



UPPSALA
UNIVERSITET

*Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Medicine 1124*

Monogenic Traits Associated with Structural Variants in Chicken and Horse

*Allelic and Phenotypic Diversity of Visually Appealing
Traits*

FREYJA IMSLAND



ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2015

ISSN 1651-6206
ISBN 978-91-554-9295-3
urn:nbn:se:uu:diva-259621

Dissertation presented at Uppsala University to be publicly examined in room B42, at the BMC, Husargatan 3, Uppsala, Friday, 25 September 2015 at 13:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English. Faculty examiner: Professor Eiríkur Steingrímsson (University of Iceland, Faculty of Medicine, School of Health Sciences).

Abstract

Imsland, F. 2015. Monogenic Traits Associated with Structural Variants in Chicken and Horse. *Allelic and Phenotypic Diversity of Visually Appealing Traits. Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine* 1124. 59 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-554-9295-3.

Domestic animals have rich phenotypic diversity that can be explored to advance our understanding of the relationship between molecular genetics and phenotypic variation. Since the advent of second generation sequencing, it has become easier to identify structural variants and associate them with phenotypic outcomes. This thesis details studies on three such variants associated with monogenic traits.

The first studies on *Rose-comb* in the chicken were published over a century ago, seminally describing Mendelian inheritance and epistatic interaction in animals. Homozygosity for the otherwise dominant *Rose-comb* allele was later associated with reduced rooster fertility. We show that a 7.38 Mb inversion is causal for *Rose-comb*, and that two alleles exist for *Rose-comb*, *R1* and *R2*. A novel genomic context for the gene *MNR2* is causative for the comb phenotype, and the bisection of the gene *CCDC108* is associated with fertility issues. The recombined *R2* allele has intact *CCDC108*, and normal fertility.

The dominant phenotype Greying with Age in horses was previously associated with an intronic duplication in *STX17*. By utilising second generation sequencing we have examined the genomic region surrounding the duplication in detail, and excluded all other discovered variants as causative for *Grey*.

Dun is the ancestral coat colour of equids, where the individual is mostly pale in colour, but carries intensely pigmented primitive markings, most notably a dorsal stripe. *Dun* is a dominant trait, and yet most domestic horses are non-dun in colour and intensely pigmented. We show that Dun colour is established by radially asymmetric expression of the transcription factor *TBX3* in hair follicles. This results in a microscopic spotting phenotype on the level of the individual hair, giving the impression of pigment dilution. Non-dun colour is caused by two different alleles, *non-dun1* and *non-dun2*, both of which disrupt the *TBX3*-mediated regulation of pigmentation. *Non-dun1* is associated with a SNP variant 5 kb downstream of *TBX3*, and *non-dun2* with a 1.6 kb deletion that overlaps the *non-dun1* SNP. Homozygotes for *non-dun2* show a more intensely pigmented appearance than horses with one or two *non-dun1* alleles. We have also shown by genotyping of ancient DNA that *non-dun1* predates domestication.

Keywords: Structural variation, Pigmentation, Domestication, Equids, *MNR2*, *CCDC108*, *STX17*, *TBX3*, *Grey*, *Dun*, non-dun, *Rose-comb*, Chicken, Genetic mapping, Phenotyping

Freyja Imsland, Department of Medical Biochemistry and Microbiology, Box 582, Uppsala University, SE-75123 Uppsala, Sweden. Science for Life Laboratory, SciLifeLab, Box 256, Uppsala University, SE-75105 Uppsala, Sweden.

© Freyja Imsland 2015

ISSN 1651-6206

ISBN 978-91-554-9295-3

urn:nbn:se:uu:diva-259621 (<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-259621>)

