Ecology and evolution* **6**, 1712-1724.

Kabir, M., and Muhammad, S. (2012). Study of Fertility and Hatchability in Shika-Brown Commercial and Parent Stock Layers. *Savannah Journal of Agriculture* **7**, 17-23.

Kassambara, A., and Mundt, F. (2017). Package 'factoextra'. *Extract and visualize the results of multivariate data analyses* **76**.

Kaufman, L., and Rousseeuw, P. J. (2009). "Finding groups in data: an introduction to cluster analysis," John Wiley & Sons.

Kawecki, T. J., and Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology letters* **7**, 1225-1241.

Kebede, F. G., Komen H, Dessie T, Alemu SW, O, H., and J, a. B. (2021). Species and Phenotypic Distribution Models Reveal Population Differentiation in Ethiopian Indigenous Chickens. *Front. Genet.* **12:723360**.

Kebede, F. G., Komen H, Dessie T, Hanotte O, Kemp S, SW, A., and J, a. B. (2022). Smallholder chicken production agroecologies defined by distribution models improve prediction of breed by environment interactions. *Manuscript submitted for publication*.

Kelly, M. (2019). Adaptation to climate change through genetic accommodation and assimilation of plastic phenotypes. *Philosophical Transactions of the Royal Society B* **374**, 20180176.

Kess, T., and Boulding, E. G. (2019). Genome-wide association analyses reveal polygenic genomic architecture underlying divergent shell morphology in Spanish Littorina saxatilis ecotypes. *Ecology and evolution* **9**, 9427-9441.

Khwatenge, C. N., and Nahashon, S. N. (2021). Recent Advances in the Application of CRISPR/Cas9 Gene Editing System in Poultry Species. *Frontiers in Genetics* **12**, 127.

Klecka, W. R., Iversen, G. R., and Klecka, W. R. (1980). "Discriminant analysis," Sage.

Knap, P., and Su, G. (2008). Genotype by environment interaction for litter size in pigs as quantified by reaction norms analysis. *Animal: an international journal of animal bioscience* **2**, 1742.

Knüpffer, H., Terentyeva, I., Hammer, K., and Kovaleva, O. (2003). Ecogeographical Diversity–a Vavilovian approach.–Diversity in Barley. Eds R. Invon Bothmer, T. Van Hintum, H. Knüpffer & K. Sato. Elsevier, Amsterdam.

Kome, G. K., Enang, R. K., Tabi, F. O., and Yerima, B. P. K. (2019). Influence of Clay Minerals on Some Soil Fertility Attributes: A Review. *Open Journal of Soil Science* **9**, 155-188.

Kumar, H., Park, W., Lim, D., Srikanth, K., Kim, J.-M., Jia, X.-Z., Han, J.-L., Hanotte, O., Park, J.-E., and Oyola, S. O. (2020). Whole metagenome sequencing of cecum microbiomes in Ethiopian indigenous chickens from two different altitudes reveals antibiotic resistance genes. *Genomics* **112**, 1988-1999.

Lafuente, E., and Beldade, P. (2019). Genomics of developmental plasticity in animals. *Frontiers in genetics* **10**, 720.

Lan, X., Hsieh, J. C., Schmidt, C. J., Zhu, Q., and Lamont, S. J. (2016). Liver transcriptome response to hyperthermic stress in three distinct chicken lines. *BMC genomics* **17**, 955.

Landon, J. R. (2014). "Booker tropical soil manual: a handbook for soil survey and agricultural land evaluation in the tropics and subtropics," Routledge.

Langerhans, R. B. (2008). Predictability of phenotypic differentiation across flow regimes in fishes. *Integrative and Comparative Biology* **48**, 750-768.

Lara, L. J., and Rostagno, M. H. (2013). Impact of heat stress on poultry production. *Animals* **3**, 356-369.

Lawal, R. A., Martin, S. H., Vanmechelen, K., Vereijken, A., Silva, P., Al-Atiyat, R. M., Aljumaah, R. S., Mwacharo, J. M., Wu, D.-D., and Zhang, Y.-P. (2020). The wild species genome ancestry of domestic chickens. *BMC biology* **18**, 1-18.

Le Corre, V., and Kremer, A. (2012). The genetic differentiation at quantitative trait loci under local adaptation. *Molecular ecology* **21**, 1548-1566.

Legendre, P., and Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia* **129**, 271-280.

Leinonen, T., Cano, J. M., Mäkinen, H., and Merilä, J. (2006). Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. *Journal of evolutionary biology* **19**, 1803-1812.

Lenth, R., Singmann, H., Love, J., Buerkner, P., and Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. *R package version* **1**, 3.

Lewontin, R. C., and Krakauer, J. (1973). Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* **74**, 175-195.

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987-2993.

Li, Y., Zhang, X.-X., Mao, R.-L., Yang, J., Miao, C.-Y., Li, Z., and Qiu, Y.-X. (2017). Ten years of landscape genomics: challenges and opportunities. *Frontiers in Plant Science* **8**, 2136.

Lin, C.-S., Binns, M. R., and Lefkovitch, L. P. (1986). Stability Analysis: Where Do We Stand? 1. *Crop science* **26**, 894-900.

Lin, H., Jiao, H., Buyse, J., and Decuypere, E. (2006). Strategies for preventing heat stress in poultry. *World's Poultry Science Journal* **62**, 71-86.

Liu, R., Sun, Y., Zhao, G., Wang, F., Wu, D., Zheng, M., Chen, J., Zhang, L., Hu, Y., and Wen, J. (2013). Genome-wide association study identifies loci and candidate genes for body composition and meat quality traits in Beijing-You chickens. *PLoS One* **8**, e61172.

Lorenzo, F. R., Huff, C., Myllymäki, M., Olenchock, B., Swierczek, S., Tashi, T., Gordeuk, V., Wuren, T., Ri-Li, G., and McClain, D. A. (2014). A genetic

mechanism for Tibetan high-altitude adaptation. *Nature genetics* **46**, 951-956.

Lotterhos, K. E., and Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology* **24**, 1031-1046.

Lozano-Jaramillo, a., Alemu, S., Dessie, T., Komen, H., and Bastiaansen, J. (2019a). Using phenotypic distribution models to predict livestock performance. *Scientific reports* **9**, 1-11.

Lozano-Jaramillo, M. (2019). Predicting breed by environment interaction using ecological modelling, Wageningen University, Wageningen.

Lozano-Jaramillo, M., Bastiaansen, J., Dessie, T., and Komen, H. (2019b). Use of geographic information system tools to predict animal breed suitability for different agro-ecological zones. *Animal* **13**, 1536-1543.

Luo, W., Xu, J., Li, Z., Xu, H., Lin, S., Wang, J., Ouyang, H., Nie, Q., and Zhang, X. (2018). Genome-Wide Association Study and Transcriptome Analysis Provide New Insights into the White/Red Earlobe Color Formation in Chicken. *Cellular Physiology and Biochemistry* **46**, 1768-1778.

Lwelamira, J., and Kifaro, G. (2010). Desired-gain selection indices for improving performance of two Tanzania local chicken ecotypes under intensive management.

Lyimo, C., Weigend, A., Msoffe, P., Eding, H., Simianer, H., and Weigend, S. (2014). Global diversity and genetic contributions of chicken populations from A frican, A sian and E uropean regions. *Animal genetics* **45**, 836-848.

MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L., and Hines, J. E. (2017). "Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence," Elsevier.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K., and Studer, M. (2013). Package 'cluster'. *Dosegljivo na*.

Magothe, T., Muhuyi, W., and Kahi, A. (2011). Genetic parameters for egg and body weights of indigenous chicken in Kenya. *Animal Production Society of Kenya*, 36.

Magothe, T., Okeno, T., Muhuyi, W., and Kahi, A. (2012). Indigenous chicken production in Kenya: I. Current status. *World's Poultry Science Journal* **68**, 119-132.

Maloney, K. O., Schmid, M., and Weller, D. E. (2012). Applying additive modelling and gradient boosting to assess the effects of watershed and reach characteristics on riverine assemblages. *Methods in Ecology and Evolution* **3**, 116-128.

Marsden, A., and Morris, T. (1987). Quantitative review of the effects of environmental temperature on food intake, egg output and energy balance in laying pullets. *British Poultry Science* **28**, 693-704.

Mayorga, E. J., Renaudeau, D., Ramirez, B. C., Ross, J. W., and Baumgard, L. H. (2019). Heat stress adaptations in pigs. *Animal Frontiers* **9**, 54-61.

Mayr, E. (1942). Systematics and the origin of species from the viewpoint of a zoologist. New York: Columbia University Press. 334 p.

Mazzi, C. M., Ferro, J. A., Ferro, M. I. T., Savino, V. J. M., Coelho, A. A. D., and Macari, M. (2003). Polymorphism analysis of the hsp70 stress gene in Broiler chickens (Gallus gallus) of different breeds. *Genetics and Molecular Biology* **26**, 275-281.

McConnell, R. (1967). The east African rift system. *Nature* **215**, 578-581.

Mdladla, K. (2016). Landscape genomic approach to investigate genetic adaptation in South African indigenous goat populations.

Mebrahtu, K. (2017). On farm evaluation of thermostable I-2 Newcastle disease vaccine and molecular characterization of NDV in smallholder poultry farms in Minjar-Shenkora wereda, Amhara Regional State, Ethiopia, Addis Ababa University.

Melesse, A., and Negesse, T. (2011). Phenotypic and morphological characterization of indigenous chicken populations in southern region of Ethiopia. *Animal Genetic Resources/Ressources génétiques animales/Recursos genéticos animales* **49**, 19-31.

Merilä, J., and Hendry, A. P. (2014). Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary applications* **7**, 1-14.

Michel, M. J. (2011). Spatial dependence of phenotype-environment associations for tadpoles in natural ponds. *Evolutionary Ecology* **25**, 915-932.

Michel, M. J., Chien, H., Beachum, C. E., Bennett, M. G., and Knouft, J. H. (2017). Climate change, hydrology, and fish morphology: predictions using phenotype-environment associations. *Clim Change* **140**, 563-576.

Mirkena, T., Walelign, E., Tewolde, N., Gari, G., Abebe, G., and Newman, S. (2018). Camel production systems in Ethiopia: a review of literature with notes on MERS-CoV risk factors. *Pastoralism* **8**, 30.

MoA (1998). "Agro-ecological Zones of Ethiopia." Ministry of Agriculture. Natural Resources Management & Regulatory Department, Ethiopia.

MoA (2000). Agro-ecological zonations of Ethiopia. (M. o. Agriculture, ed.).

Muchadeyi, F. C., and Dzomba, E. F. (2017). Genomics Tools for the Characterization of Genetic Adaptation of Low Input Extensively Raised Chickens. *In* "Poultry Science".

Mucksová, J., Reinišová, M., Kalina, J., Lejčková, B., Hejnar, J., and Trefil, P. (2019). Conservation of chicken male germline by orthotopic transplantation of primordial germ cells from genetically distant donors. *Biology of reproduction* **101**, 200-207.

Mukherjee, B., Ko, Y.-A., VanderWeele, T., Roy, A., Park, S. K., and Chen, J. (2012). Principal interactions analysis for repeated measures data: application to gene–gene and gene–environment interactions. *Statistics in medicine* **31**, 2531-2551.

Müntzing, A. (1971). Göte Wilhelm Turesson. *Taxon*, 773-775.

Murren, C., and Dudash, M. (2012). Variation in inbreeding depression and plasticity across native and non-native field environments. *Annals of Botany* **109**, 621-632.

Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., Anderson, R. P., and McPherson, J. (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity forMaxentecological niche models. *Methods in Ecology and Evolution* **5**, 1198-1205.

Mustefa, A., Assefa, A., Misganaw, M., Getachew, F., Abegaz, S., Hailu, A., and Emshaw, Y. (2020). Phenotypic characterization of donkeys in Benishangul Gumuz national regional state. *Journal of World's Poultry Research* **10**, 25-35.

Mwacharo, J. M., Bjørnstad, G., Han, J., and Hanotte, O. (2013). The history of African village chickens: an archaeological and molecular perspective. *African Archaeological Review* **30**, 97-114.

Mwacharo, J. M., Bjørnstad, G., Mobegi, V., Nomura, K., Hanada, H., Amano, T., Jianlin, H., and Hanotte, O. (2011). Mitochondrial DNA reveals multiple introductions of domestic chicken in East Africa. *Molecular phylogenetics and evolution* **58**, 374-382.

Nakamura, Y. (2016). Poultry genetic resource conservation using primordial germ cells. *Journal of Reproduction and Development*.

Nandi, S., Whyte, J., Taylor, L., Sherman, A., Nair, V., Kaiser, P., and McGrew, M. (2016). Cryopreservation of specialized chicken lines using cultured primordial germ cells. *Poultry science* **95**, 1905-1911.

Nasser, M., Lohr, J., Mebratu, G., Zessin, K.-H., Baumann, M., and Ademe, Z. (2000). Oral Newcastle disease vaccination trials in Ethiopia. *Avian Pathology* **29**, 27-34.

Neeteson-van Nieuwenhoven, A.-M., Knap, P., and Avendaño, S. (2013). The role of sustainable commercial pig and poultry breeding for food security. *Animal Frontiers* **3**, 52-57.

Nei, M. (1986). Definition and estimation of fixation indices. *Evolution* **40**, 643-645.

Nelder, J. A., and Wedderburn, R. W. (1972). Generalized linear models. *J R Stat Soc Ser A* **135**, 370-384.

Ng, C. S., Wu, P., Foley, J., Foley, A., McDonald, M.-L., Juan, W.-T., Huang, C.-J., Lai, Y.-T., Lo, W.-S., and Chen, C.-F. (2012). The chicken frizzle feather is due to an α-keratin (KRT75) mutation that causes a defective rachis. *PLoS genetics* **8**, e1002748.

Niknafs, S., Nejati-Javaremi, A., Mehrabani-Yeganeh, H., and Fatemi, S. A. (2012). Estimation of genetic parameters for body weight and egg production traits in Mazandaran native chicken. *Tropical animal health and production* **44**, 1437-1443.

Nwagu, B., Olorunju, S., Oni, O., Eduvie, L., Adeyinka, I., Sekoni, A., and Abeke, F. (2007). Response of egg number to selection in Rhode Island chickens selected for part period egg production. *Int. J. Poult. Sci* **6**, 18-22.

