UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 732

Genetic and Genomic Studies in Chicken

Assigning Function to Vertebrate Genes

JONAS ERIKSSON

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2012

ISSN 1651-6206
ISBN 978-91-554-8246-6
urn:nbn:se:uu:diva-162597

Dissertation presented at Uppsala University to be publicly examined in B42, BMC, Husargatan 3, Uppsala. Friday, February 3, 2012 at 09:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English.

Abstract
Eriksson, J. 2012. Genetic and Genomic Studies in Chicken: Assigning Function to Vertebrate Genes. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine* 732. 55 pp. Uppsala. ISBN 978-91-554-8246-6.

A major challenge in the post-genomic era is to understand how genome sequence variants (genotype) give rise to the enormous diversity observed in terms of morphology, physiology and behavior (phenotype) among living organisms. Domestic animals—with their tremendous phenotypic variation—are excellent model organisms for determining the relationships between genotype and phenotype. In this thesis, I describe the utilization of the chicken, in combination with modern genetic and genomic approaches, in developing our understanding of the genetic mechanisms underlying phenotypic variation. These studies provide novel information on the genetics behind variation in carotenoid- and melanin-based pigmentation—observed in many organisms—and also cast light on the genetic basis of chicken domestication.

In paper I, we report that the yellow skin phenotype—observed in most commercial chickens—is caused by one or several tissue-specific mutations altering the expression of *beta-carotene oxygenase 2* (*BCO2* or *BCDO2*) in skin. In addition, we present the first conclusive evidence of a hybrid origin of the domestic chicken, since the allele causing yellow skin most likely originates from the grey jungle fowl (*Gallus sonneratii*) and not from the previously described sole ancestor, the red jungle fowl (*Gallus gallus*).

In paper II, we detect a number of loci that were likely important during the domestication process of chicken and the later specialization into meat (broiler) and egg (layer) producing lines. One of the major findings was that worldwide, almost all domestic chickens carry a missense mutation in *TSHR* (*thyroid stimulating hormone receptor*) in a position that is completely conserved amongst vertebrates. We speculate that this "domestication-mutation" has played an important role in the transformation of the wild red jungle fowl ancestor into the modern domestic chicken.

In paper III, we demonstrate that the dilution of red (pheomelanin) pigmentation—observed in the plumage of the Inhibitor of Gold chicken—is caused by a frame-shift mutation in the *catechol-O-methyltransferase domain containing 1* (*COMTD1*) gene. The production and regulation of pheomelanin is poorly understood and this discovery advances our current knowledge of this pathway.

Keywords: Chicken, BCO2, TSHR, COMTD1, Phenotypic variation, Domestication, Selective sweeps, Pigmentation

Jonas Eriksson, Uppsala University, Department of Medical Biochemistry and Microbiology, Box 582, SE-751 23 Uppsala, Sweden.

ISSN 1651-6206
ISBN 978-91-554-8246-6
urn:nbn:se:uu:diva-162597 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-162597)

To my family

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I **Eriksson J.**, Larson G., Gunnarsson U., Bed'hom B., Tixier-Boichard M., Strömstedt L., Wright D., Jungerius A., Vereijken A., Randi E., Jensen P., Andersson L. (2008) Identification of the *yellow skin* gene reveals a hybrid origin of the domestic chicken. *PLoS genetics* 4(2): e1000010.

II Rubin C.J., Zody M.C., **Eriksson J.**, Meadows J.R., Sherwood E., Webster M.T., Jiang L., Ingman M., Sharpe T., Ka S., Hallböök F., Besnier F., Carlborg O., Bed'hom B., Tixier-Boichard M., Jensen P., Siegel P., Lindblad-Toh K., Andersson L. (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464: 587-591.

III **Eriksson J.**, Hellström A.R., Rubin C.J., Wang C., Shumaila S., Kerje S., Gourichon D., Bed'hom B., Tixier-Boichard M., Leif Andersson. (2010) A frameshift mutation in *COMTD1* specifically dilutes pheomelanin pigmentation in chicken. *Manuscript*.

Contents

Introduction 9
Mutations – the Source of Genetic Variation 10
Functional Categorization of Mutations 11
Identification of Sequence Elements Regulating Gene Expression 13
Domestication 14
Domestication of the Chicken 16
Pigmentation 18
Carotenoid-Based Pigmentation 19
Melanin-Based Pigmentation 20
The Chicken Genome 21
Methods for Genetic Mapping 22
Introduction 22
Linkage Analysis 23
Identical-By-Descent Mapping 23
Mutation Detection 25

Aims of this thesis 27

Present Investigations 29
Mapping Monogenic Traits in Chicken 29
Background 29
Results and Discussion 31
Signatures of Selection in the Chicken Genome 37
Background 37
Results and Discussion 38

General Discussion and Future Perspectives 43

Acknowledgements 47

References 49

Abbreviations

ASIP	agouti signaling protein
BCDO2	beta-carotene oxygenase 2
BCO2	beta-carotene oxygenase 2
bp	base pair
cAMP	cyclic adenosine monophosphate
cM	centi-Morgan
COMT	catechol-O-methyltransferase 1
COMTD1	catechol-O-methyltransferase domain containing 1
DNA	deoxyribonucleic acid
ENCODE	Encyclopedia of DNA Elements
Gb	gigabase (10^9 bases)
HPLC	high-performance liquid chromatography
IBD	identical by descent
indel	small insertion and deletion
kg	kilogram
kb	kilobase (1000 bases)
LD	linkage disequilibrium
LOD	logarithm of odds
Mb	megabase (10^6 bases)
mtDNA	mitochondrial DNA
RNA	ribonucleic acid
SNP	single nucleotide polymorphism
TSHR	thyroid stimulating hormone receptor
TYR	tyrosinase
TYRP1	tyrosinase related protein 1
TYRP2	tyrosinase related protein 2
QTL	quantitative trait locus

Introduction

Charles R. Darwin combined geological, geographical and biological observations together as support for his theory that all organisms on planet Earth share a common ancestry and that organisms are constantly evolving through natural selection (*1*). Throughout his published work, Darwin used the process of domestication of animals and plants as a proof-of-principle that natural selection exists and acts (*1, 2*). He argued that if man through selective breeding (artificial selection) has altered the morphology, physiology and behavior (phenotypes) of the domesticates compared to their wild ancestors, then the selective forces of the environment must achieve similar changes in wild organisms (natural selection), however at a much slower rate. Darwin was so amazed by the phenotypic variation observed in domestic animals that he even devoted some of his precious time writing a book about the subject: *The variation of plants and animals under domestication* (1868).

Darwin could never explain how the variation in traits arose or how the variation was inherited from parents to their offspring. The publication "*Versuche über Pflanzenhybriden"* in 1866 by Gregor Mendel gave some answers to the latter question. However, it did take almost more than half a century until Mendel's laws of inheritance were combined with Darwin's theories of natural selection to form the modern theory of evolution, known as Neo-Darwinism. A number of landmark discoveries, such as the structure of DNA (*3*) and the unearthing and deciphering of the genetic code, were required until we had a clear picture that the variation in phenotypic traits is caused by mutations in the DNA sequence that makes up our heritable genetic material, the genome.

In the year 2000, the sequence of the human genome was presented by two separate efforts; private and public consortiums directed by Craig Venter and Francis Collins respectively. The analyses of the genome sequences were published the year after in the scientific journals Nature and Science (*4, 5*). This has been followed by a number of generated genome sequences for different animal and plant species in the last ten years, with an enormous increase in the recent past as a consequence of the launch of new DNA sequencing technologies. The current sequencing technologies—massively parallel sequencing or next-generation sequencing technologies—can read ~250 billion bases in a week, compared to 5 million bases in the same time period in the year of 2000 (*6*). The big challenge for contemporary molecular geneticists is thus not to generate DNA sequences, but to decipher the ge-

nome sequence—i.e. to understand how the genetic material determines how we appear and function which is also known as the genotype-phenotype relationship.

Domestic animals are excellent model organisms for exploring genotype-phenotype relationships due to the enormous phenotypic variation among them. No other experimental organism displays so much variation in phenotypic appearance in comparison to domestic animals. Information about the genetic mechanism underlying a phenotypic trait in domestic animals—leading to the identification of the causal gene—can shed light on the function of the human orthologous genes. Humans do in fact share a lot of genes with domestic animals, and homologous vertebrate genes do often exert the same function.

In this thesis, I will present two papers where we have used the chicken as a model to identify genes that control variation in carotenoid- and melanin-based pigmentation (Papers I and III). The two identified genes are excellent candidates for explaining the variation in carotenoid- and melanin-based pigmentation that is observed in many other vertebrate species. In paper II, we utilize newly emerging technological breakthroughs in DNA sequencing to identify genomic regions that may have been under selection during both domestication and the subsequent specialization of the chicken. This helps to decipher the genetic basis of chicken domestication and the following specialization of egg and meat producing chicken lines.

Mutations – the Source of Genetic Variation

Mutation and selection are the driving forces of evolution. Mutation is the random process that increases variation, while selection primarily decreases the levels of variation in the population (*7*). Mutations are changes in the DNA sequence—or RNA in the case of some viruses—that arise during DNA replication, recombination of chromosomes in meiosis, and also through genetic transposition. Mutations can be categorized on the basis of how they affect the DNA as: single nucleotide polymorphisms, small insertions or deletions, structural variation and transpositions. Single nucleotide polymorphism(s) (SNPs) are the most common type of genetic variation, and are associated with errors made by polymerase during genome replication. Due to their high genomic abundance and ease to score, SNPs have become the primary genetic marker used for genetic mapping and comparative analyses. Small insertions and deletions (indels) arise also primarily during genome replication, whereas structural variants (large insertions and deletions, duplications and inversions) are errors made in the recombination process during meiosis. The last type of mutation, transposition, is caused by "jumping genes" named transposable elements (TE) (*8*).

Functional Categorization of Mutations

The majority of mutations that arise do not have a functional consequence. Those that can have a functional effect can be categorized as to whether they affect the biochemical properties of the encoded protein or mature RNA: coding mutations; or alter gene expression: *cis*-regulatory mutations (Figure 1) (*9*). Coding mutations fall within the exons of genes and alter the amino-acid sequence of the encoded protein or the nucleotide sequence of the mature RNA (*10*).

Cis-regulatory mutations occur in sequence elements of DNA that regulate the transcription of genes, namely *cis*-regulatory elements. The *cis*-regulatory elements—e.g. promoters, enhancers, silencers, and insulators—are scattered throughout genomes and via interaction with *trans*-acting factors (i.e. transcription factors) determine the timing and location (spatiotemporal) of gene transcription (*11*). The spatiotemporal transcription of a specific gene is under the control of a number of *cis*-regulatory elements, but primarily those in close proximity to the gene, termed the *cis*-regulatory module (CRM) (*11*). The regulation of gene expression—at the transcriptional level—via the interaction between *cis*-regulatory elements and *trans*-acting factors, is fundamental to the differentiation of various cell-types and therefore the development of complex organisms (*11*).

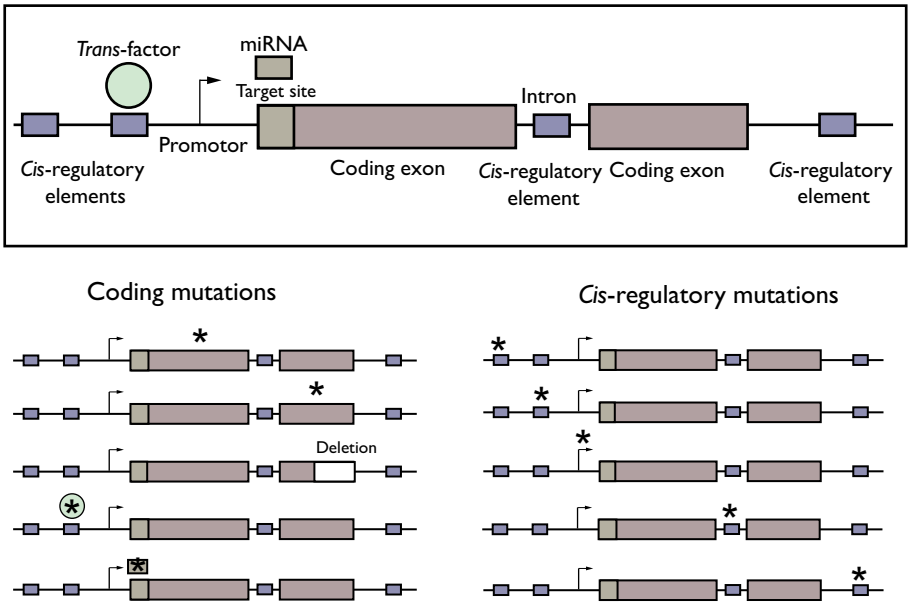

Figure 1. A figure depicting a few types of coding and *cis*-regulatory mutations. An asterisk (*) represents the location of the mutation. *Cis*-regulatory elements can be located within, upstream or downstream of the regulated gene. The mutations occurring in the genes encoding a miRNA or *trans*-factor may have secondary *trans*-acting effects on gene expression, but they are still considered to be coding mutations since they are located in coding sequences. Modified from Coyne and Hoekstra (2007) (*12*).