*Maðurinn einn er ei nema hálfur,
með öðrum er hann meiri en hann
sjálfur.*

- Einar Benediktsson

List of Papers

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

- I **Imsland F***, Feng C*, Boije H, Bed'hom B, Fillon V, Dorshorst B, Rubin CJ, Liu R, Gao Y, Gu X, Wang Y, Gourichon D, Zody MC, Zecchin W, Vieaud A, Tixier-Boichard M, Hu X, Hallböök F, Li N, Andersson L. (2012) The *Rose-comb* mutation in chickens constitutes a structural rearrangement causing both altered comb morphology and defective sperm motility. *PLoS Genetics*, 8(6):e1002775. doi: 10.1371/journal.pgen.1002775.
- II Sundström E, **Imsland F**, Mikko S, Wade C, Sigurdsson S, Pielberg GR, Golovko A, Curik I, Seltenhammer MH, Sölkner J, Lindblad-Toh K, Andersson L. (2012) Copy number expansion of the *STX17* duplication in melanoma tissue from Grey horses. *BMC Genomics*, 13:365. doi: 10.1186/1471-2164-13-365.
- III **Imsland F***, McGowan K*, Rubin CJ, Henegar C, Sundström E, Berglund J, Schwochow S, Gustafson U, Imsland P, Lindblad-Toh K, Lindgren G, Mikko S, Millon L, Wade C, Schubert M, Orlando L, Penedo MCT, Barsh GS, Andersson L. (2015) Regulatory mutations in *TBX3* disrupt asymmetric hair pigmentation underlying Dun camouflage colour in horses. (*Submitted manuscript*)

* These authors contributed equally

Related work by the Author

(Not included in this thesis)

1. Wright D, Boije H, Meadows JR, Bed'hom B, Gourichon D, Vieaud A, Tixier-Boichard M, Rubin CJ, **Imsland F**, Hallböök F, Andersson L. (2009) Copy number variation in intron 1 of *SOX5* causes the Pea-comb phenotype in chickens. *PLoS Genetics*, 5(6):e1000512. doi: 10.1371/journal.pgen.1000512.
2. Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, **Imsland F**, Lear TL, Adelson DL, Bailey E, Bellone RR, Blöcker H, Distl O, Edgar RC, Garber M, Leeb T, Mauceli E, MacLeod JN, Penedo MC, Raison JM, Sharpe T, Vogel J, Andersson L, Antczak DF, Biagi T, Binns MM, Chowdhary BP, Coleman SJ, Della Valle G, Fryc S, Guérin G, Hasegawa T, Hill EW, Jurka J, Kiialainen A, Lindgren G, Liu J, Magnani E, Mickelson JR, Murray J, Nergadze SG, Onofrio R, Pedroni S, Piras MF, Raudsepp T, Rocchi M, Røed KH, Ryder OA, Searle S, Skow L, Swinburne JE, Syvänen AC, Tozaki T, Valberg SJ, Vaudin M, White JR, Zody MC; Broad Institute Genome Sequencing Platform; Broad Institute Whole Genome Assembly Team, Lander ES, Lindblad-Toh K. (2009) Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science*, 326(5954):865-7. doi: 10.1126/science.1178158.
3. Bellone RR, Forsyth G, Leeb T, Archer S, Sigurdsson S, **Imsland F**, Mauceli E, Engensteiner M, Bailey E, Sandmeyer L, Grahn B, Lindblad-Toh K, Wade CM. (2010) Fine-mapping and mutation analysis of *TRPM1*: a candidate gene for leopard complex (*LP*) spotting and congenital stationary night blindness in horses. *Briefings in Functional Genomics*, 9(3):193-207. doi: 10.1093/bfgp/elq002.
4. Wang Y, Gao Y, **Imsland F**, Gu X, Feng C, Liu R, Song C, Tixier-Boichard M, Gourichon D, Li Q, Chen K, Li H, Andersson L, Hu X, Li N. (2012) The crest phenotype in chicken is associated with ectopic expression of *HOXC8* in cranial skin. *PLoS One*, 7(4):e34012. doi: 10.1371/journal.pone.0034012.
5. Boije H, Harun-Or-Rashid M, Lee YJ, **Imsland F**, Bruneau N, Vieaud A, Gourichon D, Tixier-Boichard M, Bed'hom B, Andersson L, Hallböök F. (2012) Sonic Hedgehog-signalling patterns the developing chicken comb as revealed by exploration of the *Pea-comb* mutation. *PLoS One*, 7(12):e50890. doi: 10.1371/journal.pone.0050890.