O'Donnell, M. S., and Ignizio, D. A. (2012). Bioclimatic predictors for supporting ecological applications in the conterminous United States. *US geological survey data series* **691**, 4-9.

Oddi, F. J., Miguez, F. E., Ghermandi, L., Bianchi, L. O., and Garibaldi, L. A. (2019). A nonlinear mixed-effects modeling approach for ecological data: Using temporal dynamics of vegetation moisture as an example. *Ecology and Evolution* **9**, 10225-10240.

Oishi, I., Yoshii, K., Miyahara, D., Kagami, H., and Tagami, T. (2016). Targeted mutagenesis in chicken using CRISPR/Cas9 system. *Scientific reports* **6**, 1-10.

Okeno, T., Kahi, A., and Peters, K. (2013). Evaluation of breeding objectives for purebred and crossbred selection schemes for adoption in indigenous chicken breeding programmes. *British Poultry Science* **54**, 62-75.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R., Simpson, G. L., Solymos, P., Stevens, M. H. H., and Wagner, H. (2013). Package 'vegan'. *Community ecology package, version* **2**, 1-295.

Oliveira, D., Lourenco, D., Tsuruta, S., Misztal, I., Santos, D., de Araújo Neto, F., Aspilcueta-Borquis, R., Baldi, F., Carvalheiro, R., and de Camargo, G. (2018). Reaction norm for yearling weight in beef cattle using single-step genomic evaluation. *Journal of animal science* **96**, 27-34.

Olivoto, T., Lúcio, A. D., da Silva, J. A., Marchioro, V. S., de Souza, V. Q., and Jost, E. (2019a). Mean performance and stability in multi-environment trials I: Combining features of AMMI and BLUP techniques. *Agronomy Journal* **111**, 2949-2960.

Olivoto, T., Lúcio, A. D., da Silva, J. A., Sari, B. G., and Diel, M. I. (2019b). Mean performance and stability in multi-environment trials II: Selection based on multiple traits. *Agronomy Journal* **111**, 2961-2969.

Olivoto, T., and Lúcio, A. D. C. (2020). metan: An R package for multi-environment trial analysis. *Methods in Ecology and Evolution* **11**, 783-789.

Overdijk, S. (2019). The effect of ecological factors on the morphology of Ethiopian chickens. MSc, Wageningen University and Research.

Pariset, L., Joost, S., Gargani, M., and Valentini, A. (2012). Landscape genomics in livestock. *In* "Analysis of Genetic Variation in Animals". IntechOpen.

Passarelli, S., Ambikapathi, R., Gunaratna, N. S., Madzorera, I., Canavan, C. R., Noor, A. R., Worku, A., Berhane, Y., Abdelmenan, S., and Sibanda, S. (2020). A chicken production intervention and additional nutrition behavior change component increased child growth in Ethiopia: a cluster-randomized trial. *The Journal of nutrition* **150**, 2806-2817.

Patterson, N., Price, A. L., and Reich, D. (2006). Population structure and eigenanalysis. *PLoS genet* **2**, e190.

Pavlidis, P., and Alachiotis, N. (2017). A survey of methods and tools to detect recent and strong positive selection. *Journal of Biological Research-Thessaloniki* **24**, 1-17.

Pebesma, E., Bivand, R., Pebesma, M. E., RColorBrewer, S., and Collate, A. (2012). Package 'sp'. *The Comprehensive R Archive Network*.

Pedroso, A., and Lee, M. (2015). The composition and role of the microbiota in chickens. *Intestinal Health. Wageningen: Wageningen Academic Publishers*, 21-50.

Pfennig, D. W. (2021). "Phenotypic Plasticity & Evolution: Causes, Consequences, Controversies," Taylor & Francis.

Phillimore, A. B., Hadfield, J. D., Jones, O. R., and Smithers, R. J. (2010). Differences in spawning date between populations of common frog reveal local adaptation. *Proceedings of the National Academy of Sciences* **107**, 8292-8297.

Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**, 231-259.

Phillips, S. J., and Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* **31**, 161-175.

Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., and Ferrier, S. (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological applications* **19**, 181-197.

Piepho, H.-P. (1994). Best linear unbiased prediction (BLUP) for regional yield trials: a comparison to additive main effects and multiplicative interaction (AMMI) analysis. *Theoretical and Applied Genetics* **89**, 647-654.

Piepho, H., and Möhring, J. (2005). Best linear unbiased prediction of cultivar effects for subdivided target regions. *Crop Science* **45**, 1151-1159.

Porto-Neto, L. R., Reverter, A., Prayaga, K. C., Chan, E. K., Johnston, D. J., Hawken, R. J., Fordyce, G., Garcia, J. F., Sonstegard, T. S., and Bolormaa, S. (2014). The genetic architecture of climatic adaptation of tropical cattle. *Plos one* **9**.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics* **38**, 904-909.

Psifidi, A., Banos, G., Matika, O., Desta, T. T., Bettridge, J., Hume, D. A., Dessie, T., Christley, R., Wigley, P., and Hanotte, O. (2016). Genome-wide association studies of immune, disease and production traits in indigenous chicken ecotypes. *Genetics Selection Evolution* **48**, 74.

Purchase, J., Hatting, H., and Van Deventer, C. (2000). Genotype× environment interaction of winter wheat (Triticum aestivum L.) in South Africa: II. Stability analysis of yield performance. *South African Journal of Plant and Soil* **17**, 101-107.

Radosavljevic, A., and Anderson, R. P. (2014). Making better Maxent models of species distributions: complexity, overfitting and evaluation. *Journal of biogeography* **41**, 629-643.

Rajora, O. P. (2019). Population genomics: concepts, approaches and applications. Springer, Cham, Switzerland.

Rasband, W. S. (1997). ImageJ 1.52a. National Institute of Health, USA.

Razgour, O. (2015). Beyond species distribution modeling: a landscape genetics approach to investigating range shifts under future climate change. *Ecological Informatics* **30**, 250-256.

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., and Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology* **24**, 4348-4370.

Resende, M. d. (2007). SELEGEN-REML/BLUP: sistema estatístico e seleção genética computadorizada via modelos lineares mistos. *Colombo: Embrapa Florestas*.

Rincent, R., Malosetti, M., Ababaei, B., Touzy, G., Mini, A., Bogard, M., Martre, P., Le Gouis, J., and van Eeuwijk, F. (2019). Using crop growth model stress covariates and AMMI decomposition to better predict genotype-by-environment interactions. *Theoretical and Applied Genetics* **132**, 3399-3411.

Robinson, S., Mason-D'Croz, D., Sulser, T., Islam, S., Robertson, R., Zhu, T., Gueneau, A., Pitois, G., and Rosegrant, M. W. (2015). The international model for policy analysis of agricultural commodities and trade (IMPACT): model description for version 3.

Rodrigues, P. C., Malosetti, M., Gauch Jr, H. G., and van Eeuwijk, F. A. (2014). A weighted AMMI algorithm to study genotype-by-environment interaction and QTL-by-environment interaction. *Crop Science* **54**, 1555-1570.

Rodrigues, P. C., Monteiro, A., and Lourenço, V. M. (2016). A robust AMMI model for the analysis of genotype-by-environment data. *Bioinformatics* **32**, 58-66.

Roffler, G. H., Schwartz, M. K., Pilgrim, K. L., Talbot, S. L., Sage, G. K., Adams, L. G., and Luikart, G. (2016). Identification of landscape features influencing gene flow: How useful are habitat selection models? *Evolutionary applications* **9**, 805-817.