Mutations in *cis*-regulatory elements have been suggested to play an important role in the evolution of phenotypic variation (*13*). King and Wilson (1975) observed that human and chimpanzee showed surprisingly few protein sequence differences, and proposed that those alone could not account for the phenotypic differences evident between the species. With the notion of transcriptional regulation in mind—discovered by the bacterial geneticists Jacob and Monod (1961) (*14*)—they suggested that mutations that affect the regulation of gene expression could account for the majority of the phenotypic differences observed between human and chimpanzee (*15*). A more recent argument is that a *cis*-regulatory mutation is expected to be less pleiotropic—affect fewer phenotypes—compared to a coding mutation, due to the modular organization of the regulatory regions (*13*). A mutation in an enhancer can, for instance, alter the expression of a gene in one tissue but leave expression unaltered in other tissues. In contrast, a functional coding mutation affects the function of the protein in all tissues in which it is expressed. This could lead to negative pleiotropic effects on many traits and such mutations would thus be the object of purifying selection and perhaps eliminated from the gene pool.

In paper I, we demonstrate that the yellow skin phenotype is caused by one or several mutation(s) in *cis*-regulatory element(s) that alters the expression of the *beta-carotene oxygenase 2* gene (*BCDO2* or *BCO2*) in skin but not in liver. A complete loss of function of *BCO2* would most likely result in negative pleiotropic effects on other traits, since the BCO2 protein has been shown to play a protective role against carotenoid-induced mitochondrial dysfunction (*16*).

Nevertheless, most of the mutations underlying phenotypic variation that have been found to date have been coding (*10*). This can partly be explained by the fact that once the gene controlling the trait has been determined, it is much easier to identify a coding mutation than a regulatory mutation. The genetic code makes it easy to predict if a mutation alters the function of the protein by searching for non-synonymous, frameshift or nonsense mutations (*10*). Mutations affecting transcription are much harder to identify due to our limited knowledge regarding the location and function of *cis*-regulatory sequence elements. Functional and biochemical experiments are also often required to confirm *cis*-regulatory mutations (*9*).

Identification of Sequence Elements Regulating Gene Expression

Comparative genomics—the comparison of genome sequences from different organisms with the aim to identify DNA sequences that are under evolutionary constraint and thus predicted to confer a biological function—has proven to be a successful approach for predicting the genomic location of sequences that control gene expression (*17*). The genome sequence comparison between human and mouse revealed that more than 5% of the human genome has been under purifying selection during the past 100 million years and is thus probably functional (*18*). Surprisingly, only ~1.5% of the human genome is protein coding, which implies that the majority of evolutionary conserved sequences maintain other functions, such as gene regulation (*18*). The study by Lindblad-Toh et al. (2011), which compared the genomes of 29 placental mammals, confirmed that ~5% of the human genome has undergone purifying selection. In addition they located constrained elements covering 4.2% of the human genome—of which more than two thirds were found to reside within non-protein-coding regions—and experimental data suggested that a large proportion of these are *cis*-regulatory elements (*19*).

Comparative analysis of genome sequences from species within the same animal lineage can be used to identify sequences that are of functional importance for that specific animal lineage. For instance, the comparison of the chicken, turkey and zebra finch genome sequences revealed that since they shared a common ancestor, 8.6% of the chicken genome has been under

purifying selection (*20*). Some of these constrained sequences may be important in maintaining a number of traits that are specific for birds.

ENCODE, the Encyclopedia of DNA Elements, is a major research effort with the aim to identify all functional elements in the human genome through a combination of comparative genomic and experimental studies (*21*). In 2007, only 1% of the human genome had been investigated, but the effort continues and new data are released continuously (*21*). The above-exemplified efforts in annotating genomes for functional elements will greatly simplify the detection of mutations that affect gene expression.

Domestication

Domestication refers to the human-directed process of making wild plants and animals adapted to new environments and human needs. This adaption has been achieved through selective breeding, or artificial selection, and has resulted in an enrichment of mutations that control the phenotypes desired by man (*22*).

The early domestication of animals was most likely unintentional, where individuals carrying genetic variants predisposing them to be less afraid of human approach were likely rewarded with food and protection (*23*). This was later followed by intentional human selection for tame animals and other desirable phenotypes, such as coat color variants (*24, 25*). Other typical phenotypic changes associated with animal domestication are as listed in the review article by Jensen and Andersson (2005):

I External morphological changes (e.g. altered body size, fur and feather color and growth pattern)
II Internal morphological changes (e.g. reduced brain size)
III Physiological changes (e.g. altered reproductive cycle and endocrine response)
IV Developmental changes (e.g. earlier sexual maturity)
V Behavioral changes (e.g. reduced fear, increased sociability and reduced antipredator responses)

Many of these phenotypes frequently occur together as a block in numerous different domesticated species, suggesting that they may represent general adaptations to captivity and domestication (*26*). Belyaev and colleagues showed in their classical selection experiment on silver foxes (*Vulpes fulvus*) that many of the domestication phenotypes—as described above—could be achieved by simply selecting against aggression (reviewed in Trut, 1999). The reduction in aggression achieved after a few generations of selective breeding was accompanied by a number of physiological and morphological changes similar to the domestication phenotypes, such as coat color altera-

tions, dropping ears, shorter legs and tail, earlier sexual maturity, longer mating seasons and alterations in skull morphology (*27*).

The rapid change in traits observed by Belyaev and colleagues', together with the catalogue of shared domestication phenotypes among domestic animals, have led some scientists to speculate that there are so-called "domestication genes" with large pleiotropic effects and that these created the domestication phenotypes (*28*).

It has been proposed that thyroid hormones play an important role in animal domestication (*29*). Susan J. Crockford—the main advocate of this theory—came to this conclusion after observing the similarities between domestication phenotypes and those associated with hypothyroidism (*29*). Crockford suggested that only a few changes of the regulation of thyroid hormones during development, would be required to achieve many of the phenotypic changes associated with the process of domestication. In paper II of this thesis, we present data suggesting that the thyroid stimulating hormone receptor (TSHR)—a key player in the regulation of thyroid hormone expression—has played an important role in shaping the domestic chicken.

As described above, the domestication of animals has led to dramatic changes in the phenotypic appearance of these individuals compared to their wild ancestors (*30*). This long (~10 000 years) and intense selective breeding has also created an enormous range of phenotypic variation within the domesticated species. The majority of these selected mutations have a favorable phenotypic effect on the selected trait (e.g. egg production or coat color) but rarely a deleterious effect on other traits (*22*). Mutations with deleterious effects are most often eliminated from the domestic animal breeding pool, and so very few remain in the populations at large. The dog may be considered as an exception to that general rule, since selective breeding for specific characteristics has resulted in pathological outcomes—either due to pleiotropism or linked deleterious mutations—that have been accepted by the breeders (*31, 32*).

The domestication of plants and animals is considered to be the most important development during the last 10 000 years of human history (*33*) and was a prerequisite for the rise of civilization. Maintaining animals and plants under our provision at the farm instead of hunting and gathering our food, led to a population explosion due to the increased accessibility to food (*33*). Permanent settlements could be established due to the decreased need for migration to follow the seasonal shift in wild food supplies. It is no coincidence that inventions such as the wheel, metal tools, writing and advanced social systems arose in the same areas as early domestication (*33*).

Domestication of the Chicken

The domestic chicken belongs to the genus of *Gallus* that comprises the four wild species of jungle fowls: red jungle fowl (*Gallus gallus*), grey jungle fowl (*G. sonneratii*), Ceylon jungle fowl (*G. lafayetii*) and green jungle fowl (*G. varius*) (*34*). All four jungle fowl species inhabit different geographical regions of South Asia and exhibit quite extensive differences in morphology (Figure 2).

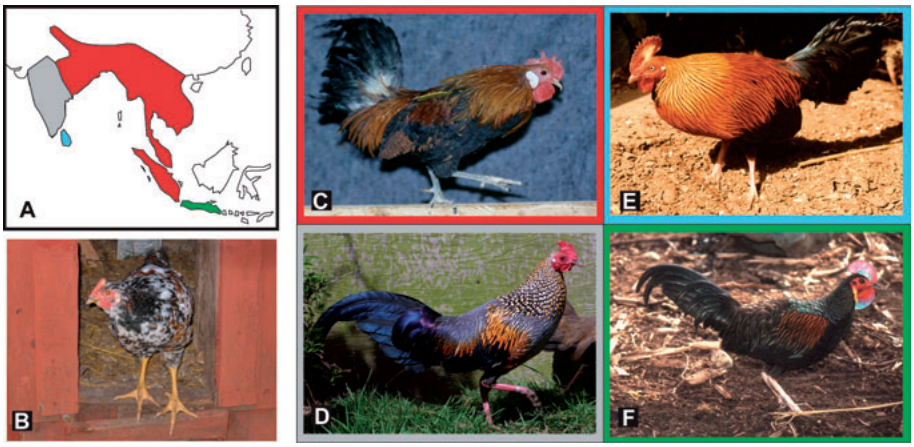

Figure 2. Panel A depicts a map of South Asia onto which the ranges of the four species of jungle fowl are drawn. Panel B depicts a European domestic chicken. Red, grey, blue and green regions represent the respective ranges of red, grey, Ceylon and green jungle fowls. Images of these birds are presented in panels C through F respectively, within colored borders that correspond to the colors on the map. The figure was adapted from Eriksson et al (2008).

The origin of the domestic chicken, in terms of geographical origin and ancestry, has been an issue of debate since Darwin first proposed the Indian red jungle fowl (*Gallus gallus*) as the single ancestor of all domestic chicken, and thus a monophyletic origin (*2*). He based his theory on the marked morphological similarities between the domestic chicken and the red jungle fowl, and the fact that crosses between the two generated fertile offspring, while crosses between domestic chickens and the three other jungle fowl species resulted in low survival of the chicks (*2*).

The alternative hypothesis for the ancestry of chicken—the multiple-species or polyphyletic origin—was first proposed by F.B. Hutt in his book *Genetics of the Fowl* (1949). He claimed that potentially all four wild species of jungle fowl may have contributed material to the genetic makeup of the domestic chicken. He mainly based his hypothesis on the fact that a number of phenotypes observed in domestic chickens are not found in the red jungle fowl population, but are present in the other species of jungle fowl (*35*).

Early molecular genetic studies of mitochondrial DNA (mtDNA) supported the initial view that the red jungle fowl was the sole ancestor of the chicken (*36, 37*). However, the study by Nishibori et al. (2005)—which compared sequences of mtDNA and five autosomal loci from domestic chickens and all four jungle fowl species—suggested that two additional species of jungle fowl (grey and Ceylon jungle fowl) might hybridize with the domestic chicken (*38*). Nevertheless, this study did not provide an answer to the question of whether they had contributed to the domestication of chicken. In paper I of this thesis, we present the first conclusive genetic evidence supporting the theory of a hybrid or polyphyletic origin of the domestic chicken.

The earliest archeological findings of chicken remains—bones larger than the red jungle fowl—were found in Southeast Asia and were estimated to be almost 8000 years old (*39*). Additional archeological findings—bones and art objects depicting chickens in the Indus Valley dated to about 2500 B.C.—suggested that there have been multiple domestication events of the chicken (*34*). The multiple geographical origin of the chicken is also supported by molecular genetic data that further implies that the majority of European and Middle Eastern domestic chickens originate from India (*40, 41*). Further molecular studies present evidence that American, Oceanic and African chickens also originated from India, suggesting that India was the original platform for the worldwide dispersal of chicken (*42-45*).

Chicken was most likely initially domesticated for mainly cultural reasons such as religion, art and entertainment and much later as a source of food (*34*). This would consequently imply an early selection for increased body size, an important feature in cockfighting, and for plumage variants, used as decorative art. Most of the hundreds of chicken breeds currently in existence were formed during the 1800s, when selection for fancy characters—such as plumage and skin color variation, amount of feathering, size of various external organs—became high fashion among the royalties and the upper classes in Europe and America (*34*). Many Asian breeds were at that time imported to Europe, and very often hybridized with European breeds to create new stocks with fascinating characters (*34*).

It was first during the 20th century when more systematic breeding schemes focused on production traits were initiated (*34*). Chicken lines were created with the sole task of producing either eggs or meat—to overcome the negative genetic and phenotypic relationship between reproduction and growth (*46*). This has resulted in extreme differences in reproduction between the modern egg-layer chicken and its main wild ancestor, the red jungle fowl. The modern layer used in commercial egg production produces more than 300 eggs per year, while a pure red jungle fowl lays around 5 eggs once per year (*47*). The same extreme difference is observed in body weight between the modern commercial broiler and the red jungle fowl, 4-5 kilogram (kg) versus 1 kg respectively (*34*). This great response to selective

breeding for egg and meat production has thus been achieved by sophisticated breeding programs, and is made possible due to the high genetic variability and large effective population sizes within the domestic chicken populations (*46*).

In paper II we, utilize massively parallel sequencing technology to search for haplotypes in the chicken gene pool, which are thought to be responsible for some of the phenotypic changes that have taken place during the domestication process, and the subsequent selection for specialized breeds used for food production.

Chicken as Model Organism

The chicken has a long and proud history as a model organism. It was the first animal in which the Mendelian inheritance of traits was demonstrated (*48*), and scientific discoveries made with the chicken as an experimental system in developmental biology (*49*), virology (*50*) and immunology (*51*) have been awarded with Nobel Prizes.