6. Andersson LS*, Larhammar M*, Memic F*, Wootz H*, Schwochow D, Rubin CJ, Patra K, Arnason T, Wellbring L, Hjälms G, **Imsland F**, Petersen JL, McCue ME, Mickelson JR, Cothran G, Ahituv N, Roepstorff L, Mikko S, Vallstedt A, Lindgren G, Andersson L*, Kullander K*. (2012) Mutations in *DMRT3* affect locomotion in horses and spinal circuit function in mice. *Nature*, 488(7413):642-6. doi: 10.1038/nature11399.
7. Promerová M, Andersson LS, Juras R, Penedo MC, Reissmann M, Tozaki T, Bellone R, Dunner S, Hořín P, **Imsland F**, Imsland P, Mikko S, Modrý D, Roed KH, Schwochow D, Vega-Pla JL, Mehrabani-Yeganeh H, Yousefi-Mashouf N, G Cothran E, Lindgren G, Andersson L. (2014) Worldwide frequency distribution of the 'Gait keeper' mutation in the *DMRT3* gene. *Animal Genetics*, 45(2):274-82. doi: 10.1111/age.12120.
8. Jäderkvist K, Holm N, **Imsland F**, Arnason T, Andersson L, Andersson LS, Lindgren G. (2015) The importance of the *DMRT3* 'Gait keeper' mutation on riding traits and gaits in Standardbred and Icelandic horses. *Livestock Science*, doi:10.1016/j.livsci.2015.03.025.
9. Velie BD*, Jäderkvist K*, **Imsland F**, Viluma A, Andersson LS, Mikko S, Eriksson S, Lindgren G. (2015) Frequencies of polymorphisms in *myostatin* (*MSTN*) vary in Icelandic horses according to the use of the horses. *Animal Genetics*, 46(4):467-8. doi: 10.1111/age.12315

* These authors contributed equally

Contents

Introduction	11
Good material is of material importance	12
Genetic variation.....	15
Chicken combs	16
Melanic pigmentation	17
Equid pigmentation.....	19
Introduction of papers	21
Paper I.....	21
Paper II	22
Paper III	24
Discussion	26
The rooster's Rose-comb	26
The strikingly Grey steed	28
Analysis.....	28
Rate and mode of Greying	28
The disappearing Dun.....	31
The Dun phenotype.....	31
Primitive markings and the non-dun horse	32
Differences and similarities	33
Predomestic horse colour	34
Lessons learnt	36
Allelic evolution.....	36
Exploring variation	37
Materials	39
Phenotyping	40
Conversations with the public	42
Future and reflections.....	44
Tables	47
Acknowledgements	49
References	53

Abbreviations

α -MSH	α -melanocyte-stimulating hormone
<i>ASIP</i>	<i>agouti signalling protein</i>
bp	basepair
<i>CCDC108</i>	<i>coiled-coil domain containing 108</i>
cDNA	complementary DNA
DNA	deoxyribonucleic acid
<i>EOMES</i>	<i>eomesodermin</i>
F ₁	first filial generation
FISH	fluorescence <i>in situ</i> hybridization
<i>FKBP7</i>	<i>FK506 binding protein 7</i>
GC-	guanine-cytosine-
IBD	identity by descent
kb	kilobase
KIT	<i>v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog</i>
<i>KITLG</i>	<i>KIT ligand</i>
Mb	Megabase
<i>MC1R</i>	<i>melanocortin 1 receptor</i>
<i>MITF</i>	<i>microphthalmia-associated transcription factor</i>
<i>MNR2</i>	<i>MNR2 homeodomain protein</i>
mRNA	messenger ribonucleic acid
<i>NR4A3</i>	<i>nuclear receptor subfamily 4, group A, member 3</i>
PCR	polymerase chain reaction
<i>PLEKHA3</i>	<i>pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 3</i>
<i>POMC</i>	<i>proopiomelanocortin</i>
SNP	single nucleotide polymorphism
<i>SOX5</i>	<i>SRY (sex determining region Y)-box 5</i>
<i>STX17</i>	<i>syntaxin 17</i>
<i>TBX3</i>	<i>T-BOX3 transcription factor</i>
<i>TRPM1</i>	<i>transient receptor potential cation channel, subfamily M, member 1</i>