Ross, J., Hale, B., Gabler, N., Rhoads, R., Keating, A., and Baumgard, L. (2015). Physiological consequences of heat stress in pigs. *Animal Production Science* **55**, 1381-1390.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* **20**, 53-65.

Rovelli, G., Ceccobelli, S., Perini, F., Demir, E., Mastrangelo, S., Conte, G., Abeni, F., Marletta, D., Ciampolini, R., and Cassandro, M. (2020). The genetics of

phenotypic plasticity in livestock in the era of climate change: A review. *Italian Journal of Animal Science* **19**, 997-1014.

Ruane, J., and Sonnino, A. (2007). Marker-assisted selection as a tool for genetic improvement of crops, livestock, forestry and fish in developing countries: an overview of the issues. *Marker-assisted Selection-Current Status and Future Perspectives in Crops, Livestock, Forestry and Fish.-Rome (Food and Agriculture Organization of the United Nations (FAO))*, 3-13.

Sa'diyah, H., and Hadi, A. F. (2016). AMMI Model for yield estimation in multi-environment trials: A comparison to BLUP. *Agriculture and Agricultural Science Procedia* **9**, 163-169.

Sabeti, P. C., Schaffner, S. F., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O., Palma, A., Mikkelsen, T., Altshuler, D., and Lander, E. (2006). Positive natural selection in the human lineage. *science* **312**, 1614-1620.

Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E. H., McCarroll, S. A., and Gaudet, R. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913-918.

Safalaoh, A. (2001). Village chicken upgrading programme in Malawi. *World's Poultry Science Journal* **57**, 179-188.

SAS (2002). User Installation Guide for the SAS® System Version 9 for Microsoft® Windows®SAS Institute Inc., Cary, NC.

Savolainen, O., Lascoux, M., and Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics* **14**, 807.

Scheiner, S. M. (1993). Genetics and evolution of phenotypic plasticity. *Annual review of ecology and systematics* **24**, 35-68.

Scherf, B. D., and Pilling, D. (2015). The second report on the state of the world's animal genetic resources for food and agriculture.

Schlichting, C., and Levin, D. (1986). Effects of inbreeding on phenotypic plasticity in cultivated Phlox. *Theoretical and Applied Genetics* **72**, 114-119.

Schmid, M., and Guillaume, F. (2017). The role of phenotypic plasticity on population differentiation. *Heredity* **119**, 214-225.

Schou, M. F., Kristensen, T. N., and Loeschcke, V. (2015). Trait-specific consequences of inbreeding on adaptive phenotypic plasticity. *Ecology and evolution* **5**, 1-6.

Selmoni, O., Vajana, E., Guillaume, A., Rochat, E., and Joost, S. (2020). Sampling strategy optimization to increase statistical power in landscape genomics: A simulation-based approach. *Molecular ecology resources* **20**, 154-169.

Sgro, C. M., Terblanche, J. S., and Hoffmann, A. A. (2016). What can plasticity contribute to insect responses to climate change? *Annual review of entomology* **61**, 433-451.

Shane, S. M. (1988). Factors influencing health and performance of poultry in hot climates. *Critical Reviews in Poultry Biology* **1**, 247-267.

Shapiro, B. I., Gebru, G., Desta, S., Negassa, A., Negussie, K., Aboset, G., and Mechal, H. (2015). Ethiopia livestock master plan: Roadmaps for growth and transformation.

Sheng, Z., Pettersson, M. E., Hu, X., Luo, C., Qu, H., Shu, D., Shen, X., Carlborg, Ö., and Li, N. (2013). Genetic dissection of growth traits in a Chinese indigenous× commercial broiler chicken cross. *BMC genomics* **14**, 1-12.

Shlomo, Y. (2000). Domestic fowl-strategies to confront environmental conditions. *Poultry and Avian Biology Reviews* **11**, 81-95.

Shukla, G. (1972). Some statistical aspects of partitioning genotype environmental components of variability. *Heredity* **29**, 237-245.

Sieriebriennikov, B., Prabh, N., Dardiry, M., Witte, H., Röseler, W., Kieninger, M. R., Rödelsperger, C., and Sommer, R. J. (2018). A developmental switch generating phenotypic plasticity is part of a conserved multi-gene locus. *Cell reports* **23**, 2835-2843. e4.

Singh, D. (2002). Utilization of Indian Native Chicken for development of egg type scavenging chicken. *In* "Second National Seminar on Rural Poultry for Adverse Environment, University of Agricultural Sciences, Bangalore, India".

Singh, M., Islam, R., and Avasthe, R. (2017). Production performance of vanaraja birds under traditional tribal production system of Sikkim Himalayan Region. *International Journal of Livestock Research* **7**, 153-157.

Smith, A. B., Alsdurf, J., Knapp, M., Baer, S. G., and Johnson, L. C. (2017). Phenotypic distribution models corroborate species distribution models: A shift in the role and prevalence of a dominant prairie grass in response to climate change. *Global change biology* **23**, 4365-4375.

Smith, J. M. (1956). Acclimatization to high temperatures in inbred and outbkedDrosophila subobsoura. *Journal of genetics* **54**, 497-505.

Soberón, J. (2007). Grinnellian and Eltonian niches and geographic distributions of species. *Ecology letters* **10**, 1115-1123.

Sonaiya, E. B., and Swan, S. (2007). Small scale poultry production: technical guide.

Soussana, J.-F., Tichit, M., Lecomte, P., and Dumont, B. (2015). "Agroecology: integration with livestock," FAO, Rome.

Storfer, A., Patton, A., and Fraik, A. K. (2018). Navigating the Interface Between Landscape Genetics and Landscape Genomics. *Frontiers in genetics* **9**, 68.

Storz, J. F. (2002). Contrasting patterns of divergence in quantitative traits and neutral DNA markers: analysis of clinal variation. *Molecular Ecology* **11**, 2537-2551.

Sultan, S. E. (2021). Phenotypic plasticity as an intrinsic property of organisms. *In* "Phenotypic Plasticity & Evolution", pp. 3-24. CRC Press.

Tadelle, D., Alemu, Y., and Peters, K. J. (2000). Indigenous chickens in Ethiopia: genetic potential and attempts at improvement. *World's Poultry Science Journal* **56**, 45-54.

Tadesse, M., Alemu, B., Bekele, G., Tebikew, T., Chamberlin, J., and Benson, T. (2006). "Atlas of the Ethiopian rural economy," Intl Food Policy Res Inst.

Tadesse Mulugeta, A. B., Bekele Gashaw, Tebikew Tewodros, Chamberlin Jordan, Benson Todd (2006). "Atlas of the Ethiopian rural economy," Intl Food Policy Res Inst.

Talbot, B., Chen, T.-W., Zimmerman, S., Joost, S., Eckert, A. J., Crow, T. M., Semizer-Cuming, D., Seshadri, C., and Manel, S. (2017). Combining genotype, phenotype, and environment to infer potential candidate genes. *Journal of Heredity* **108**, 207-216.

Team, R. D. C. (2013). R development core team. *In* "RA Lang Environ Stat Comput", Vol. 55, pp. 275-286.