The chicken possesses a number of features which make it particularly suited to the study of genotype-phenotype relationships. Firstly, there is a great diversity in phenotypes—such as plumage and skin color, body composition, production traits—providing material for dissecting the genetic mechanisms underlying phenotypic variation. Secondly, the chicken is easy and cheap to breed, permitting the construction of large pedigree material which segregates for the trait of interest (*46*). Third, the chicken genome has a high recombination rate compared to many mammalian species (*46, 52*). Fourth, many mutations underlying traits are shared between chicken breeds due to gene flow and the recent history of breed creation (*22*). Fifth, there is high genetic diversity within the chicken gene pool (*53, 54*). The last four points provide excellent conditions for powerful genetic analyses to determine genotype-phenotype relationships.

Identifying the genes underlying some of the enormous variation in chicken phenotypes is of great importance since it can be used to improve the future breeding of domestic animals and also as it can provide information about the function of orthologous genes in vertebrates.

Pigmentation

One distinctive feature of all domestic animals is the huge variation in coloration among and within these individuals compared to their wild ancestors (*55*). Selective breeding, since the early days of domestication (*24, 25*), has achieved this phenomenon. The reason for that phenotypic selection might have been to reduce camouflage in order to facilitate animal husbandry, and/or to make the domesticated forms distinct from their wild ancestors, or just as a human preference for novelty (*24*). The mutations controlling pig-

mentation variation were either present in the wild species prior to domestication or arose during domestication (*55*).

Pigmentation serves many different functions in animals, such as camouflage, mimicry, intraspecific communication, and protection against ultraviolet radiation (UVR) (*56, 57*). The mouse is the prime organism for the identification of genes controlling pigmentation due to the extensive mutagenesis projects that have generated a tremendous number of pigment mutants. To date (November, 2011), 171 genes have been identified which influence pigmentation in the mouse in various ways, and more than 200 loci remain to be mapped (*58*). Many of these genes have also been reported to affect pigmentation in other vertebrate species, suggesting a high degree of conservation for the genes and molecular pathways controlling pigmentation in vertebrates (*55*).

The chicken has also proven to be an excellent model organism for understanding the genetics of pigmentation—illustrated by the number of pigmentation genes mapped in chicken (*59-68*). The identification of genes underlying plumage and skin color variants in the domestic chicken will provide new insight about basic pigmentation biology.

Carotenoid-Based Pigmentation

Carotenoids are pigments—yellow, orange or red in color—that are synthesized by plants and microorganisms where they, for instance, facilitate photosynthesis and in plants protect chlorophyll from photo-damage (*69*). Animals lack the ability to synthesize carotenoids and must therefore acquire them from their feed (*69*). Carotenoids serve several important functions in vertebrate physiology. They act as precursors to vitamin A and retinoic acid (*70*); anti-oxidative molecules (*71*); and photoprotectors of the human retina (*72*).

Carotenoids are also important as pigments and add coloration to the tissue/organ where they are deposited, as observed in insects, fish, reptiles, birds and mammals (*69, 73*). For instance, carotenoids are responsible for the color of the egg yolk, lobster shell, salmon flesh and the flamingo plumage (*69*). In birds carotenoid-based pigmentation is known to act as an honest signal of the condition of a male carrier and functions as a sexually selected trait (*74, 75*). There is a trade-off between the allocation of carotenoids for colorful ornamentation versus the maintenance of carotenoids for important physiological processes—e.g. as antioxidants in the immune system (*74, 75*)—and only males in a good physical condition can afford to allot carotenoids to ornamentation.

The genetic mechanisms involved in the metabolism and deposition of carotenoid pigments in animals were elusive until recently. In paper I, we present the first genetic mechanism underlying variation in carotenoid-based pigmentation in animals. We conclude that a decreased expression of the

enzyme BCO2 (beta-carotene oxygenase 2) in the skin—allowing the depositions of colorful carotenoids—causes the yellow skin and beak phenotype in the domestic chicken.

Melanin-Based Pigmentation

The melanocyte is the cell type responsible for the synthesis of melanin pigment in birds and mammals. Two types of melanin—the brown/black eumelanin and the red/yellow pheomelanin—are synthesized in the melanocyte-specific organelle, the melanosome (*57*). The melanin packed melanosome is later transferred to surrounding keratinocytes in the skin, eye, hair or feather, where it provides color and photoprotection (*57*). It is the number, size, cellular distribution and type of melanosome—pheomelanosome or eumelanonsome—that determines the color of the pigmented organ (*76*).

Research in the field of melanin-based pigmentation biology is extensive, and so I will focus my review on the pathway that affects my work, namely the synthesis of melanin—also know as melanogenesis—or more specifically the production of pheomelanin pigment.

Melanogenesis

The production of melanin pigment is primarily initiated by the interaction of external signaling proteins—e.g. α−melanocyte stimulating hormone (α-MSH) or agouti signaling protein (ASIP)—to the melanocortin-1 receptor (MC1R) that is located in the membrane of the melanocyte (*77*). The interaction between α-MSH and MC1R leads to elevated levels of intracellular cyclic AMP (cAMP) that increases the transcription of the genes encoding tyrosinase (Tyr) and the tyrosinase related protein 1 (Tyrp1) and 2 (Tyrp2/Dct) (*78*). Tyrosinase (Tyr) catalyzes the conversion of L-tyrosine or L-dopa to dopaquinone, which is the precursor molecule to both eumelanin and pheomelanin. The Tyrp1 and Tyrp2 enzymes catalyze reactions in the biosynthesis of eumelanin (eumelanogenesis) with dopaquinone as the substrate—leading to the production of the brown/black eumelanin pigment.

The interaction between ASIP and MC1R leads to lowered intracellular levels of cAMP resulting in pheomelanin synthesis, which is also known as pheomelanogenesis. The biosynthesis of pheomelanin requires the amino acid cysteine in addition to dopaquinone (*79*). But unlike eumelanogenesis, no additional enzymes seem to be required for pheomelanogenesis to proceed: pheomelanin is produced in the presence of dopaquinone and cysteine (*80*). The potential lack of enzymes in pheomelanogenesis could explain why only a scarce number of mice mutants (*dwarf gray, lethal grey, grizzled, subtle grey*) (*58*) displaying a specific dilution or removal of pheomelanin pigmentation have been identified. In contrast, numerous mice with mutations affecting only eumelanin pigmentation have been identified, and of these 56 are caused by a mutated *Tyrp1* gene (*81*).

Dopaquinone—the melanin precursor molecule—and intermediate molecules produced in melanogenesis are known to be toxic to the melanocyte (*57*). To cope with those cytotoxic intermediates, melanogenesis takes place in the closed environment of the melanosome (*57*). Occasionally some leakage to cytosol from the melanosome occurs, and protective mechanisms have evolved to incapacitate the cytotoxic intermediates (*57*). One example of such a mechanism is the O-methylation of the melanogenic intermediates mediated by catechol-O-methyltransferase (COMT), leading to a reduced cytotoxicity of the intermediates (*57*).

In paper III of this thesis, we propose that catechol-O-methyltransferase domain containing 1 protein (COMTD1) is affecting pheomelanin pigmentation either; by the detoxification of cytotoxic pheomelanin intermediates leaked from the melanosome; or by altering the cellular levels of cysteine; or by acting as a regulator of pheomelanogenesis—this is because a frameshift mutation in *COMTD1* is underlying the Inhibitor of Gold phenotype in chicken.

The Chicken Genome

The draft genome sequence of a single red jungle fowl female (UCD 001) and the associated analysis was published in 2004 (*52*). The rationale for sequencing a female bird was in order to generate sequence for both the Z and W sex chromosomes, since in birds the female is the heterogametic sex (*82*). In addition to the sex chromosomes, the chicken karyotype is made up of 38 autosomal chromosome pairs that are divided into 10 large macrochromosome pairs (chromosome 1-10) and 28 small microchromosome pairs (chromosome 11-38) (*82*). The microchromosomes generally have a higher G+C content, gene density and recombination rate compared to the macrochromosomes (*52*). High G+C content is normally associated with low sequence representation in genome assemblies, which is reflected by the lack of sequences assigned to ten of the 28 microchromosomes in the current genome assembly (version 2.1, Washington University).

The recombination rate varies between 2.5 and 21 cM Mb^{-1} across the chicken genome, but is on average higher compared to human (~1 cM Mb^{-1}) and mouse (~0.5 cM Mb^{-1}) (*52*). The number of genes in the chicken genome is estimated at 20 000–23 000 of which 60% have a single human orthologue (*52*). The haploid size of the chicken genome is around $1x10^9$ basepairs (bp), which is almost one third of most mammalian genomes (*83*). This size difference is likely due to lower numbers of interspersed repeats, segmental duplications and pseudogenes (*52*). Small genome size is a feature observed in most bird species and has been proposed to be an adaptation to the higher rate of oxidative metabolism required to meet the energetic demand of flight (*83*).

In order to assess the nucleotide diversity between and within domestic chicken breeds, low coverage genome sequences (0.25 X) from three domestic chicken breeds (a Chinese silkie chicken, a broiler and a layer), were generated by random-shotgun sequencing (*53*). The study identified 2.8 million single nucleotide polymorphisms (SNPs) and the analyses revealed a remarkably high nucleotide diversity between and within breeds; on average 5 SNPs were detected every 1000 base pair (bp) in the pairwise comparisons between and within breeds. This was comparable to the same amount of variation that was observed in the domestic breeds to red jungle fowl comparisons. In conjunction, this suggested that there had been a historically large population size of domestic chicken and that a significant amount of the genetic variation has been captured from the wild ancestors (*53*).

Methods for Genetic Mapping

Introduction

A major challenge in biology and medicine is the assignment of function to the sequences in the genome, with the ultimate goal being to understand the link between genotype and phenotype. This has proven to be very challenging in the case of quantitative traits, to which multiple loci with small effects, environmental effects and epistatic interactions between loci all contribute to the resultant phenotype. Monogenic traits—those that are determined by a single locus—are on the other hand relatively easy to determine the genotype-phenotype relationship of, and there are numerous molecular and bioinformatic tools available that can be utilized for such studies. In paper I and III, we identify the genes responsible for two monogenic traits, using some of the methods that are available for gene mapping.

The new generation of sequence technologies has opened up a number of possibilities for determining genotype-phenotype relationships. For instance it is now possible, and cost-effective, to use massively parallel sequencing and pooled samples to estimate the allele frequencies at all polymorphic loci in different populations (*84*). Bioinformatic analyses of these data sets can reveal genomic regions showing signs of positive selection—as a reduced degree of heterozygosity—in phenotypically similar populations. These regions are likely enriched for adaptive genetic variants that control the phenotypic differences observed between the compared populations. The regions can be used as candidate regions to be further scrutinized by follow-up experiments in order to determine the link between the selected locus and a phenotype.

In paper II, we use bioinformatics—on whole genome sequence data generated from different DNA pools of domestic chicken breeds and the red jungle fowl—to identify genomic regions that most likely have been under

positive selection during the domestication process and the subsequent specialization of meat or egg producing chicken lines.

Linkage Analysis

Linkage analysis is a method used to identify a chromosomal region harboring a gene that underlies a phenotypic trait of interest. In order to perform a linkage analysis three prerequisites need to be fulfilled: (i) a pedigree material segregating for the trait, (ii) genotype data from polymorphic genetic markers and (iii) phenotype data of the individuals in the pedigree material.

Linkage analysis utilizes the "phenomena" of genetic linkage to identify genetic markers that co-segregate with the phenotype of interest in a pedigree material. Genetic linkage refers to the observation that the closer two genetic loci physically are to each other, the greater chance that they will be inherited together—i.e. they will not be separated by recombination during meiosis (*85*). The genetic distance between marker pairs is measured in centi-Morgans and can be directly inferred from the recombination fraction: 1% recombination between markers equals 1 centi-Morgan. The recombination fraction (Θ) between markers is calculated by dividing the number of recombinant gametes with the total number of informative meiosis. Markers are considered linked if the recombination fraction between them is significantly less than 0.5.

When performing linkage analysis, the phenotype is treated as a marker with a specific mode of inheritance, and the aim is to detect genetic markers, of known genomic positions, that are inherited together with the phenotypic trait. To each marker-phenotype combination, a logarithm of odds (LOD) score is assigned to statistically test the chance of obtaining a similar outcome as if the markers were unlinked (Θ=0.5**)** (*85*). A LOD score of 3 is generally considered to be significant evidence of linkage and means that there is a 1 in 1000 chance that the observed linkage would have occurred by chance (*85*).

Identical-By-Descent Mapping

Identical-by-descent (IBD) mapping, or linkage-disequilibrium analysis, is a natural follow-up to linkage mapping and is used to refine the position of a locus controlling a phenotypic trait. IBD mapping has proven to be especially powerful for detecting genes underlying phenotypic traits in domestic animals (*22*).

IBD mapping utilizes the fact that many mutations underlying selected phenotypes within domestic animals have been inherited in a manner termed identical-by-descent (IBD) from a common ancestor and are therefore shared between populations of a domesticated species (*30*). The intense selection for such mutations in domestic animals has created so-called selective

sweeps in populations selected for the same phenotypes. A selective sweep is characterized by a loss of heterozygosity at the selected locus and nearby linked polymorphic loci, due to genetic hitch-hiking (*86*). The size of region affected by a selective sweep is determined by both the local recombination rate and the number of generations it took for the favorable haplotype to become fixed in the selected population (*87*).