Introduction

Generally speaking, most animal species are comprised of individuals that are relatively uniform in appearance and form, with individual differences mostly existing in small quantities along a cline. Individuals of the same species may be on the opposite ends of a cline and thus exhibit a pronounced difference in appearance, but one can expect to find a multitude of intermediate forms between the two extremes. Only rarely are there pronounced differences between individuals of a species without intermediate forms being more abundant than the extremes. Perhaps the most commonly observed bimodally distributed difference between individuals of a single animal species is sexual dimorphism, where the sexes differ, for instance in size, colouration or physical morphology, and there is no intermediate form to be found, discounting rare developmental exceptions such as chimeras and intersexed individuals.

This does not hold true of domestic animals. There one can find numerous traits that affect the appearance of the individual in such a pronounced way that the animal is radically different from the next individual. Yet the trait in question is discrete, so that the phenotypic extremes by far outnumber the intermediate forms. When the concept of Mendelian genetics was first gaining acceptance amongst scientists, these monogenic traits so common in domestic species were invaluable in elucidating how inheritance works. Consequently, our modern understanding of genetics is built upon the foundation of monogenetic variation in domestic animals and plants, as well as in model organisms.

The probable reason for the prevalence of monogenic variants in domestic animals, in contrast to their relatively monomorphic wild relatives, is in all likelihood to be found in the effect humans have exerted upon the species in question. The protective environment provided by humans has relaxed selective pressure on the animals through increased survival of individuals with unusual characteristics. Our fascination with the unusual has also led to selective propagation and dissemination of novel phenotypic traits in domestic animals.

As the field of genetics matured, geneticists began moving onto traits with more complex inheritance than the monogenic traits, attempting to elucidate how polygenic traits are affected by inheritance. When molecular genetics first emerged, monogenic traits were again in vogue as their inheritance and phenotypic effects had often been well characterised, and

their linkage to other monogenic loci had led to the elucidation of linkage groups that were facilitative of molecular analysis. Analogous traits had been discovered in different species, so that when a disruption of a gene was found to cause a phenotypic variant in one species, that particular gene could be examined in other species exhibiting similar phenotypes. This led to a great number of monogenic traits being explained on the molecular level, where the sequence of a gene was in some way altered, resulting in a protein product that was either subtly or radically different from the unaltered form. However, not all monogenic traits immediately offered up easily targeted genes for examination, and were thus not explained on the molecular level in those early days.

In recent years the field of molecular genetics has moved forwards, not in strides, but in leaps and bounds. An assembled sequence of the human genome was made available in the year 2001, after an enormous collaborative effort¹. In the decade and a half that has passed since, genome sequences of a multitude of eukaryotic species have been made public, and whole genome sequences from multiple individuals of certain species have been analysed and made available to the public. This has provided a new avenue of investigation into both mono- and polygenic traits. A good reference sequence for a species enables not only improved capacity for molecular mapping of genetic traits, but also for the elucidation of regulatory variants and structural variants that would be hard to assess without a reference genome assembly, thus giving new insights into the biology affected by such variants. As before, the monogenic trait is an ideal candidate for the forefront of a new and exciting wave of genetic discoveries.

Good material is of material importance

When attempting to discover the genetic cause of a trait, the potential for success is only ever going to be as good as the genetic material upon which the investigation is founded. One factor affecting the choice of material is the selected mapping strategy. Two major mapping strategies are generally employed when examining monogenic traits.

Identical by descent (IBD) mapping relies upon haplotype information. Assuming that a trait has descended down through the generations from a common ancestor, all individuals exhibiting the trait should share a specific sequence surrounding the causative variant. The best results with IBD mapping are obtained when the individuals used are as distantly related as possible, on the principle that the genomic region uniting the individuals becomes gradually smaller as more recombination events take place.

Pedigree mapping, on the other hand, focuses on examining pedigrees of individuals, associating the inheritance of a trait through the generations with

the inheritance of a genomic region through the generations, relying upon recombination events occurring within the pedigree for narrowing down the region. Both methods have their merits, but in many cases it is advisable to utilise both if possible, using the methods to complement one another, and thus strengthening the results.