Thomas, J., Hepp, T., Mayr, A., and Bischl, B. (2017). Probing for sparse and fast variable selection with model-based boosting. *Computational and mathematical methods in medicine* **2017**.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* **58**, 267-288.

Tilahun, H., and Schmidt, E. (2012). Spatial analysis of livestock production patterns in Ethiopia. Ethiopia Strategy Support Program II (ESSP II) Working Paper 44.

Torrado, H., Carreras, C., Raventos, N., Macpherson, E., and Pascual, M. (2020). Individual-based population genomics reveal different drivers of adaptation in sympatric fish. *Scientific reports* **10**, 1-14.

Turesson, G. (1922). The species and the variety as ecological units. *Hereditas* **3**, 100-113.

Vajana, E., Barbato, M., Colli, L., Milanesi, M., Rochat, E., Fabrizi, E., Mukasa, C., Del Corvo, M., Masembe, C., Muwanika, V. B., Kabi, F., Sonstegard, T. S., Huson, H. J., Negrini, R., Joost, S., Ajmone-Marsan, P., and NextGen, C. (2018). Combining Landscape Genomics and Ecological Modelling to Investigate Local Adaptation of Indigenous Ugandan Cattle to East Coast Fever. *Frontiers in genetics* **9**, 385.

Valette, T., Leitwein, M., Lascaux, J.-M., Desmarais, E., Berrebi, P., and Guinand, B. (2020). Spotting genome-wide pigmentation variation in a brown trout admixture context. *bioRxiv*.

Vallejo-Trujillo, A., Kebede, A., Lozano-Jaramillo, M., Dessie, T., Smith, J., Hanotte, O., and Gheyas, A. (2021). A framework for defining livestock ecotypes based on ecological modelling and exploring genomic environmental adaptation: the example of Ethiopian village chicken. *bioRxiv*.

Vallejo-Trujillo, A., Kebede, A., Lozano, M., Dessie, T., Sparks, N., Smith, J., Hanotte, O., and Gheyas, A. (2018). "Ecological niche modelling applies for the characterization of indigenous livestock species: the example of Ethiopian village chickens."

Van Eeuwijk, F. A., Bustos-Korts, D. V., and Malosetti, M. (2016). What should students in plant breeding know about the statistical aspects of genotype× environment interactions? *Crop Science* **56**, 2119-2140.

Vangestel, C., Eckert, A. J., Wegrzyn, J. L., St. Clair, J. B., and Neale, D. B. (2018). Linking phenotype, genotype and environment to unravel genetic components underlying cold hardiness in coastal Douglas-fir (Pseudotsuga menziesii var. menziesii). *Tree Genetics & Genomes* **14**.

Venables, W., and Ripley, B. (2002). Modern applied statistics with S fourth edition. World.

Venkatramaiah, A., Mohapatra, S., Sinha, R., Ayyagari, V., and Choudhuri, D. (1986). Selection response for part period egg number and egg mass in chickens— a comparison. *Theoretical and applied genetics* **72**, 129-134.

Voight, B. F., Kudaravalli, S., Wen, X., and Pritchard, J. K. (2006). A map of recent positive selection in the human genome. *PLoS biology* **4**, e72.

Wang, J., Street, N. R., Scofield, D. G., and Ingvarsson, P. K. (2016). Variation in linked selection and recombination drive genomic divergence during allopatric speciation of European and American aspens. *Molecular Biology and Evolution* **33**, 1754-1767.

Wang, Q., Pi, J., Pan, A., Shen, J., and Qu, L. (2017). A novel sex-linked mutant affecting tail formation in Hongshan chicken. *Scientific reports* **7**, 1-8.

Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association* **58**, 236-244.

Warren, D., Glor, R., and Turelli, M. (2011). ENMTools user manual v1. 3. 1-34.

Warren, D. L., Glor, R. E., and Turelli, M. (2008). Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution: International Journal of Organic Evolution* **62**, 2868-2883.

Warren, D. L., Glor, R. E., and Turelli, M. (2010). ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* **33**, 607-611.

Warren, D. L., and Seifert, S. N. (2011). Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological applications* **21**, 335-342.

Warren, D. L., Wright, A. N., Seifert, S. N., and Shaffer, H. B. (2014). Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 C alifornia vertebrate species of concern. *Divers Distrib* **20**, 334-343.

Wiens, J. A., Stralberg, D., Jongsomjit, D., Howell, C. A., and Snyder, M. A. (2009). Niches, models, and climate change: assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences* **106**, 19729-19736.

Wiley, M., and Wiley, J. F. (2019). "Advanced R Statistical Programming and Data Models: Analysis, Machine Learning, and Visualization," Apress.

Williams, G. (1966). Adaptation and natural selection Princeton University Press. *Princeton, NJ*.

Windhorst, H.-W. (2008). A projection of the regional development of egg production until 2015. *World's Poultry Science Journal* **64**, 356-376.

Woldekiros, H., and D'Andrea, A. (2017). Early evidence for domestic chickens (Gallus gallus domesticus) in the Horn of Africa. *International Journal of Osteoarchaeology* **27**, 329-341.

Woltereck, R. (1909). Weitere experimentelle Unter-suchungen uber Artveranderung, speziell uber das Wesen quantitativer Artunterschiede bei Daphniden. *Verh. D. Tsch. Zool. Ges.* **1909**, 110-172.

Wondmeneh, E., Dawud, I., and Adey, M. (2011). Comparative Evaluation of Fertility and Hatchability of Horro, Fayoumi, Lohmann Silver and Potchefstroom Koekoek Breeds of. *Asian Journal of Poultry Science* **5**, 124-129.

Wood, S. (2012). mgcv: Mixed GAM Computation Vehicle with GCV/AIC/REML smoothness estimation.

Wood, S. N. (2017). "Generalized additive models: an introduction with R," CRC press.

Wood, S. N., and Augustin, N. H. (2002). GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecological modelling* **157**, 157-177.

Wricke, G. (1964). Zur berechnung der okovalenz bei sommerweizen und hafer. *Zeitschrift Fur Pflanzenzuchtung-Journal of Plant Breeding* **52**, 127-&.

Wright, D., Boije, H., Meadows, J. R., Bed'Hom, B., Gourichon, D., Vieaud, A., Tixier-Boichard, M., Rubin, C.-J., Imsland, F., and Hallböök, F. (2009). Copy number variation in intron 1 of SOX5 causes the Pea-comb phenotype in chickens. *PLoS genetics* **5**.

Wu, Z., Bortoluzzi, C., Derks, M. F., Liu, L., Bosse, M., Hiemstra, S. J., Groenen, M. A., and Crooijmans, R. P. (2021). Heterogeneity of a dwarf phenotype in Dutch traditional chicken breeds revealed by genomic analyses. *Evolutionary applications* **14**, 1095-1108.

Xie, L., Luo, C., Zhang, C., Zhang, R., Tang, J., Nie, Q., Ma, L., Hu, X., Li, N., and Da, Y. (2012). Genome-wide association study identified a narrow chromosome 1 region associated with chicken growth traits. *PLoS one* **7**, e30910.

Xiong, J., Thenkabail, P., Tilton, J., Gumma, M., Teluguntla, P., Congalton, R., Yadav, K., Dungan, J., Oliphant, A., and Poehnelt, J. (2017). NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) Global Food Security-support Analysis Data (GFSAD) Cropland Extent 2015 Africa 30 m V001.