IBD mapping aims to detect a minimum haplotype shared across individuals from different breeds selected for the same phenotype (Figure 3). This is usually performed by genotyping a dense set of polymorphic markers in the trait-associated region—across a large set of animals that are segregating for the phenotype of interest—in order to identify a minimum haplotype that is completely associated with the phenotype across breeds. It is also feasible to use the massively parallel sequencing technique on pooled samples to detect selected loci—which we prove in paper II of this thesis.

IBD mapping provides a much higher resolution in trait mapping than linkage analysis, because the size of the associated region is based on many historical recombination events, whereas linkage analysis is based on recombination events occurring within the pedigree and is limited by the size of the pedigree material.

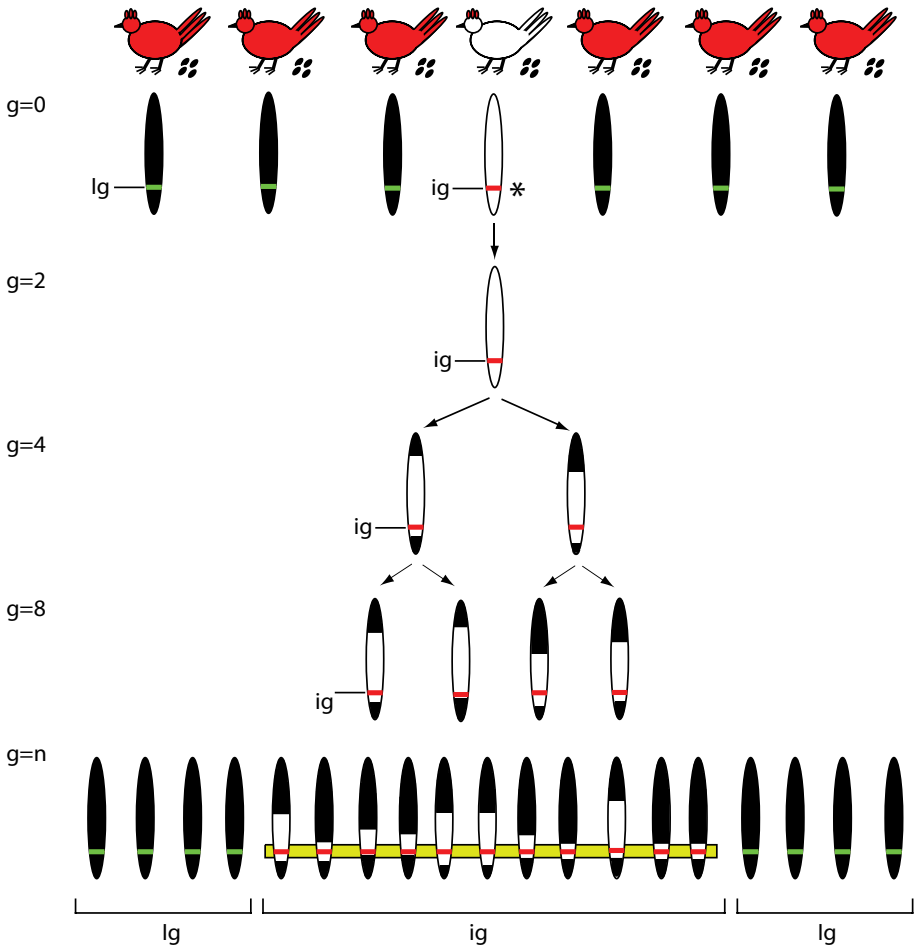

Figure 3. Identical-by-Descent Mapping. An allele (ig) with a phenotypic effect arises by mutation (*) in generation 0 (g=0). The mutant allele (ig) will be in complete linkage disequilibrium (LD) with alleles at other loci on the same chromosome. LD will decay as a consequence of recombination, as the mutant chromosome is passed down through the generations, but will although persist between the mutant allele and alleles at closely linked loci. In generation n, a sample of ig and Ig chromosomes are genotyped to determine the minimum haplotype that is shared among all animals carrying the ig allele. The minimum shared haplotype (yellow box) is assumed to harbor the casual mutation. The figure was kindly provided by Dr. A. R. Hellström who modified it from Andersson and Georges (2004).

Mutation Detection

Mutation detection used to be a daunting task due to the time-consuming processes of primer-design, PCR amplification and nucleotide sequence determination of the trait-associated region. Today, using the latest sequencing technologies provided by companies such Illumina and Life Technolo-

gies, it is possible to generate sequence for the trait-associated region within a week (*88*). The generated sequence reads are mapped against a reference genome sequence and sequence variants (SNPs, indels and structural variants) can be identified using various bioinformatic tools. The detected polymorphisms may then be intersected with exons of protein coding genes and investigated as to how they affect the encoded protein sequence using public software e.g. Variant Effect Predictor (*89*) and SIFT (*90*). Polymorphisms may also be intersected with sequences that have been identified to be under evolutionary constraint—suggesting that they possess a function, e.g. *cis*-regulatory elements. Candidate mutations can, in the case of domestic animals, be further scrutinized by experimentation if they are shared between breeds that display the same phenotype but are absent in those breeds for which the phenotype is also absent.

Aims of this thesis

The overall aim of this thesis was to utilize genetic and genomic methods to develop our understanding of the genetic mechanisms underlying phenotypic variation and consequently add function to vertebrate genes.

The specific aims were to:

- Identify the casual genes for two pigmentation phenotypes, yellow skin (W) and Inhibitor of Gold (IG), in the domestic chicken (Paper I and III).

- Utilize massively parallel sequencing technology to search for signatures of positive selection during chicken domestication, and the subsequent artificial selection for egg and meat producing chicken lines (Paper II).

Present Investigations

Mapping Monogenic Traits in Chicken

Background

The long period of intense phenotypic selection has created an enormous resource of variation in phenotypes within domestic animals. This rich resource is ready to be utilized, with the modern techniques in gene mapping and mutation detection, in order to explore the genetic mechanisms underlying the variation in phenotypic traits. Whereas major efforts—both financial and time—have been put into understanding the genetic mechanisms underlying quantitative traits of agricultural significance in domestic animals (*22*), very few resources have been allocated to study the genes behind the numerous monogenetic traits observed in domestic animals. This is somewhat surprising since considerable amounts of money are spent on mouse mutagenesis programs to develop mutants that are used to assign function to genes (*22*). The great number of monogenetic traits in domestic animals is therefore an underutilized resource with a great potential to assign function to vertebrate genes and to determine genotype-phenotype relationships (*22*).

Numerous genes underlying coat color variation in mouse have been identified (*58*). This is primarily due to the ease of detecting coat color phenotypes in mutagenesis screens, in comparison to many other internal and external phenotypes that require more effort to detect. In addition, pigmentation phenotypes are also easy to score within the pedigree materials that are established to identify the genes underlying pigmentation phenotypes. The chicken has also proven to be an excellent model organism in which to identify genes underlying variation in plumage and skin pigmentation (*59-66*, *68*). In paper I and III of this thesis, we report the identification of two genes—never previously associated with variation in pigmentation—to control the pigmentation phenotypes yellow skin and Inhibitor of Gold in chicken.

Understanding the genetics behind pigmentation in model organisms is, for instance, of great importance when it comes to understanding what goes wrong in human pigmentation disorders and skin cancers.

Yellow Skin Phenotype

The yellow skin phenotype in chicken was one of the first traits in animals described to exhibit Mendelian inheritance, and was at the same time demonstrated to be caused by the recessive autosomal *yellow skin* (*W*Y*) allele at the *W* locus (*48*). Birds homozygous for the *yellow skin* allele display both yellow skin of the legs and body, and also a yellow beak. This is the result of depositions of yellow carotenoids that the birds acquire from their feed (*35*). The dominant *white skin* (*W*W*) allele does not allow depositions of carotenoids in the skin and beak, but does not seem to affect carotenoid depositions in other tissues (*35*). Today yellow skin is a very common phenotype, and nearly all chicken used in commercial meat and egg production in the Western world possess yellow skin—corresponding to billions of chickens.

The red jungle fowl (*Gallus gallus*) has long been considered to be the sole wild ancestor of the domestic chicken (*1, 2*). This was however challenged by Hutt (1949), who proposed that as many as three additional jungle fowl species—the grey (*G. sonneratii*), Ceylon (*G. lafayeteii*) and green (*G. varius*) jungle fowls—may have contributed to the genetic makeup of the domestic chicken. He based his view on the observation that a number of phenotypes in the chicken are either found in the grey, Ceylon, and green jungle fowl, but missing in red jungle fowl (*35*)—including yellow skin. Early molecular genetic studies have supported the view of the red jungle fowl as the sole ancestor (*36, 37*). However a study from 2005 proposed the possibility of hybridization between chicken and grey and Ceylon jungle fowl (*38*), but it did not answer the question as to if they contributed to chicken domestication.

Inhibitor of Gold Phenotype

The Inhibitor of Gold (IG) or Cream phenotype is a relatively common phenotype amongst chicken and is characterized by the dilution of the red pheomelanin pigmentation of the plumage. Taylor first described IG in 1932 to be caused by the autosomal recessive *IG*IG* allele (*91*), which was later confirmed by Punnett in 1948 (*92*). The wild type allele—leaving the pheomelanin pigmentation unaffected—was later denoted *IG*N*.

The regulation of the biochemical pathway leading to the production of the red/yellow pheomelanin pigmentation is considered to be less understood than the production of the black/brown eumelanin pigmentation. Only a few mutations in the genes *Slc7a11*, *Ggt, Ostm1* and *SLC45A2* have been associated with a specific dilution of pheomelanin pigmentation (*60, 93-95*). Identification of the gene responsible for the Inhibitor of Gold phenotype in chicken would thus provide a primer to decipher the biochemical regulation of the synthesis of pheomelanin pigment. This is of great importance since

pheomelanin is thought to be involved in the development of UV-induced skin cancer due to its phototoxic properties (*96*)—however this is under debate (*97*).

Results and Discussion

Paper I – Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken

In paper I, we demonstrate that the yellow skin phenotype is caused by one or several tissue-specific *cis*-regulatory mutation(s) that leads to a reduced expression of *BCO2* (formerly *BCDO2*) in the skin, but not in liver. Surprisingly, we also demonstrate that the *yellow skin* allele most likely originates from the grey jungle fowl and not from red jungle fowl, and thus present the first evidence of a hybrid origin of the domestic chicken.

A combination of linkage analysis and IBD mapping across breeds with the yellow skin phenotype was used to identify the gene responsible for yellow skin. Linkage analysis was performed using a backcross pedigree-material (*W*/*Y*×*Y*/*Y*), comprised of 91 individuals, that was provided by Hendrix Genetic B.V. The *yellow skin* locus—previously assigned to chicken chromosome 24 (*98*)—was found to be in linkage with a SNP on chromosome 24 at nucleotide position 5,237,523 (LOD score=16.4; recombination fraction=6.9%). The *BCO2* gene located in proximity to the linked SNP stood out as a candidate for yellow skin, as *BCO2* encodes the enzyme beta-carotene oxygenase 2 that had been reported to turn colorful carotenoids into colorless apocarotenoids (*99*).

Partial resequencing—followed by SNP genotyping—of the *BCO2* gene in birds representing breeds segregating for *yellow skin*, revealed a 23.8 kilobase (kb) haplotype that all tested yellow skinned chicken were fixed for. The haplotype—assumed to have been inherited identical-by-descent from a common ancestor—spanned the major part of *BCO2* and the putative gene *BX935617*.

Reverse Transcription PCR followed by pyrosequencing, of six heterozygous birds (*Y*/*W*), revealed that more than 90% of the *BCO2* transcripts in skin originated from the *white skin* (*W***W*) allele, whereas in liver, no difference in transcription from the *Y* and *W* alleles was detected. This suggests the presence of one or several *cis*-regulatory mutation(s) affecting the transcription of *BCO2* in skin, but not in liver. Consequently, we postulated that both genotypes take up colorful carotenoids in the skin, but in white skinned birds the carotenoids are degraded to colorless apocarotenoids by the action of the BCO2 enzyme. After this study was published, mutations in *BCO2* have been reported to be associated with increased carotenoid levels in: cow milk (*100, 101*); cow fat (*101*); and sheep fat (*73*).

Sequence comparisons between the *yellow skin* and wild type alleles revealed the surprisingly high sequence difference of 0.81%, which was much higher than the previously reported genome average of 0.5% (*53*). This observation, combined with the notion that both Ceylon and grey jungle fowl have yellow/red legs, led us to sequence the 23.8 kb IBD region in the additional three jungle fowl species. The subsequent phylogenetic analysis revealed that the sequences clustered into two distinct clades (Figure 4). The first comprised of *BCO2* sequences from red jungle fowls and white skinned (*W*/*W*) chickens, while the second clade comprised of sequences from yellow skinned chickens and Ceylon and grey jungle fowls. We concluded that the most plausible explanation to the observation was that the *yellow skin* allele originates from the grey jungle fowl—possibly Ceylon jungle fowl—and was introgressed into the chicken genome sometime during early domestication. Resequencing of the *yellow skin* interval in Chinese Shek-ki chickens—known to express the yellow skin phenotype—revealed that they carried alternative *yellow skin* haplotype that was slightly different to *yellow skin* haplotype found in Western breeds. The Chinese *yellow skin* haplotype did however cluster with the Western *yellow skin* haplotype and *BCO2* sequences from grey and Ceylon jungle fowls. The existence of two different *yellow skin* haplotypes suggested that there have been multiple introgression events of the *yellow skin* gene into the chicken gene pool.