Another element of sample selection that is of critical importance is accurate phenotyping. Even if phenotyping is done with extreme carefulness, mapping results can always throw a new spin on the phenotype, and therefore it is advisable to have as detailed a record of sampled individuals as possible, including pedigree information, detailed descriptions, and photographs. Depending upon the trait being studied, samples of tissues may be necessary for proper elucidation of the genetic information. It is also wise to ensure that animals sampled for purposes other than genetic mapping are phenotyped with care, and that a DNA sample is secured from them in the eventuality that an unforeseen genetic complication is discovered.

Mapping a trait to a genomic locus will require some kind of genotyping. Smaller scale methods like Sanger sequencing can be used to genotype a few individuals at a limited number of loci. Intermediate scale methods such as electrophoretic fragment length analysis of both SNPs and structural variants, and SNP genotyping with TaqMan or pyrosequencing are suitable to type a larger number of individuals for several loci. Larger scale methods, for instance SNP chip analysis or variant calling from large scale sequencing data, are used when more in-depth information is required. The genotyping results gained with all of these methods can then be used in conjunction with phenotypic data to establish linkage of a trait to a genomic region, and consequently to narrow the region down, hopefully ending with the causative variant being discovered.

Once a genomic region has been found to be associated with a trait, figuring out which genetic lesion in the region is the causative one can prove a daunting task. This holds true in particular when none can be found in an obvious place, such as in the coding sequence of a gene, or within a regulatory element that is well conserved in a number of distantly related species. Thus it can be easier to figure out causative variants for traits that are of fundamental biological importance than it is to figure out causative variants for traits that are particular to a narrower taxon, or even to a single species. This is where novel methods in molecular genetics can be of value, enabling researchers to come upon new biological insights that were previously obscured due to the limitations in scale inherent to classical molecular genetics.

Novel methods of sequencing that create enormous quantities of sequence in a short period of time have arisen in the past few years. Continuous improvements are being made to these second generation sequencing methods. The wet-laboratory end of the process generates more sequence with each update. The sequences are also becoming longer, more

informative, and more accurate. The datasets generated are far larger in scale than anything previously available in molecular genetics, and the bioinformatic methods used to make sense of them are likewise being advanced and developed.

The enormous quantity of data produced in a single second generation sequencing reaction is both an asset and a liability. When a genome, or a genomic region, is sequenced in its entirety with sufficient read depth it can fairly safely be said that there will not be a lot of information lost. However, this means that the researcher will have to face the luxury problem of having too much information. Polymorphisms can be variably dense between species, populations and genomic regions, but when a researcher is given information on every single genetic variant in a target region, the list of possible candidate variants can become dauntingly large.

Several approaches can be made to minimise this problem. Samples from individuals with and without the phenotype under investigation can be selected beforehand to be as similar in the target genome region as possible. As an example, if a sample is available that has the ancestral haplotype on which the causative variant is expected to have arisen, sequencing it will eliminate a great number of variants present on other haplotypes that are not relevant to the trait being studied. However, this is not always a feasible strategy, in particular when the variant being studied is sufficiently old that accumulation of additional variants since the time of divergence begins to pose a problem. Drift and selection may also have caused the ancestral haplotype to become rare and hard to find.

Another method is to narrow the genomic region of interest down as much as possible prior to sequencing. Alternately, a larger cohort of individuals likely to narrow the region further down can be screened for a subset of discovered variants from the sequencing data, for instance with a SNP screen. This can provide improved resolution and power to detect potentially causative variants.

It can also be worthwhile to limit what is sequenced in an effort to minimise irrelevant data. Sequencing an entire genome in search of a single genetic variant within a relatively limited genomic region would be fairly wasteful if the remaining data will not be used for any other ends, when it is possible to gain more data on the region of interest by excluding genomic areas not under investigation. Targeted sequencing can for instance be accomplished by sequence enrichment, where the target sequence is captured with probes and subsequently isolated from the genome. Whilst this will give greater read depth in the regions of interest, the use of probes means that regions within the area of interest that are not amenable to probe design for one reason or another are lost from the sample to be sequenced. The same can be said if the area to be investigated is amplified with PCR methods. The placement of primers, the structure of the genomic area and GC-content can all affect how well the DNA obtained for sequencing

represents the *in vivo* state of the area. For instance, hard to amplify sequences, and insertions and duplications in a heterozygous state might be underrepresented or go entirely uncaptured.