Xue, B., and Leibler, S. (2018). Benefits of phenotypic plasticity for population growth in varying environments. *Proceedings of the National Academy of Sciences* **115**, 12745-12750.

Zobel, R. W., Wright, M. J., and Gauch Jr, H. G. (1988). Statistical analysis of a yield trial. *Agronomy journal* **80**, 388-393.

Zuur, A., Ieno, E. N., and Smith, G. M. (2007). "Analyzing ecological data," Springer.

Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A., and Smith, G. M. (2009). "Mixed effects models and extensions in ecology with R," Springer Science & Business Media.

Summary

Smallholder chicken production is dominant in many tropical developing countries and contributes significantly to sustainable livelihoods. As a low input-low output production system, it has low efficiency to meet growing demands for meat and eggs in these regions. The lack of productive breeds and inadequate understanding of the roles of local adaptation are major factors contributing to poor performance. Knowledge on adaptive phenotypic and genetic variation is essential to design sustainable chicken genetic improvement and development programmes.

In this thesis I aim to address two overarching research questions:
What are the environmental drivers of local adaptation, and phenotypic and genetic differentiation in indigenous chickens?
How do improved chicken populations that are introduced into smallholder systems respond phenotypically to environmental variations? To answer these questions, I follow a landscape genomics approach and integrate genetic, phenotypic, and environmental information in my study design and statistical analyses.

In the first part of the thesis I investigate phenotypic and genetic differentiation in indigenous chickens. In **chapter 2**, I perform species distribution models (SDMs) and apply correlative methods to identify environmental predictors associated with habitat suitability and phenotypic differentiation in Ethiopian indigenous chickens. I report that the presence of population differentiation among Ethiopian chickens is probably in response to environmental variation. I use habitat suitability maps produced by SDMs to show that populations went through different environmental selective pressures. Based on the matching between the presence of distinct phenotypes and availability of unique environmental niches, I classify the Ethiopian indigenous chicken populations into three ecotypes.

In **chapter 3,** I look for candidate genes and regions under positive selection in different agroecologies (lowland, midaltitude, and highland) and environmental gradients (clines in different geographies). I show that phenotypic differentiation in Ethiopian indigenous chickens has a genetic basis. I look at independent results and overlaps between two methods of signatures of selection (F_{ST} and XP-EHH). I show that that Ethiopian chicken populations differentiated the most between gradients but selection pressures leading to adaptive variation are stronger between agroecologies. These results lead to the hypothesis that evolutionary processes other than natural selection, such as gene flow and genetic drift may have contributed to genetic divergence among populations sampled from different gradients.

I perform Redundancy analysis (RDA) and show that SDM-identified environmental predictors and quantitative traits are useful to explain variations in the genome. I demonstrate that RDA can be used as an alternative approach to GWAS in random mating, indigenous livestock populations which have sufficiently interacted with the environment. I indicate that the results from RDA are supported by the outputs from signatures of selection analyses (F_{ST} and XP-EHH). I demonstrate that signatures of selection analysis with the two methods (F_{ST} and XP-EHH) can be used complementarily with RDA to shed light on the relationship between genomic, phenotypic, and environmental variation in local adaptation studies in indigenous chickens.

In the second part of the thesis, I evaluate the performance of improved chicken breeds introduced into smallholder systems. In **Chapter 4,** I apply distribution models to compare performances of improved chicken breeds introduced into smallholder systems. I show that classifying agroecologies based on environmental variables associated with habitat suitability and phenotypic differentiation of a livestock species improves model fit in GxE predictions. I demonstrate that phenotypic distribution models (PDMs) like boosted GAMs and boosted GLMs are valuable tools in animal breeding to integrate environmental and phenotypic information and predict phenotypic values.

Finally, in **chapter 5,** I utilize the concept of phenotypic plasticity to evaluate yield stability among improved chickens distributed to smallholder farmers. I implement two methods of multi-environment breed performance analysis (MEPA), namely, additive main effects and multiplicative interaction model (AMMI) and linear mixed-effects model (LMM) to identify and recommend more productive and stable breeds for wider dissemination into smallholder systems. I report that that improved chicken breeds introduced into different agroecologies significantly vary in growth performance and yield stability probably owing to their different genetic backgrounds.

Together, I demonstrate in this thesis how adaptive phenotypic and genetic variation can be exploited to enhance performance of chickens in smallholder systems.

Curriculum vitae

About the author

Fasil Getachew Kebede was born on the 1st of October 1978 in Ethiopia. He earned his BSc degree in Animal Science in 2001 and MSc degree in Animal Genetics and Breeding in 2006, both from Haramaya University, Ethiopia. In the last twenty-one years, Fasil contributed to the Ethiopian livestock sector at different capacities: as a researcher at the Amhara Regional Agricultural Research Institute (ARARI) (2001-2006); as a lecturer of animal breeding and genetics at the College of Veterinary Medicine and Animal Sciences, University of Gondar (2006-2011); as an animal breeder at the National Artificial Insemination Centre (NAIC) of the Ethiopian Ministry of Agriculture (MoA) (2011); as the national coordinator for domestic animal biodiversity research at the Ethiopian Biodiversity Institute (EBI) (2011-2015); and as a research officer in the Livestock Genetics Programme of the International Livestock Research Institute (ILRI) (2015-2018).

From April 2018 to August 2022, he has been a PhD candidate at Animal Breeding and Genomics (ABG) of Wageningen University and Research (WUR) and a Graduate Fellow at the International Livestock Research Institute (ILRI), Ethiopia.

Collaborative works with national and international project partners and extensive field activities created wonderful opportunities for him to understand the complexities of smallholder livestock systems in several tropical countries. Fasil is passionate about transforming the lives of smallholder farmers through agricultural research and development programmes. His present research focuses on integrating phenotypic, genomic, and environmental information to enhance livestock productivity and adaptability in Africa and Southeast Asia. In his free time, Fasil enjoys reading philosophy books, listening to music, hanging out with his family, and hiking.

Peer reviewed publications

Kebede FG, Komen H, Dessie T, Alemu SW, Hanotte O and Bastiaansen JWM. 2021. Species and Phenotypic Distribution Models Reveal Population Differentiation in Ethiopian Indigenous Chickens. *Frontiers in Genetics,* 12:723360.

Getachew F, Assefa A, Getachew T, Kebede SA, Hailu A, Mesganaw M, Emishaw Y, Tessema M. 2020. On-farm phenotypic characterization of Begaria cattle population and their production system in Guba district, North Western Ethiopia. *Ethiopian Journal of Animal Production*, 20(1), 1-17.

Getachew F, Abegaz S, Assefa A, Misganaw M, Emshaw Y, Hailu A, Tessema M, and Okore C. 2016. Multivariate analyses of morphological traits in indigenous chicken populations of Metekel zone, Northwestern Ethiopia. *Animal Genetic Resources*, 59, 15-25.

Getachew F, Assefa A, Mesganaw M and Fekansa T. 2014. On-farm phenotypic characterization of Ogaden cattle populations of Jigjiga zone, southeastern Ethiopia. *Ethiopian Journal of Animal Production* 14, 66-83.

Alemu SW, Hanotte O, **Kebede FG**, Esatu W, Abegaz S, Bruno JE, Abrar B, Alemayehu T, Mrode, R, and Dessie T. 2021. Evaluation of live-body weight and the number of eggs produced for introduced and local chickens in Ethiopia. *Acta Agriculturae Scandinavica, Section A—Animal Science* 70, 71-77.