Phylogenetic analyses of sequences of mtDNA and random nuclear loci (*38*) derived from the four species of jungle fowl and yellow and white skinned chickens, revealed a tight cluster of chicken and red jungle fowls sequences separated from the other jungle fowl species—which suggested a limited contribution of genetic material from the grey jungle fowl to the genetic makeup of domestic chicken.

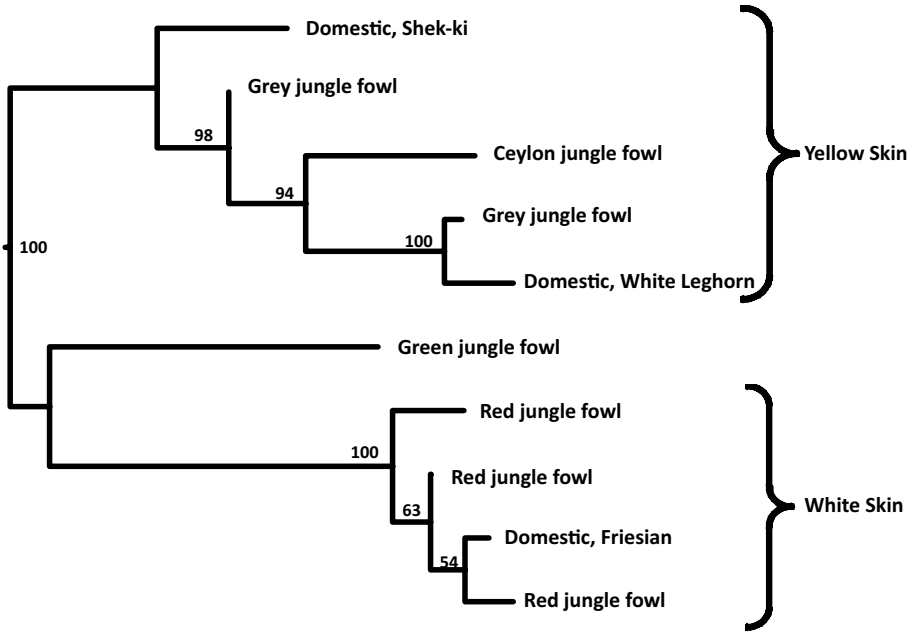

Figure 4. A neighbor-joining tree depicting the relationships between sequences derived from 23.8 kb of the *BCO2* locus. Wild and domestic samples possessing white and yellow skin clearly separate into two divergent clades. Node support values were generated from 1000 bootstrap replicates.

Since yellow skin is a very common phenotype—it has nearly replaced white skin in commercial populations—we wanted to examine if the *yellow skin* allele had any pleiotropic effects on other traits. Quantitative trait locus (QTL) analyses on a pedigree material segregating at the *yellow skin* locus revealed no significant association between the approximately 80 recorded traits—including egg production and growth—and *yellow skin*. We speculate that the strong positive selection for yellow skin may be for cosmetic reasons only, or that yellow skin had been used as an indicator of the egg-laying capacity of a hen: egg-laying bleaches the skin color, as all acquired carotenoids are allocated to the egg-yolks (*35*), and this may have been utilized by the early farmers to assess the length of a hens egg-laying period, a trait correlated with total egg-production.

Paper III – A frameshift mutation in *COMTD1* specifically dilutes pheomelanin pigmentation in chicken

In paper III (manuscript), we present genetic data implying that the Inhibitor of Gold phenotype is caused by a frameshift mutation in the previously uncharacterized *catechol-O-methyltransferase domain containing 1* (*COMTD1*) gene. This adds a new gene to the list of few genes that affect the red pheomelanin pigmentation but which leave the brown/black eumelanin pigmentation unaffected.

The *Inhibitor of Gold* locus was mapped—by linkage analysis in a pedigree segregating for IG—to a 3.58 megabase (Mb) region located between the nucleotide positions 12,388,399–15,970,174 on chicken chromosome 6. No recombination was observed between genetic markers within the 3.58 Mb region and *IG*, which was a much larger non-recombinant region than expected considering the size of the pedigree material (166 informative meiosis) and the high average recombination rate of the chicken genome (*52*). However, a low recombination rate in the chromosomal region was also observed in the consensus map for chicken (*102*), which made us conclude that *IG*IG* allele is not associated with a large inversion—inversions are known to suppress recombination.

Using IBD mapping—under the assumption that most *IG*IG* chromosomes trace back to one common ancestor—we identified a shared haplotype among IG birds from 5 different populations that spanned 438,921 bp. The region harbored three ENSEMBL gene predictions corresponding to the protein-coding genes *COMTD1*, *C10ORF11* and *ZNF503*. A forth ENSEMBL gene prediction (*ENSGALG00000024010*) was also located in the region but was only supported by a weak sequence similarity to a protein encoded by the baboon herpes virus (Uniprot ID; Q9Q5K9). An even smaller haplotype associated with *Inhibitor of Gold*—containing only *COMTD1*—was also determined. This IBD region was just supported by one recombinant *IG*IG* chromosome and further analyses were performed on the high confidence 439 kb minimum shared region.

To search for potential causal mutations for the IG phenotype, the whole genome sequence of a bird homozygous *IG*IG* was generated, using next-generation Illumina sequencing technology. Polymorphic sites, including single nucleotide polymorphisms, small insertions and deletions and large structural variants, were identified and intersected with coding and evolutionary conserved sequences within the high confidence IBD region—in which the causative mutation was supposed to be located. The only mutation that fell in the coding or highly conserved sequences was a 2-bp-insertion in the 5th exon of the *COMTD1* gene. The insertion was predicted to alter the reading frame of *COMTD1* resulting in a premature stop codon.

Reverse Transcriptase PCR analysis of RNA samples derived from feather follicles representing the three genotypes at the *IG* locus, revealed the presence of two different *COMTD1* transcripts associated with the *IG*IG* allele: one full-length transcript with the 2-bp insertion, and a second with the 2-bp insertion that was lacking the 6th exon of *COMTD1*. The proteins (Chicken_IG1 and Chicken_IG2) translated from the two *IG*IG* associated transcripts, were both predicted to be completely or partially non-functional, as both were lacking C-terminal sequences that were highly conserved across vertebrate orthologs (Figure 5).

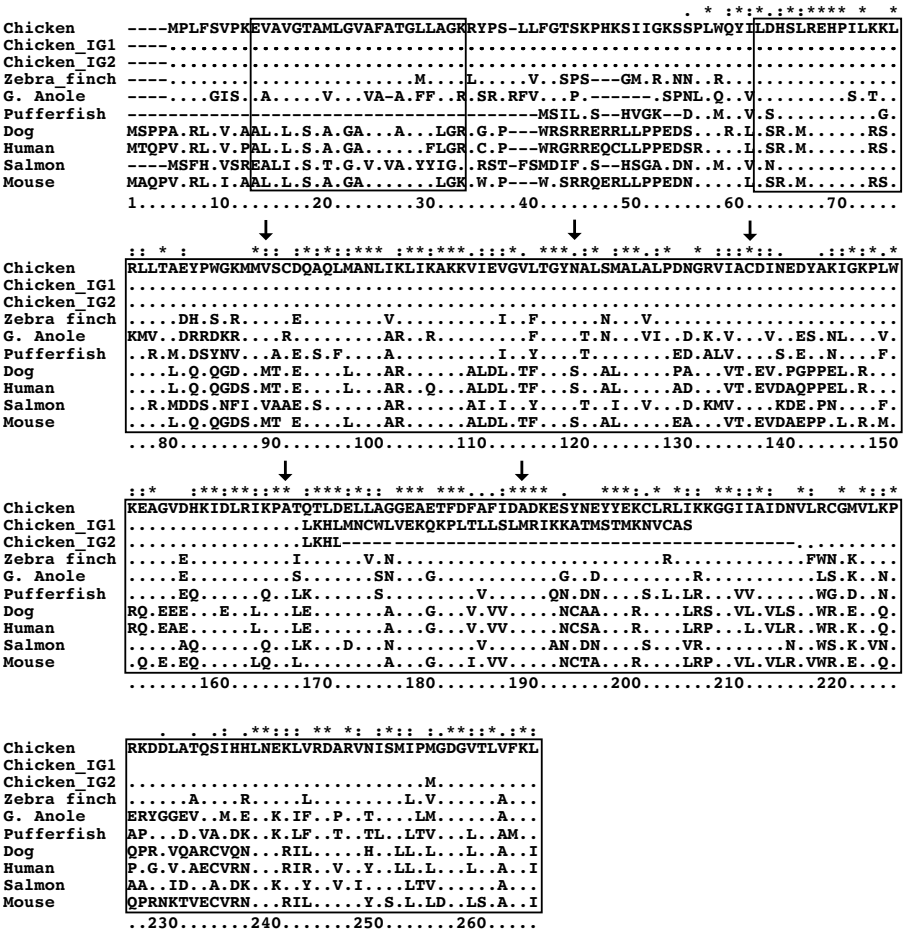

Figure 5. A protein sequence alignment of COMTD1 homologs in vertebrates, including the wild type chicken COMTD1 and the two predicted protein sequences encoded by the *IG*IG* allele. Periods (.) in the alignment indicate the same amino acid as the full-length chicken master sequence and dashes (–) indicate gaps in the alignment. An asterisk (*) indicates positions which have a single, fully conserved residue. A colon (:) indicates conservation between amino acid groups of strongly similar properties and a dot (.) indicates conservation between groups of weakly similar properties according to the Gonnet Pam matrix made by Clustal X (*103*). The arrows indicate residues that are in direct contact with the cofactor, S-Adenosyl-Methionine, based on the crystal structure model of COMTD1 (*104*). The black square surrounding residues 13–33 indicates the putative membrane-binding domain while the black square spanning residues 62–265 indicates the putative O-methyltransferase domain.

Homozygosity for the 2-bp insertion in *COMTD1* was shown to be in complete association with the Inhibitor of Gold phenotype in a pedigree material segregating for IG. The 2-bp insertion was also demonstrated to be in near complete linkage disequilibrium with *IG*IG* across breeds segregating for the phenotype. These results combined with the predicted severity of the mutation—a complete or partial loss-of-function—led us to propose that the 2-bp insertion in *COMTD1* is the causative mutation for the Inhibitor of Gold phenotype.

We present three hypotheses that could explain how a non-functional *COMTD1* would result in specific dilution of pheomelanin pigmentation. These theories are based on the known function of the well-characterized catechol-O-methyltransferase (COMT) protein, which significantly resembles COMTD1 in terms of sequence and structure, including the predicted O-methyltransferase domain.

We first propose that the COMTD1 protein may have the ability to alter cysteine homeostasis—cysteine is required for the synthesis of pheomelanin—due to its predicted function as an O-methyltransferase. O-methyltransferases are known to catalyze the transfer of a methyl group from S-Adenosyl-Methionine (SAM) to a variety of substrates (*105*). This transfer results in the conversion of S-Adenosyl-Methionine to S-Adenosyl-Homocysteine (SAH), which is an intermediate molecule in the cysteine biosynthesis pathway (*106*). A non-functional COMTD1 protein would thus potentially decrease the levels of cysteine—due to reduced SAM to SAH conversion—and that could subsequently affect the production of pheomelanin.

The second hypothesis is that the COMTD1 protein regulates one of the many reactions in the pheomelanin biosynthesis pathway. The COMT enzyme has been suggested to act as a regulator of melanogenesis due to its ability to O-methylate L-dopa, which is a precursor molecule to melanin (*107, 108*). We therefore speculate that the COMTD1 protein may act as a pheomelanin synthesis specific pathway regulator.

The last theory we put forward is that COMTD1 exerts a protective role against the cytotoxic metabolites produced during pheomelanongenesis. The actual production of melanin—both eu- and pheomelanin—has been shown to be cytotoxic due to the production of toxic intermediates that occasionally leak from the melanosome into the cytosol (*57*). COMT has been proposed to be involved in the detoxification process of these cytotoxic melanogeneic intermediates (*107*) and COMTD1 may act in a similar manner during pheomelanogenesis, but lack specificity to the intermediates produced as a result of eumelanogenesis.

Signatures of Selection in the Chicken Genome

Background

A selective sweep—characterized by a reduced degree of nucleotide variation at linked loci in a gene pool of interest—is the product of strong positive selection. Selective sweeps arise when a favorable allele at a genetic locus, and hitchhiking neutral variants at neighboring loci, reach fixation or close to fixation in a population (*86*). This can occur as a consequence of strong natural or artificial selection and the size of the selective sweep region depends on the local recombination rate and the number of generations it took for the selected allele to reach fixation.

Selective sweeps are a common feature in populations of domestic animals due to both the strong and recent selection for mutations with phenotypic effects—exemplified by the selective sweeps at the *IGF2* locus in lean pigs (*109*) and at the *BCO2* locus in yellow skinned chickens (*64*). In 2004, it was proposed that low-coverage whole-genome-sequences of domestic animals—selected for the same purpose—could be used to identify selective sweeps on a genome-wide basis (*30*).