A researcher can also aim to kill two birds with one stone when choosing individuals to sequence. If several traits are simultaneously under investigation, samples can be chosen so that cases at one locus serve as controls for other loci. Whole genome sequences can also be utilised for other ends than direct search for causative variants, the data collected in such a reaction can also be used for numerous other research endeavours, such as population genetics and genomics, as well as studies on genome biology and to aid in modelling the biochemistry of DNA.

Genetic variation

Genetic variation at the sequence level comes in many different forms. Not so long ago microsatellites, short stretches of oligonucleotide repeats of variable lengths, were what researchers relied upon in their mapping endeavours. The most commonly investigated form nowadays is the SNP, where a single base pair differs between individual chromosomes. Other types of variants are less commonly employed for mapping due to their nature not lending itself to large-scale genotyping, and lower density in genomes.

Copy number variants are structural changes to a DNA sequence, where the amount of DNA varies between homologous chromosomes. These can take the form of deletions, insertions and duplications. A deletion event results in loss of sequence from a chromosome, an insertion adds novel sequence to a chromosome, and a duplication event expands sequence already present in a genome.

Some structural variants do not necessarily involve a change in the quantity of DNA, but rather affect the spatial relationship of DNA. An inversion involves a section of a chromosome changing orientation relative to the rest of the chromosome, whereas a translocation moves a section of a chromosome to another genomic location, either elsewhere on the same chromosome, or more commonly to another chromosome entirely. Translocations can be DNA quantity neutral, or involve partial deletions or translocational duplications.

The nature of structural variants can have implications for how the chromosome behaves during replication and recombination, with larger structural variants having more pronounced effects than smaller.

Duplications can wind up misaligning, leading either to a reversion of a wild-type state of the chromosome, or to an expansion of the duplication². Translocations lead to a different problem, as quadrivalent chromosomal alignment during meiosis in an individual with a large reciprocal

translocation will often result in gametes with unbalanced chromosomal complement, where large portions of chromosomes can either be duplicated or deleted during recombination and chromosome segregation³.

In individuals heterozygous for inversions, successful recombination events inside the inverted sequence are unlikely. On one hand, such recombination events can lead to unbalanced gametes, via the generation of a- and dicentric chromosomes, or deletions and duplications, dependent upon whether the inversion is para- or pericentric. Another possible outcome is repression of recombination within the inverted sequence due to lack of sequence homology if an inversion loop fails to form during recombination⁴. In surviving gametes and zygotes this results in an excess of recombination events just beyond the inversion breakpoints.

Chicken combs

There are four different extant species of wild junglefowl, the males of which sport a comb, a sexual ornament. The combs are different in size, shape, texture, and colour between the species⁵. In domestic chickens (*Gallus gallus domesticus*), both hens and roosters have combs. This is in contrast to what has been assumed to be true of their ancestor, the Red Junglefowl, where the assumption is that hens are essentially combless⁶. The ancestral state of the comb is, however, hard to assess, as most populations of wild Red Junglefowl are assumed to have crossbred with domestic chickens to some degree⁷.

Although the comb is present in both domestic hens and roosters, it is a sexually dimorphic trait. Roosters in a given population will generally have more sizeable combs than hens from the same genetic background. In spite of this, the shape of the comb is relatively similar in hens and roosters. However, comb morphology in domestic chickens is far from being uniform. The wild-type comb shape is called a Single-comb; it is a flap of tissue extending upward from the head, serrated as a very coarse comb. There are three major loci known to have a pronounced effect on comb shape in chickens, each of which is associated with a particular comb form dominant to the Single-comb; the Rose-comb, Pea-comb, and Duplex-comb (Figure 1).

William Bateson described the very first traits to behave according to Mendel's laws of inheritance in animals in 1902, and among those was the Rose-comb⁸. The standard Rose-comb is wider than a Single-comb, large and fleshy, has numerous papillae on the dorsal surface, and sports a spike at the posterior of the comb.