Aseged T, Hailu A, Assefa A, Getachew T, Misganaw M, Sinke S, **Getachew F**, and Abegaz S. 2021. Goat breeding practice and production constraints in Boset and Minjar Shenkora districts of Ethiopia. *Genetics and Biodiversity Journal* 5, 107-117.

Ayele TA, Assefa A, Hailu A, Getachew T, Misganaw M, Sinke S, **Getachew F**, and Abegaz S. 2021. Phenotypic Characterization for Identification, Conservation and Sustainable Utilization of Boset and Minjar Shenkora Goats of Ethiopia. *Turkish Journal of Agriculture-Food Science and Technology*, 9(10), 1876-1882.

Birhanu MY, Alemayehu T, Bruno JE, **Kebede FG**, Sonaiya EB, Goromela EH, Bamidele O, and Dessie T. 2021a. Technical Efficiency of Traditional Village Chicken Production in Africa: Entry Points for Sustainable Transformation and Improved Livelihood. *Sustainability* 13, 8539.

Birhanu MY, Bruno JE, Alemayehu T, Esatu W, Geremew K, **Kebede FG,** Dessie Tadelle. 2022. Beyond diffusion to sustained adoption of innovation: A case of smallholder poultry development in sub-Saharan Africa. *International Journal of Agricultural Sustainability*, DOI: 10.1080/14735903.2022.2041235

Mustefa A, Assefa A, Misganaw M, **Getachew F**, Abegaz S, Hailu A, and Emshaw Y. 2020. Phenotypic characterization of donkeys in Benishangul Gumuz national regional state. *Online Journal of Animal and Feed Research*,10 (1), 25-35.

Manuscripts in preparation

Kebede FG, Komen H, Dessie T, Hanotte O, Kemp S, Alemu SW, and Bastiaansen JWM. 2022. Smallholder chicken production agroecologies defined by distribution models improve prediction of breed by environment interactions. *Under review.*

Kebede FG, Komen H, Dessie T, Hanotte O, Kemp S, Alemu SW, and Bastiaansen JWM. 2022. Multi-environment performance analysis (MEPA) to evaluate productivity and yield stability in livestock. *To be submitted.*

Kebede FG, Komen H, Crooijmans RPMA, Dessie T, Hanotte O, Pita Barros C, Derks MFL, and Bastiaansen JWM. 2022. Landscape analyses reveal the genome regions associated with morphological and climatic adaptation in Ethiopian indigenous chickens. *To be submitted.*

Conference proceedings and research reports

Getachew F, Worku, S., Esatu, W. and Dessie, T. 2019. Unleashing the power of data in transforming livestock agriculture in Ethiopia. Paper presented at the 27th Annual Conference of the Ethiopian Society of Animal Production (ESAP), EIAR, Addis Ababa, 29-31 August.

Kebede F, Komen H, Dessie T, Hanotte O, Kemp S, Pita Barros C, Crooijmans R, Derks M, Alemu S, Bastiaansen J. 2022. Exploiting phenotypic plasticity in animal breeding. In: 12th World Congress on Genetics Applied to Livestock Production (WCGALP). 3-8 July 2022. Wageningen Academic Publishers, Rotterdam, the Netherlands.

Kebede F, Komen H, Alemu S, Hanotte O, Dessie T and Bastiaansen J. 2020. Distribution models unravel drivers of local adaptation in chickens. In: 71st Annual Meeting of the European Federation of Animal Science, pp. 375. Wageningen Academic Publishers, Virtual meeting.

Abegaz, S., Esatu, W., Assefa, G., Goromela, E.H., Sonaiya, E.B., Mbaga, S.H., Adebambo, O., Bamidele, O., Teressa, A., Bruno, J., Poole, J., **Getachew F**, Kasaye, H. and Dessie, T. 2019. On-farm performance testing of tropically adaptable chicken strains under small holder management in three countries of sub-Saharan Africa. In: 7th All Africa Conference on Animal Agriculture, Accra, Ghana, 29 July - 2 August.

Alemayehu T, Bruno J, **Getachew F**, and Dessie T. 2018a. Socio-economic, marketing and gender aspects of village chicken production in the tropics: a review of literature. ILRI Research Report, Nairobi, Kenya.

Alemayehu T, Bruno JE, Poole EJ, **Getachew F**, Goromela E, Mbaga S, and Dessie T. 2018b. Household baseline data in Ethiopia: Monitoring delivery of chicken genetic gains. ILRI Research Report, Nairobi, Kenya.

Alemayehu T, Bruno JE, Poole EJ, **Getachew F**, Goromela E, Mbaga S, and Dessie T. 2018c. Household baseline data in Tanzania: Monitoring delivery of chicken genetic gains. ILRI Research Report, Nairobi, Kenya.

Alemayehu, T, Bruno JE, Poole EJ, **Getachew F**, Sonaiya EB, Bamidele O, Adebambo O, Adeyinka, IA, and Dessie T. 2018d. Household baseline data in Nigeria: Monitoring delivery of chicken genetic gains. ILRI Research Report, Nairobi, Kenya.

Birhanu MY, Esatu W, Geremew K, Worku S, **Kebede FG**, Unger F, and Dessie T. 2021b. Poultry production, marketing, and consumption in Myanmar: A review of literature. ILRI Research Report 89. Nairobi, Kenya: ILRI.

Birhanu MY, Geremew K, Esatu W, Worku S, **Getachew F**, Nguyen V, Ngo T, Unger F, and Dessie T. 2021c. Poultry production, marketing, and consumption in Vietnam: A review of literature. ILRI Research Report 80. Nairobi, Kenya: ILRI.

Birhanu MY, Geremew K, Wondmeneh E, Worku S, **Kebede FG**, Ty C, and Dessie T. Poultry Production, Marketing, and Consumption in Cambodia: A Review of Literature. ILRI Research Report 81. Nairobi, Kenya: ILRI.

Goromela, E.H., Sonaiya, E.B., Assefa, G., Mbaga, S.H., Adebambo, O. Abegaz, S., Bamidele, O., Esatu, W., Alemayehu, T., Bruno, J., Poole, J., **Getachew F** and Dessie, T. 2019. Understanding the Entry Points for Improving the Smallholder Chicken Production in Sub-Saharan Africa. In: 7th All Africa Conference on Animal Agriculture, Accra, Ghana, 29 July - 2 August.

Sonaiya EB, Goromela EH, Assefa G, Bamidele O, Mbaga SH, Abegaz S, Alemayehu T, **Getachew F** and Dessie T. 2019. Doing it differently: A farmer-focused, private sector-driven approach to delivering innovation that works for women in the smallholder chicken value chain. In: 7th All Africa Conference on Animal Agriculture, Accra, Ghana, 29 July - 2 August.