In paper II of this thesis, we describe the use of next-generation sequencing technology on pools of DNA samples—representing different domestic chicken populations and red jungle fowl populations—to identify haplotypes that may have undergone selective sweeps during chicken domestication and the latter specialization of egg and meat producing chicken lines. These haplotypes may contain alleles that could explain: how the domestic chicken can produce up to 300 eggs per year in comparison to its major ancestor, the red jungle fowl that lay a single clutch of 5 eggs per year (*47*); or how the commercial broiler easily can weigh 4–5 kg compared to the ~1 kg of the red jungle fowl (*34*).

The chicken was selected for this experiment due to its small genome size (~1Gb) and high neutral variation rate (~5SNPs/kb) within and between chicken populations. For this study four broiler chicken lines were selected: two commercial broiler lines collected in the AvianDiv project (*110*) and the High and Low growth line (*111*) that have been divergently selected for body weight at 8 weeks for more than 50 years. In addition, four egg-layer chicken lines were selected: two White Leghorn populations, the OS-line (*112*) and a commercial Rhode Island Red line (*110*). We also collected samples from two red jungle fowl zoo populations. For each population we selected the DNA from 8-11 birds to pool prior to the resequencing on the Applied Biosystems SOLiD platform.

Results and Discussion

Paper 2 – Whole-genome resequencing reveals loci under selection during chicken domestication.

In paper II, we identify a number of loci showing a low degree of heterozygosity in eight domestic chicken populations, and which likely are the products of positive selection during chicken domestication. These regions are assumed to be enriched for functional variants that have facilitated the transformation of the wild red jungle fowl to the domestic chicken. We also detected regions that have been under positive selection during the formation of egg and meat producing chicken lines. In addition, we present a large number of deletions in eight chicken populations, some of which may have functional importance. The putative selective sweep regions provide an important raw material that can be used to investigate the differences observed between the red jungle fowl and domestic chickens, and may also shed light onto the genetic basis of the domestication process.

The whole-genome sequences were generated from the pools of DNA samples representing eight different domestic chicken populations and a pool of red jungle fowls. The genome sequence of the red jungle fowl (UCD 001) bird—previously used to generate the chicken genome assembly (*52*)—was also resequenced to function as a quality control for the sequence data generated with the SOLiD technology (Applied Biosystems, U.S.A). The short 35-bp sequences were aligned to the reference genome sequence and single nucleotide polymorphisms (SNPs) were identified. Rigorous quality checks—including the experimental verification of 318 SNPs using an Illumina GoldenGate assay—suggested that the majority of the identified ~7,5 million SNP loci were true variants.

Next, the allele counts at all polymorphic sites were determined in three different groups (breed pools) selected to represent the same purpose: all eight domestic lines, three egg-layer lines (WL-A, WL-B and RIR) and two broiler lines (CB-1 and CB-2). The rationale behind this move was to minimize the risk of detecting genomic regions fixed due to genetic drift. Putative selective sweeps were identified by calculating the pooled heterozygosity (H_p) in 40 kb windows—using the allele counts within the three breed pools—sliding along the autosomes with the step size of 20 kb. The H_p values of each window were Z-transformed (ZH_p) based on the observation that the distributions of the H_p values in the three comparisons resembled normal distributions. A window displaying a ZH_p score of –6 or less was considered to be a potential locus of selection, since these were in the extreme lower end of the distributions.

The previously identified *yellow skin* allele at the *BCO2* locus (*64*) was used as a proof-of-principle for our approach to detect true selective sweeps—since all sequenced chicken populations were assumed to be fixed for *yellow skin*. The 40-kb window that overlapped the *yellow skin*/*BCO2*

region showed the forth-lowest ZH_p score (−8.2) among all ZH_p scores in the all-domestic comparison, and thus confirmed the capacity of our approach to detect true selective sweeps.

For the all-domestic comparison, a total of 58 windows representing 21 loci, displayed a ZH_p score lower than −6. The putative selective sweep with the lowest ZH_p score (−9.2) occurred at the locus encoding the thyroid stimulating hormone receptor (*TSHR*) on chromosome 5. This was a noteworthy discovery since TSHR has a well-established role in metabolic regulation and reproduction (*113-116*), changes in which are associated with the domestication process (*26*). To examine the geographical distribution of the selected TSHR haplotype, genotyping of eight SNP loci at the *TSHR* locus in 271 birds representing 36 different chicken populations of a worldwide distribution, was performed. This revealed an amazingly high homozygosity at the *TSHR* locus: 264 of the 271 birds were homozygous for the TSHR sweep haplotype and the remaining 7 were heterozygous (Figure 6). This extreme fixation observed in distantly related chicken populations led us to propose that *TSHR* may be a so-called domestication locus, and have played an important role in shaping domestic chicken.

To search for selected mutations, Sanger sequencing was utilized to cover gaps in the *TSHR* sweep region. The non-synonymous mutation at nucleotide position 43,250,347 on chromosome 5—resulting in the glycine (Gly) to arginine (Arg) substitution at residue 558 in the TSHR protein—was considered a top candidate for the selective sweep at the *TSHR* locus. The glycine at that position was found to be completely conserved among all known vertebrate TSHR sequences, which suggested that the substitution most likely would alter the function of TSHR.

Figure 6. Allele frequencies of the selected *TSHR* haplotype—presented as pie charts—in chicken populations grouped on their continental origin (Europe, Asia, Australia, North America and Africa). The grey color of the pie charts represents the selected domestic *TSHR* haplotype while black represents non-selected haplotypes. The allele frequencies are taken from Rubin et al (2010) and are based on data from 271 birds representing 36 different populations.

A previous experiment aimed at identifying QTL for growth used an intercross between the red jungle fowl and White Leghorn (*117*). Even though the *TSHR* locus was determined to segregate in the intercross, it did not overlap with any of the QTL regions previously reported, suggesting that the mutation is not affecting growth. Instead, we speculate that the domestic *TSHR* allele potentially acts to reduce the regulation placed on seasonal reproduction. TSHR signaling has an established role in the regulation of photoperiodic control of reproduction (*114-116*), and the absence of strict regulation of seasonal breeding is in fact a common feature in domestic animals (*118*).

Several of the putative sweep regions identified in the commercial broiler lines occurred at genes with roles in obesity (*TBC1D1),* growth (*IGF1, INSR*), and regulation of appetite (*PMCH*). This was anticipated since broiler chickens have primarily been selected for muscle growth. An additional 5 putative selected loci with ZH_p scores less than –6 were also identified in the broiler lines.

We also searched for loss-of-function mutations—nonsense mutations and deletions—in the sequence data, since these have been proposed as an important factor in rapid evolution (*119*). Very little evidence was found that the selection of loss-of-function mutations had been important during the domestication process of chicken. Out of 1300 identified deletions, only seven affected coding sequences, and two of those were considered to be functional. The first deleted sequence in the growth hormone receptor (*GHR*) gene, which was previously described to cause sex-linked dwarfism (*120*).

The second was a deletion in the uncharacterized gene *SH3RF2* (SH3 domain containing ring finger 2). This deletion was considered to be especially interesting since it sat within a QTL region for growth identified in the F2 generation of a cross between the High and Low growth lines. The High growth line was fixed for the deletion, while it occurred at a low frequency in the Low Growth line, and expression data from the hypothalamus showed *SH3RF2* expression in the Low growth line but not in the High growth line. A QTL analysis with genotype data from the *SH3RF2* locus in the F8 generation of the High and Low growth line intercross revealed a strong association between the deletion and increased growth. This suggests that the deletion in *SH3RF2* is the causal mutation for the QTL on chromosome 20.

General Discussion and Future Perspectives

In papers I and III of this thesis, we present two genetic mechanisms underlying some of the variation in carotenoid and melanin pigmentation observed in the domestic chicken. The genes identified to control the two monogenic traits are excellent candidates for deciphering the genetic basis of the variation in carotenoid and pheomelanin pigmentation that is observed in many other organisms. In paper II, we take one step closer to unraveling the genetic basis for the transformation of the wild red jungle fowl into the highly productive domestic chicken. In combination, all three papers have advanced our understanding of the genetic mechanisms underlying phenotypic variation.

The yellow skin study (Paper I) provides the first evidence of a hybrid origin of the domestic chicken—as proposed by Hutt in 1949. Despite that the amount of genetic introgression from the grey jungle fowl seems to be limited, it would be worthwhile to explore the extent of introgression from the grey jungle fowl and the additional two jungle fowl species on a genome-wide basis. This may be achieved by generating the genome sequences for the four jungle fowl species, and a set of chicken breeds, and searching for genomic regions where the chicken sequences cluster tightly with one of the jungle fowls, rather than red jungle fowl. Such regions are thought to be enriched for functional alleles since they have not been removed from the chicken population through the extensive backcrossing that most likely took place after the introgression. These alleles would probably also exert large phenotypic effects on the selected traits, if we assume that the introgression occurred early during domestication and before the worldwide dispersal of chicken, when visual selection was the only selection method present. Mutations with small phenotypic effects on traits such as growth and reproduction have probably first increased in frequency within the last 200 years or so, when more advanced methods for phenotypic selection were developed.

This study also provides insight into the molecular mechanisms underlying variation in carotenoid pigmentation between two distinct wild species—namely the red and grey jungle fowls—and thus sheds some light on the evolution of carotenoid-based pigmentation in wild species. Our study did in fact compare sequences from two different species even though the sequences were segregating in the gene pool of a single species. With the fact that grey jungle fowl has yellow/red skin and red jungle fowl white skin, we can be quite confident that the mutation controlling the variation in carotenoid-

pigmentation arose prior to domestication. However, we are not able to determine if it is yellow or white skin that represents the ancestral state—there is no reliable information of leg color of the closely related green jungle fowl. It is quite remarkable that a study aimed at to identify the gene responsible for a phenotype observed within a single species, led to the detection of the genetic mechanism underlying variation in a phenotypic trait, observed between two distinct species.

The *BCO2* gene is hence an excellent candidate for examining the genetic mechanisms underlying variation in carotenoid-based pigmentation in other species. We have recently started a project—in collaboration with evolutionary biologists—to test if *BCO2* underlies the variation in beak color (yellow or white/pink) observed in another bird species.

Our study also adds support to the view that *cis*-regulatory mutations have played a significant role in the evolution of phenotypic variation—the yellow skin phenotype is caused by a *cis*-regulatory mutation affecting the expression of *BCO2* in skin but not in liver. BCO2 has recently been shown to be a mitochondrial protein with an important role in carotenoid homeostasis and the protection of the mitochondria against carotenoid induced dysfunction (*16*). In addition *Bco2* knockout mice display enlarged spleens and reduced testis size compared to wild-type mice after being fed a carotenoid-rich diet (*121*). A complete loss-of-function mutation of *BCO2* would most likely confer deleterious effects on many traits and hence would be removed from the gene pool. Interestingly, no adverse effects on health or development are associated with the complete loss-of-function (nonsense) mutations of *BCO2* causing increased carotenoid levels in cow milk (*100, 101*) and yellow fat in cow (*101*) and sheep (*73*). A more thorough investigation of numerous phenotypic traits for the different genotypes is required to truly rule out the presence of deleterious effects associated with loss-of-function mutations in *BCO2* in these species.

Regarding the *Inhibitor of Gold* locus (Paper III), several functional experiments are required to develop our understanding of how the 2-bp insertion in *COMTD1* results in reduced pheomelanin pigmentation. Characterization of the morphology and the number of melanocytes in feather follicles derived from IG and wild type birds, would perhaps reveal if the COMTD1 protein exerts a protective role against the toxic intermediates produced during pheomelanogenesis. Whereas knock-down or over-expression experiments of *COMTD1* in melanocyte cell lines, followed by the quantification of pheomelanin using HPLC, would indicate if *COMTD1* possesses a regulatory role in the synthesis of pheomelanin. To assess if COMTD1 is affecting pheomelanin pigmentation, through regulating the levels of available cysteine for the pheomelanin synthesis, an appropriate experiment would be to measure the cysteine levels in blood plasma from IG and wild type birds (*122*).

It is interesting to speculate why *COMTD1* has never previously been associated with pheomelanin pigmentation in other species—even through extensive mouse mutagenesis screens. One explanation could be that in other species, a mutant *COMTD1* would only have a subtle effect on pheomelanin pigmentation and is thus not detected in the mouse mutagenesis programs. Another explanation is that COMTD1 mutations are perhaps lethal due to negative pleiotropic effects on essential physiological processes. The Inhibitor of Gold bird represents a unique resource in which to study the function of *COMTD1*.

In paper II, we identified a number of loci in the chicken genome that likely have been under selection during the domestication process of chicken. These regions are thus most likely enriched for alleles that can explain some of the major phenotypic differences observed between the major wild ancestor, red jungle fowl, and the domestic chicken. We also detected a number of putative selected loci specific to meat or egg producing chicken lines. The selected loci, found in broiler and layer chickens, may be utilized to advance our knowledge of how energy metabolism, appetite and reproduction are regulated in vertebrate species.