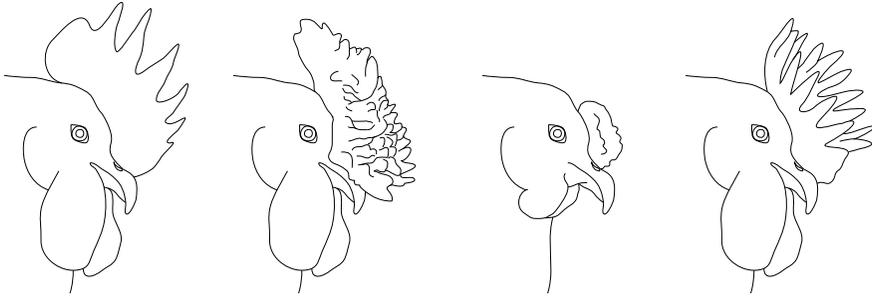


Figure 1. Illustration of chicken combs. From left to right: Single-comb, Rose-comb, Pea-comb, Duplex-comb. (Drawing: Freyja Imsland.)

A few years after the inheritance of Rose-comb was first described, Bateson and Reginald Punnett reported for the very first time an epistatic interaction between two genetic loci in an animal. Their seminal paper described how the co-presence of two dominant alleles at separate loci, *Rose-comb* and *Pea-comb*, results in a Walnut-comb, a comb shape quite unlike Rose, Pea, and Single^{9,10}. Illustration of these four different comb shapes can be found in Figure 1 of Paper I.

Pea-comb is caused by a copy number variant in the first intron of *SOX5*, where a 3.2 kb sequence duplication has expanded, leading to ectopic expression of *SOX5* during embryogenesis in the comb-forming region of the mesenchyme¹¹. Duplex-comb is caused by a 20 kb duplication upstream of *EOMES*, leading to ectopic expression of *EOMES* in the ectoderm of the comb-forming region during embryogenesis¹².

Further studies of *Rose-comb* after Bateson and Punnett's initial report revealed it to be a complicated trait. It is monogenic, but with two facets to it. The most readily apparent one is the dominant inheritance of an altered comb shape observed in both roosters and hens. Less apparent is that severely reduced rooster fertility is a recessive trait associated with the Rose-comb phenotype¹³⁻²⁵. This duality in phenotype and dominance associated with a single trait suggests that a causative variant might not be of a nature so simple as disrupting the coding sequence of a single gene.

In spite of the negative effect homozygosity for *Rose-comb* can have on rooster fertility, the Rose-comb is a trait that has been selected for in many chicken breeds, presumably due to the decorative appearance of the comb itself.

Melanic pigmentation

Many species of domestic animals have great pigmentary variation, affecting skin, irides, dermal appendages such as hairs, feathers, and scales, and even

other keratinised appendages such as beaks, hooves, horns, spurs, and claws. Birds exhibit an array of different colourants; ingested carotenoids and porphyrins provide pigmentation in addition to structural colouration²⁶. Another major pigment class in birds are the melanins.

Unlike in birds, mammalian pigmentation is nearly only determined by melanins. Mammalian surface melanins can be divided into two major classes, the eumelanins and the phaeomelanins. In general, eumelanins are black or brown in colour, and phaeomelanins are red, orange, and yellow²⁷.

Melanins are produced by melanocytes, specialised cells of neural crest origin²⁸. Melanocytic precursors in the neural crest migrate from the dorsal midline of the embryo. Along the way they proliferate and differentiate to eventually populate the epidermis, including hair and feather follicles^{26,29}. This process of migration and survival of melanocytes and their precursors is tightly regulated, with several genes known to be of vital importance. Defects affecting this pathway tend to lead to a lack of melanocytes in patterns that are characteristic for each particular defect. Many variants of these genes have been found to be under selection in various domesticated species, as humans have often favoured the spotting phenotypes associated with them. *KIT* and *MITF*, for instance, have known variants associated with coat colour in cattle³⁰⁻³³, pigs³⁴⁻³⁹, donkeys⁴⁰, dogs⁴¹⁻⁴⁴ and cats^{45,46}.