Training and Education

The Basic Package (1.8 credits)	
WIAS introduction day	2018
Scientific integrity & ethics and animal sciences	2018

Disciplinary Competences (17.6 credits)	
Genomics	2018
Quantitative and population genetics	2018
PhD research project proposal	
2018	
A PhD discussion group on quantitative genetics	2018
Meta-analysis	2018
Livestock genomics and bioinformatics	2019
Bioinformatics for biologists	2021
Geographic information systems (GIS) to livestock	2019

Professional Competence (18.6 credits)	
Searching and organising literature	2018
English for academic writing	2019
Writing grant proposals	2022
Third-generation genomics and bioinformatics fellowship at ILRI	2021

Presentation Skills (3 credits)	
Ethiopian Society of Animal Production conference (oral)	2019
European Association of Animal Production conference (poster)	2020
Wageningen Institute of Animal Sciences conference (oral)	2021
World Congress on Genetics Applied to Livestock Production (oral)	2022

Teaching competences (2 credits)	
Supervising MSc thesis student (1 times)	2018

Total credits	43.0

Acknowledgements

Several people have helped me to finalize my four-year PhD project. Without their tremendous support, I would not have successfully completed my studies and developed personally during those years. First, I would like to thank my promotor Prof. Hans Komen for the trust he put on me, together with Dr John Bastiaansen, my co-promotor, since our first meeting in Addis Ababa when they considered me as their PhD student. Unlike many other students, I did not have sufficient financial resources at the start of the project to implement my research activities. I only overcame the challenges through his kindness. Hans provided great leadership that enabled me to work on high quality data, to collaborate with more researchers, and to attended prestigious international conferences. These processes have immensely boosted my self-confidence. Hans, I thank you for all the opportunities you opened for me. You always fascinated me with your out-of-the-box ideas and reasoning. I have learnt from you the importance of having diverse and higher-level perspectives in doing science. The large number of PhD students preceding me who worked in Africa and Asia on different species witness you are genuinely interested in tropical smallholder livestock systems. Your insights and wisdom always brought me back whenever I lost focus. I thought that I knew what I wanted to achieve when I initially developed my PhD proposal. Ironically, it was when I was about to finish the four years that I fully recognized what I was doing and why. I also thank you for your concern regarding my future career and for exploring opportunities toward my intellectual development after completing my PhD.

Secondly, I thank Dr John Bastiaansen for devoting a considerable amount of his time with me as my daily supervisor and mentor. John, I admire your open-mindedness in our quest for novel analytical approaches. Your Socratic questioning during our regular meetings helped me to critically evaluate and re-evaluate my research questions, experimental methods, and interpretations. At the same time, I have enjoyed my intellectual freedom to contemplate ideas on my own. Your supervisory approach was more as a colleague and less pedantic. I thank you for giving me a space to think, plan, and implement my activities as a sandwich PhD student who started his studies a after some years of work experience. I also thank you for your empathy and support whenever I encountered technical and administrative hurdles. Thinking about you reminds me about the importance of storytelling in academic writing. You have taught me how to write more coherent paragraphs and produce well-structured manuscripts. I have also learnt how to be a good academic supervisor.

This brings me to another important person who immensely contributed to my study, Dr Tadelle Dessie. Tadelle used to be my supervisor when I was a research officer at the International Livestock Research Institute (ILRI). He continued with his supervisory role during my PhD research. Tadelle collaborated closely with Hans to put in place all the technical and administrative arrangements needed to start and successfully implement my PhD project. He put the African Chicken Genetic Gains (ACGG) programme dataset at my disposal and allowed me to explore it further. Tadelle is my role model regarding research for development aiming at improving the livelihoods of smallholder farmers. Tadelle, I thank you for your encouragement that made me start my PhD study at a time when I was about to fall into complacency. Without your all-rounded support, I probably may not have embarked on this pleasant journey and successfully finished it. Your actions have been more fatherly than anything else.

My other mentor at ILRI, Prof. Olivier Hanotte has been such an inspiration. I thank him for the knowledge and insights he shared with me since the inception of my PhD project. Olivier has always been passionate about tropical livestock genetics and sets me and fellow PhD students at ILRI on the path to unravelling the evolutionary roots of African livestock. Working with him as a student has been greatly rewarding. Olivier, I also cherish your motto that we all shall try to continuously improve our personalities intellectually and as human beings. I thank you for your great mentorship in the last four years.

Let me return back to my other colleagues in Animal Breeding and Genomics (ABG), Wageningen. I kept my samples for months when Dr Richard Crooijmans came to Addis Ababa to give a training. Richard looked at my chicken blood samples which were stored in a freezer for weeks. Considering the large number of samples which will take too many hours for manual processing, he suggested for an automated extraction of the DNA. Richard, I am highly grateful for the technical and administrative support you and your lab members have provided me in shipping and processing my samples. I also thank you for all your contributions to the success of my project.

My other colleagues at ABG who assisted me in finalizing my project are Dr Martijn Derks and Carolina Pita Barros. I do not know how I would have gone through all the bioinformatic pipelines without your support. I realize that the field of bioinformatics is perplexing for someone like me, with a background in animal science. I thank you for your guidance in the analyses and re-analyses of our data. Martijn, I also thank

for your novel research ideas and approaches which kept me stimulated. I look forward to exciting opportunities of continued collaboration.

I am thankful for the administrative support I received from my colleagues at ABG (Fadma Had-Dadi, Rosilde Dijkhof, Lisette Bourquin, and Marielle Hoogstede) in the last four years. I am especially grateful for your support during the shipment of my biological samples from Addis Ababa to Wageningen and during the planning of my PhD defense. You tried your best to make my stay in ABG and The Netherlands as comfortable as possible. I also appreciate the administrative support I received from colleagues in ILRI's LiveGene programme (Wubalem Dejene, Lidya Abebe, and Kelem Mamo) and Capacity Development Office (Willington Ekaya, Jennifer Kinuthia, and Lea Symekher). Following up through the graduate fellowship contracts and transferring my stipend and other payments timely to different locations was a lot of work. I also miss the time I spent at ILRI with Dr Setegn Worku as a colleague and a friend who was genuinely interested in my academic endeavours and filled me with positive spirits. I enjoyed the time I spent with fellow PhD students at ILRI (Dessalew Tadesse, Agraw Amane, Gebru Gebreslassie, Endashaw Terefe, Endashaw Jebessa, Tsadkan Zegeye, Sumaya Kambal, Maria Lozano Jaramllo and Shumuye Belay), and at ABG (Xiofei Yu, Samuel Mengistu, Renzo Bonifazi, Ibrahim Jibrila, Priadi Setyawan, Benan Gulzari, Pauline Jehannet, and Chantal Roozeboom) and other colleagues at ILRI (Kumlachew Geremew, Mulugeta Yitayih, Wondmeneh Esatu and Mekonnen Girma). We had lots of nice chats and coffee. I specially thank my paranymphs Farid Aththar and Wisdom Agbeti for their help in the final stretch of my PhD; you guys came to the rescues at a time when I was exhausted.

I cannot thank enough Nego Hurma and Million Abebe (veterinary technicians) and Michael Temesgen and Anteneh Hailemariam (drivers at ILRI) who supported me in my field work in Ethiopia during times of high political uncertainty. We repeatedly put our lives at risk in search of chicken samples☺ You are the only people who really understand and remember what we went through on the Ethiopian highlands and lowlands. We drove for thousands of kilometres and trekked on foot for miles across the twenty-six villages in the four regional states, sometimes without food and water. It sounds like a miracle that we all returned safe, and the PhD project has come to completion without a serious problem. I thank you for your commitments and for not giving up. I also thank all the agricultural experts and farmers who assisted me in my sample collection and wished that something meaningful will come out of my research.

Colophon