First, further investigations in terms of genetic and functional experiments are required to establish if sweep signals are the result of selection or genetic drift. This may be followed up with experiments to determine the link between the selected haplotypes and phenotypes. One obvious effort that would add support to the premise that selection has created the sweep signals, is if the selected mutations responsible for the sweep signatures were found. For instance, non-synonymous single nucleotide polymorphisms (nsSNP(s)) are fairly easy to detect and may be examined using SIFT (*90*) or PolyPhen software (*123*), to predict the functional effect of the polymorphism. The nsSNPs—predicted to have a functional effect—could be further investigated using intercrosses between different chicken lines (e.g. egg layer/broiler or White Leghorn/red jungle fowl) if they were found to segregate within these populations. QTL analyses with the SNP loci genotypes could reveal if they are associated with any of the phenotypes collected from the intercross. All putative sweep regions—with or without the presence of candidate mutations—should be genotyped in intercrosses that are believed to segregate at the sweep loci, in order to search for associations between the sweep haplotypes and the recorded traits. In fact, such analyses are now being performed with broiler sweep loci in an intercross between egg layer and broiler chickens.

The genes within, or in close proximity to, the putative sweep regions should also be examined to see if they display differential expression between the appropriate populations—i.e. red jungle fowl vs. domestic and layer vs. broiler—that may unearth if the selected mutations have a *cis*-acting effect on gene expression. The combination of results from the above-described analyses will perhaps provide the evidence required to decide

which of the putative selective sweeps that are the result of selection or genetic drift, and also possibly link the true selected loci to their phenotypes.

The supposed domestication locus at the *TSHR* gene should be followed up with functional experiments and segregation analysis with data of phenotypic traits that are proposed to be affected by altered TSHR signaling. It is tempting to speculate that the regulation of thyroid hormone signaling is altered in other domesticated species—resulting in many of phenotypes associated with the domestication process. This may be settled when the data concerning the genetic basis of domestication in other animal species are examined and made public.

Our study provides evidence that it is possible to detect loci that have been under strong positive selection during chicken domestication using next-generation sequencing technology on pools of DNA. The same should now be possible for mammalian domesticated species with larger genomes sizes than chicken, since the cost per sequenced base has dropped considerably in the last two years. It would also be interesting to utilize our approach to detect signatures of selection in wild populations, in order to identify haplotypes controlling phenotypic traits that are important for adaptation to different environments.

Acknowledgements

This work was performed at the Department of Medical Biochemistry and Microbiology, Uppsala, Sweden.

I would like to express my sincere gratitude to all the people that I have had the pleasure to work with during my PhD studies. In particular I would like to acknowledge the following people:

First, my supervisor Leif Andersson, for letting me be a part of so many stimulating projects, and for guiding me during the current exciting period of genome biology. It has been very inspiring to work with an excellent scientist such as yourself; you are a prime example of the correlation between passion and success. A big thank you for the all the support you gave me as I was struggling with finishing up a manuscript and my thesis at the same time.

My co-supervisor Göran Hjälm, for sharing your knowledge of protein biology and for keeping the local gym in excellent condition. Thanks for ordering "handsprit" for me so that I could stay healthy during this busy period.

Greger Larson, my second co-supervisor, for sharing your extensive knowledge of domestication and phylogenetic analyses—which resulted in the yellow skin paper—and for giving me access to your fine iTunes library.

Jennifer Meadows, colleague, for always having smart answers to my questions, and for your extensive input on my thesis, despite being down with a cold and fighting jet lag.

Carl-Johan Rubin, colleague, for your inspiring "all-in attitude" and precious help with the analyses of whole-genome sequence data.

All other excellent scientists—none mentioned, none forgotten—who have contributed to the findings presented in this thesis.

Anders Hellström, former colleague. We really had a good time in the lab with our deep scientific discussions mixed with rubber band wars and "Kejsarens kycklingsallad" lunches. I'm happy that you still are my friend, now when you left the lab, and I wish you all the luck in south of Stockholm.

Lisa, Maria and Anders, my old office mates, for the hot dog sessions at IKEA and the interesting and amusing discussions in the office.

All current and former members of the lab, for all your excellent support and for making the years in the lab to an exciting time in my life. I will miss our discussions on topics ranging from chicken breeding, HiFi equipment,

mushroom picking, rugby, Silvio Berlusconi, Chinese pop music, breast-feeding, beekeeping, to kitchen blenders.

Jag vill tacka min mentor Christer Harbeck som fått mig att inse att mycket spännande väntar efter disputation. Samt för att ha påtalat vikten av planering och förberedelse inför stora prövningar, som att skriva avhandling och att sedan försvara den.

Jag vill också framföra ett stort tack till min omfattande familj i Upplands Väsby, Leksand, Solna, Norrtälje, Täby, Julita, Tullinge, o.s.v. Jag vill speciellt tacka pappa, mamma, och mina bröder för all glädje och allt stöd ni gett mig under årens gång. Ett gigantiskt tack till min mamma för alla de timmar du tragglade ”plutifikationstabellen” och högläsning med decemberpojken Jonas. Tragglandet visade ju sig ge resultat!

Ett stort tack till alla mina vänner som utstått min extra påtagliga frånvaro det senaste halvåret. Nu kommer det säkert finnas mer tid för ost- och vinkvällar och Halo sessioner.

Sist men absolut inte minst, vill jag tacka min underbara Anna, för att du ordnat med allt hemma, köpt mina julklappar, korrekturläst min avhandling och till och med hunnit ordna en nobelmiddag åt mig. Utan din matlagningskonst hade jag nog tappat ytterligare 5 kg i vikt, och utan dina lugnande ord hade jag varit mycket mer stressad. Det ska bli underbart att nu få mer tid med dig och jag ser framemot allt spännande vi tillsammans har framför oss i livet.

References

1. C. Darwin, On the origin of species by means of natural selection: or the preservation of favoured races in the struggle for life. (John Murray, London, ed. 1., 1859), pp. 502, 32 s.
2. C. Darwin, The variation of animals and plants under domestication. (O. Judd & company, New York,, ed. Authorized, 1868), pp. 2 v.
3. J. D. Watson, F. H. Crick, Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 171, 737 (Apr 25, 1953).
4. E. S. Lander et al., Initial sequencing and analysis of the human genome. Nature 409, 860 (Feb 15, 2001).
5. J. C. Venter et al., The sequence of the human genome. Science 291, 1304 (Feb 16, 2001).
6. E. S. Lander, Initial impact of the sequencing of the human genome. Nature 470, 187 (Feb 10, 2011).
7. D. L. Hartl, A primer of population genetics. (Sinauer Associates, Sunderland, Mass, ed. 3rd, 2000), pp. xvii, 221 p.
8. W. S. Klug, M. R. Cummings, Essentials of genetics. (Prentice Hall, Upper Saddle River (N.J.), ed. 4th, 2002).
9. G. A. Wray et al., The evolution of transcriptional regulation in eukaryotes. Mol Biol Evol 20, 1377 (Sep, 2003).
10. D. L. Stern, V. Orgogozo, The loci of evolution: how predictable is genetic evolution? Evolution 62, 2155 (Sep, 2008).
11. E. H. Davidson, The regulatory genome : gene regulatory networks in development and evolution. (Academic, Burlington, MA ; London, 2006), pp. xi, 289 p.
12. J. A. Coyne, H. E. Hoekstra, Evolution of protein expression: new genes for a new diet. Current biology : CB 17, R1014 (Dec 4, 2007).
13. G. A. Wray, The evolutionary significance of cis-regulatory mutations. Nature reviews. Genetics 8, 206 (Mar, 2007).
14. F. Jacob, J. Monod, Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol 3, 318 (Jun, 1961).
15. M. C. King, A. C. Wilson, Evolution at two levels in humans and chimpanzees. Science 188, 107 (Apr 11, 1975).
16. J. Amengual et al., A mitochondrial enzyme degrades carotenoids and protects against oxidative stress. The FASEB journal 25, 948 (Mar, 2011).
17. R. P. Alexander, G. Fang, J. Rozowsky, M. Snyder, M. B. Gerstein, Annotating non-coding regions of the genome. Nature reviews. Genetics 11, 559 (Aug, 2010).
18. R. H. Waterston et al., Initial sequencing and comparative analysis of the mouse genome. Nature 420, 520 (Dec 5, 2002).
19. K. Lindblad-Toh et al., A high-resolution map of human evolutionary constraint using 29 mammals. Nature 478, 476 (Oct 27, 2011).

20. R. A. Dalloul et al., Multi-platform next-generation sequencing of the domestic turkey (Meleagris gallopavo): genome assembly and analysis. PLoS Biol 8, (2010).
21. E. Birney et al., Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447, 799 (Jun 14, 2007).
22. L. Andersson, in Comparative Genomics : Basic and Applied Research. (Taylor & Francis Group, Boca Raton, 2007), pp. 22.
23. C. A. Driscoll, D. W. Macdonald, S. J. O'Brien, From wild animals to domestic pets, an evolutionary view of domestication. Proc Natl Acad Sci U S A 106 Suppl 1, 9971 (Jun 16, 2009).
24. M. Fang, G. Larson, H. S. Ribeiro, N. Li, L. Andersson, Contrasting mode of evolution at a coat color locus in wild and domestic pigs. PLoS Genet 5, e1000341 (Jan, 2009).
25. A. Ludwig et al., Coat color variation at the beginning of horse domestication. Science 324, 485 (Apr 24, 2009).
26. P. Jensen, L. Andersson, Genomics meets ethology: a new route to understanding domestication, behavior, and sustainability in animal breeding. Ambio 34, 320 (Jun, 2005).
27. L. Trut, Early canid domestication: The farm-fox experiment. American Scientist 87, (1999).
28. W. R. Stricklin, in Social behaviour in farm animals. (CABI Publishing, Wallingford, 2001), pp. xvii, 406 p.
29. S. J. Crockford, in Human evolution through developmental change, N. Minugh-Purvis, K. McNamara, Eds. (Johns Hopkins University Press, Baltimore, 2002), pp. 122 – 153.
30. L. Andersson, M. Georges, Domestic-animal genomics: deciphering the genetics of complex traits. Nature reviews 5, 202 (Mar, 2004).
31. N. H. Salmon Hillbertz et al., Duplication of FGF3, FGF4, FGF19 and ORAOV1 causes hair ridge and predisposition to dermoid sinus in Ridgeback dogs. Nat Genet 39, 1318 (Nov, 2007).
32. M. Olsson et al., A novel unstable duplication upstream of HAS2 predisposes to a breed-defining skin phenotype and a periodic fever syndrome in Chinese Shar-Pei dogs. PLoS Genet 7, e1001332 (Mar, 2011).
33. J. Diamond, Evolution, consequences and future of plant and animal domestication. Nature 418, 700 (Aug 8, 2002).
34. R. D. Crawford, Poultry breeding and genetics. Developments in animal and veterinary sciences, (Elsevier, Amsterdam ; New York, 1990), pp. 1123 s.
35. F. B. Hutt, Genetics of the fowl. (McGraw-Hill book co, New york, ed. 1st, 1949), pp. 590 s.
36. A. Fumihito et al., One subspecies of the red junglefowl (Gallus gallus gallus) suffices as the matriarchic ancestor of all domestic breeds. Proc Natl Acad Sci U S A 91, 12505 (Dec 20, 1994).
37. A. Fumihito et al., Monophyletic origin and unique dispersal patterns of domestic fowls. Proc Natl Acad Sci U S A 93, 6792 (Jun 25, 1996).
38. M. Nishibori, T. Shimogiri, T. Hayashi, H. Yasue, Molecular evidence for hybridization of species in the genus Gallus except for Gallus varius. Anim Genet 36, 367 (Oct, 2005).
39. B. West, B.-X. Zhou, Did chickens go North? New evidence for domestication. Journal of Archaeological Science 15, 515 (1988).
40. Y. P. Liu et al., Multiple maternal origins of chickens: out of the Asian jungles. Mol Phylogenet Evol 38, 12 (Jan, 2006).

41. S. Kanginakudru, M. Metta, R. D. Jakati, J. Nagaraju, Genetic evidence from Indian red jungle fowl corroborates multiple domestication of modern day chicken. BMC Evol Biol 8, 174 (2008).
42. C. Berthouly-Salazar et al., Vietnamese chickens: a gate towards Asian genetic diversity. BMC Genet 11, 53 (2010).
43. J. Gongora et al., Indo-European and Asian origins for Chilean and Pacific chickens revealed by mtDNA. Proc Natl Acad Sci U S A 105, 10308 (Jul 29, 2008).
44. K. N. Dancause, M. G. Vilar, R. Steffy, J. K. Lum, Characterizing genetic diversity of contemporary pacific chickens using mitochondrial DNA analyses. PLoS ONE 6, e16843 (2011).
45. F. C. Muchadeyi et al., Mitochondrial DNA D-loop sequences suggest a Southeast Asian and Indian origin of Zimbabwean village chickens. Anim Genet 39, 615 (Dec, 2008).
46. P. B. Siegel, J. B. Dodgson, L. Andersson, Progress from chicken genetics to the chicken genome. Poult Sci 85, 2050 (Dec, 2006).
47. P. A. Johnsgard, The pheasants of the world : biology and natural history. (Smithsonian Institution Press, Washington, DC, ed. 2nd ed, 1999), pp. xvii, 398 p.
48. W. Bateson, Experiments with poultry. Repts. Evol. Comm. Roy. Soc. 1, 87 (1902).
49. S. Cohen, R. Levi-Montalcini, A Nerve Growth-Stimulating Factor Isolated from Snake Venom. Proc Natl Acad Sci U S A 42, 571 (Sep, 1956).
50. P. Rous, A Sarcoma of the Fowl Transmissible by an Agent Separable from the Tumor Cells. J Exp Med 13, 397 (Apr 1, 1911).
51. M. D. Cooper, D. A. Raymond, R. D. Peterson, M. A. South, R. A. Good, The functions of the thymus system and the bursa system in the chicken. J Exp Med 123, 75 (Jan 1, 1966).
52. International Chicken Genome Sequencing Consortium, Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432, 695 (Dec 9, 2004).
53. G. K. Wong et al., A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. Nature 432, 717 (Dec 9, 2004).
54. C. J. Rubin et al., Whole-genome resequencing reveals loci under selection during chicken domestication. Nature 464, 587 (Mar 25, 2010).
55. M. Cieslak, M. Reissmann, M. Hofreiter, A. Ludwig, Colours of domestication. Biological reviews of the Cambridge Philosophical Society, (Mar 28, 2011).
56. J. K. Hubbard, J. A. Uy, M. E. Hauber, H. E. Hoekstra, R. J. Safran, Vertebrate pigmentation: from underlying genes to adaptive function. Trends Genet 26, 231 (May, 2010).
57. J. J. Nordlund, The pigmentary system : physiology and pathophysiology. (Blackwell, Malden, Mass. ; Oxford, ed. 2nd, 2006), pp. xviii, 1229 p., [84] p. of plates.
58. L. Montoliu, Oetting, W. S. & Bennett, D. C. (European Society for Pigment Cell Research, 2011).
59. S. Kerje, J. Lind, K. Schutz, P. Jensen, L. Andersson, Melanocortin 1-receptor (MC1R) mutations are associated with plumage colour in chicken. Anim Genet 34, 241 (Aug, 2003).
60. U. Gunnarsson et al., Mutations in SLC45A2 cause plumage color variation in chicken and Japanese quail. Genetics 175, 867 (Feb, 2007).

61. U. Gunnarsson et al., The Dark brown plumage color in chickens is caused by an 8.3-kb deletion upstream of SOX10. Pigment cell & melanoma research 24, 268 (Apr, 2011).
62. M. Vaez et al., A single point-mutation within the melanophilin gene causes the lavender plumage colour dilution phenotype in the chicken. BMC Genet 9, 7 (2008).
63. A. R. Hellstrom et al., Sex-linked barring in chickens is controlled by the CDKN2A /B tumour suppressor locus. Pigment Cell & Melanoma Research 23, 521 (Aug, 2010).
64. J. Eriksson et al., Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. PLoS Genet 4, e1000010 (Feb, 2008).
65. S. Sato, T. Otake, C. Suzuki, J. Saburi, E. Kobayashi, Mapping of the recessive white locus and analysis of the tyrosinase gene in chickens. Poult Sci 86, 2126 (Oct, 2007).
66. S. Kerje et al., The Dominant white, Dun and Smoky color variants in chicken are associated with insertion/deletion polymorphisms in the PMEL17 gene. Genetics 168, 1507 (Nov, 2004).
67. T. Tobita-Teramoto et al., Autosomal albino chicken mutation (ca/ca) deletes hexanucleotide (-deltaGACTGG817) at a copper-binding site of the tyrosinase gene. Poult Sci 79, 46 (Jan, 2000).
68. B. M. Dorshorst, A-M. Johansson, AM. Strömstedt, L. Pham, M-H. Chen, CF. Hallböök, F. Ashwell, C. Andersson, L., A complex genomic rearrangement involving the Endothelin 3 locus causes dermal hyperpigmentation in the chicken. PLoS genetics in press, (2011, 2011).
69. G. A. Armstrong, J. E. Hearst, Carotenoids 2: Genetics and molecular biology of carotenoid pigment biosynthesis. Faseb J 10, 228 (Feb, 1996).
70. J. M. Berg, J. L. Tymoczko, L. Stryer, Biochemistry. (W. H. Freeman, New York, ed. 6th, 2007).
71. B. Demmig-Adams, W. W. Adams, 3rd, Antioxidants in photosynthesis and human nutrition. Science 298, 2149 (Dec 13, 2002).
72. N. I. Krinsky, J. T. Landrum, R. A. Bone, Biologic mechanisms of the protective role of lutein and zeaxanthin in the eye. Annu Rev Nutr 23, 171 (2003).
73. D. I. Vage, I. A. Boman, A nonsense mutation in the beta-carotene oxygenase 2 (BCO2) gene is tightly associated with accumulation of carotenoids in adipose tissue in sheep (Ovis aries). BMC Genet 11, 10 (2010).
74. J. D. Blount, N. B. Metcalfe, T. R. Birkhead, P. F. Surai, Carotenoid modulation of immune function and sexual attractiveness in zebra finches. Science 300, 125 (Apr 4, 2003).
75. B. Faivre, A. Gregoire, M. Preault, F. Cezilly, G. Sorci, Immune activation rapidly mirrored in a secondary sexual trait. Science 300, 103 (Apr 4, 2003).
76. D. Scherer, R. Kumar, Genetics of pigmentation in skin cancer--a review. Mutat Res 705, 141 (Oct, 2010).
77. C. Wasmeier, A. N. Hume, G. Bolasco, M. C. Seabra, Melanosomes at a glance. J Cell Sci 121, 3995 (Dec 15, 2008).
78. W. P. Walker, T. M. Gunn, Shades of meaning: the pigment-type switching system as a tool for discovery. Pigment Cell & Melanoma Research 23, 485 (Aug, 2010).
79. J. D. Simon, D. Peles, K. Wakamatsu, S. Ito, Current challenges in understanding melanogenesis: bridging chemistry, biological control, morphology, and function. Pigment Cell & Melanoma Research 22, 563 (Oct 1, 2009).

80. S. Ito, K. Wakamatsu, Chemistry of mixed melanogenesis--pivotal roles of dopaquinone. Photochem Photobiol 84, 582 (May-Jun, 2008).
81. J. A. Blake, C. J. Bult, J. A. Kadin, J. E. Richardson, J. T. Eppig, The Mouse Genome Database (MGD): premier model organism resource for mammalian genomics and genetics. Nucleic Acids Res 39, D842 (Jan, 2011).
82. L. Stevens, Genetics and evolution of the domestic fowl. (Cambridge University Press, Cambridge, 2005), pp. xiii, 306 p.
83. A. L. Hughes, M. K. Hughes, Small genomes for better flyers. Nature 377, 391 (Oct 5, 1995).
84. L. Andersson, How selective sweeps in domestic animals provide new insight into biological mechanisms. J Intern Med, (Sep 3, 2011).
85. T. Strachan, A. P. Read, D. J. Matthes, Human molecular genetics 2. (BIOS Scientific, Oxford, ed. 2nd, 1999), pp. xxiii, 576 p.
86. J. M. Smith, J. Haigh, The hitch-hiking effect of a favourable gene. Genet Res 23, 23 (Feb, 1974).
87. L. Andersson, Genetic dissection of phenotypic diversity in farm animals. Nature reviews 2, 130 (Feb, 2001).
88. M. L. Metzker, Sequencing technologies - the next generation. Nature reviews. Genetics 11, 31 (Jan, 2010).
89. W. McLaren et al., Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics 26, 2069 (Aug 15, 2010).
90. P. C. Ng, S. Henikoff, Predicting deleterious amino acid substitutions. Genome Res 11, 863 (May, 2001).
91. L. W. Taylor, in Sixth International Congress of Genetics (Ithaca, New York, 1932), pp. 197-199.
92. R. C. Punnett, Genetic studies in poultry; cream plumage. J Genet 48, 327 (Jan, 1948).
93. S. Chintala et al., Slc7a11 gene controls production of pheomelanin pigment and proliferation of cultured cells. Proc Natl Acad Sci U S A 102, 10964 (Aug 2, 2005).
94. M. W. Lieberman et al., Growth retardation and cysteine deficiency in gamma-glutamyl transpeptidase-deficient mice. Proc Natl Acad Sci U S A 93, 7923 (Jul 23, 1996).
95. N. Chalhoub et al., Grey-lethal mutation induces severe malignant autosomal recessive osteopetrosis in mouse and human. Nat Med 9, 399 (Apr, 2003).
96. M. R. Chedekel, P. P. Agin, R. M. Sayre, Photochemistry of pheomelanin: action spectrum for superoxide production. Photochem Photobiol 31, 553 (1980).
97. J. D. Simon, D. N. Peles, The red and the black. Acc Chem Res 43, 1452 (Nov 16, 2010).
98. M. Schmid et al., First report on chicken genes and chromosomes 2000. Cytogenet Cell Genet 90, 169 (2000).
99. C. Kiefer et al., Identification and characterization of a mammalian enzyme catalyzing the asymmetric oxidative cleavage of provitamin A. J Biol Chem 276, 14110 (Apr 27, 2001).
100. S. D. Berry et al., Mutation in bovine beta-carotene oxygenase 2 affects milk color. Genetics 182, 923 (Jul, 2009).
101. R. Tian, W. S. Pitchford, C. A. Morris, N. G. Cullen, C. D. Bottema, Genetic variation in the beta, beta-carotene-9', 10'-dioxygenase gene and association with fat colour in bovine adipose tissue and milk. Anim Genet 41, 253 (Jun 1, 2010).

102. M. A. Groenen et al., A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. Genome Res 19, 510 (Mar, 2009).
103. J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin, D. G. Higgins, The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25, 4876 (Dec 15, 1997).
104. J. W. Min, H. Zeng, H. Loppnau, P. Sundstrom, M. Arrowsmith, CH. Edwards, AM. Bochkarev, A. Plotnikov, AN., The crystal structure of human catechol-o-methyltransferase domain containing 1 in complex with s-adenosyl-l-methionine., Protein Data Bank (2006).
105. T. T. Myöhänen, P. T. Männistö, Distribution and Functions of Catechol-O-Methyltransferase Proteins: Do Recent Findings Change The Picture? International Review of Neurobiology - Volume 95 95, 29 (Jan 1, 2010).
106. D. L. Nelson, A. L. Lehninger, M. M. Cox, Lehninger principles of biochemistry. (Worth, New York, ed. 3rd, 2000), pp. xxix, 1231 p. in various pagings.
107. N. Smit et al., O-methylation of L-dopa in melanin metabolism and the presence of catechol-O-methyltransferase in melanocytes. Pigment Cell Res 7, 403 (Dec, 1994).
108. A. Slominski, D. J. Tobin, S. Shibahara, J. Wortsman, Melanin pigmentation in mammalian skin and its hormonal regulation. Physiol Rev 84, 1155 (Oct, 2004).
109. A. S. Van Laere et al., A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. Nature 425, 832 (Oct 23, 2003).
110. J. Hillel et al., Biodiversity of 52 chicken populations assessed by microsatellite typing of DNA pools. Genet Sel Evol 35, 533 (Sep-Oct, 2003).
111. E. A. Dunnington, P. B. Siegel, Long-term divergent selection for eight-week body weight in white Plymouth rock chickens. Poult Sci 75, 1168 (Oct, 1996).
112. R. K. Cole, Hereditary hypothyroidism in the domestic fowl. Genetics 53, 1021 (Jun, 1966).
113. K. Dobney, G. Larson, Genetics and animal domestication : new windows on an elusive process. Journal of zoology. 269, 261 (2006).
114. T. Yoshimura et al., Light-induced hormone conversion of T4 to T3 regulates photoperiodic response of gonads in birds. Nature 426, 178 (Nov 13, 2003).
115. N. Nakao et al., Thyrotrophin in the pars tuberalis triggers photoperiodic response. Nature 452, 317 (Mar 20, 2008).
116. E. A. Hanon et al., Ancestral TSH mechanism signals summer in a photoperiodic mammal. Current biology : CB 18, 1147 (Aug 5, 2008).
117. S. Kerje et al., The twofold difference in adult size between the red junglefowl and White Leghorn chickens is largely explained by a limited number of QTLs. Anim Genet 34, 264 (Aug, 2003).
118. E. O. Price. Animal Domestication and Behavior (CABI Pub., Wallingford, Oxon, UK ; New York, 2002), pp. 297 p.
119. M. V. Olson, When Less Is More: Gene Loss as an Engine of Evolutionary Change. The American Journal of Human Genetics 64, 18 (1999).
120. S. K. Agarwal, L. A. Cogburn, J. Burnside, Dysfunctional growth hormone receptor in a strain of sex-linked dwarf chicken: evidence for a mutation in the intracellular domain. J Endocrinol 142, 427 (Sep, 1994).
121. N. A. Ford, S. K. Clinton, J. von Lintig, A. Wyss, J. W. Erdman, Jr., Loss of carotene-9',10'-monooxygenase expression increases serum and tissue

lycopene concentrations in lycopene-fed mice. The Journal of nutrition 140, 2134 (Dec, 2010).
122. C. M. Pfeiffer, D. L. Huff, E. W. Gunter, Rapid and accurate HPLC assay for plasma total homocysteine and cysteine in a clinical laboratory setting. Clin Chem 45, 290 (Feb, 1999).
123. V. Ramensky, P. Bork, S. Sunyaev, Human non-synonymous SNPs: server and survey. Nucleic Acids Res 30, 3894 (Sep 1, 2002).

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 732

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine.

Distribution: publications.uu.se
urn:nbn:se:uu:diva-162597

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2012