It is not enough for melanocytes to migrate to the correct location for an animal to become pigmented. Genetic mapping of pigmentary variants in both model organisms and other species has shown that the entire process of melanogenesis within the melanocyte is highly complex, involving cell survival and proliferation, as well as regulation, production, processing and deposition of pigment, with many genes affecting each part of the process⁴⁷.

Perhaps the most readily recognisable regulatory effect involving melanogenesis is that which results from pigment type switching, the process that determines the proportions of eu- and phaeomelanin produced by the melanocyte. The major players determining pigment type are the trio of *MC1R*⁴⁸, *ASIP*^{49,50} and *POMC* (encoding for α -MSH)⁵¹, where *MC1R* signalling determines whether the cell is directed towards eumelanogenesis or phaeomelanogenesis. α -MSH is an agonist of *MC1R*, initiating production of eumelanins, whereas *ASIP* is an antagonist, blocking α -MSH from binding to *MC1R*, leading to production of phaeomelanins⁵².

The pigmentary appearance of an animal is established by how melanocytes migrate, survive and proliferate in the tissues of an animal, and subsequently by melanogenesis. The determinants of the colour phenotype of each individual are the absolute intensity of pigmentation, and the relative intensity and distribution of eu- and phaeomelanins, in both time and space.

To gain a better understanding of what that entails, it can be of use to consider the pigmentation of different species. The red fox (*Vulpes vulpes*) provides a good example of different spatial distribution and intensity patterns of eu- and phaeomelanins. It has a red phaeomelaninic body fur, with

the legs covered in black eumelanic hairs, showing differing spatial distribution of the classes of melanins. The chin, chest and belly of the red fox are adorned with white hairs with sparse, if any, pigmentation, demonstrating different relative intensity of pigmentation.

The arctic fox (*Vulpes lagopus*) presents a different profile, illustrating temporally differing pigment intensity. One of the colour morphs is a dark brown eumelanistic colour when it has a short summer coat, but when autumn arrives and the long winter coat grows out, it is unpigmented and white as the driven snow. Then, come spring, the white winter coat is shed, and underneath there is a dark brown eumelanistic coat again.

Equid pigmentation

The domestic horse (*Equus ferus caballus*) exhibits a multitude of pigmentary variants, many of which have been associated with sequence variations⁵³⁻⁶⁶. Most of these variants affect the coding sequence of genes, or have an effect on the sequence of a gene's mRNA. Other pigmentary traits in the horse are caused by structural variants that have an effect on gene regulation. For instance, the Tobiano spotting pattern is associated with a large inversion that disrupts the extensive regulatory region of the gene *KIT*⁶⁷, the Leopard Spotting Complex has been associated with a retroviral insertion in *TRPM1*⁶⁸, and Greying with Age is associated with a 4.6 kb intronic sequence duplication in *STX17*⁶⁹.

Greying with Age is a depigmentation phenotype, where a horse born with any given coat colour gradually loses hair pigmentation due to what might be depletion of melanocytic stem cells in the hair bulbs⁶⁹, but retains general pigmentation of skin, irides and hooves. However, Grey horses can develop vitiligo of the skin as they age, and they are also prone to develop melanomas with the years⁷⁰. The melanomas of Grey horses are in general not cancerous or aggressive, with discomfort or debilitation due to location of tumours being more of a threat to the animals' health than aggressive metastasis⁷¹⁻⁷³. Some melanomas in Grey horses are however far from benign, and show aggressive profiles⁷⁴.

Another equid coat colour, one with no molecular research upon it published, is Dun. It is the ancestral wild-type pattern of pigmentation in equids, characterized by specific patterns of intensely pigmented primitive markings on a diluted background (Figure 2). The manifestation of Dun subtly differs in various ways between extant equid species, but with an underlying theme of striping and mottling attributable to variable pigment intensity of both eu- and phaeomelanins. The Dun pattern has been lost in the majority of domestic horses, and they are as a consequence non-dun. Given that diversity of Dun patterning is seen in wild equids, it is not unreasonable to assume that a regulatory variant might be causative for the

loss of pigmentary patterning in non-duns, as differences in regulation of the gene responsible for the patterning are presumably important to establish the variety of patterning observed in wild equids.

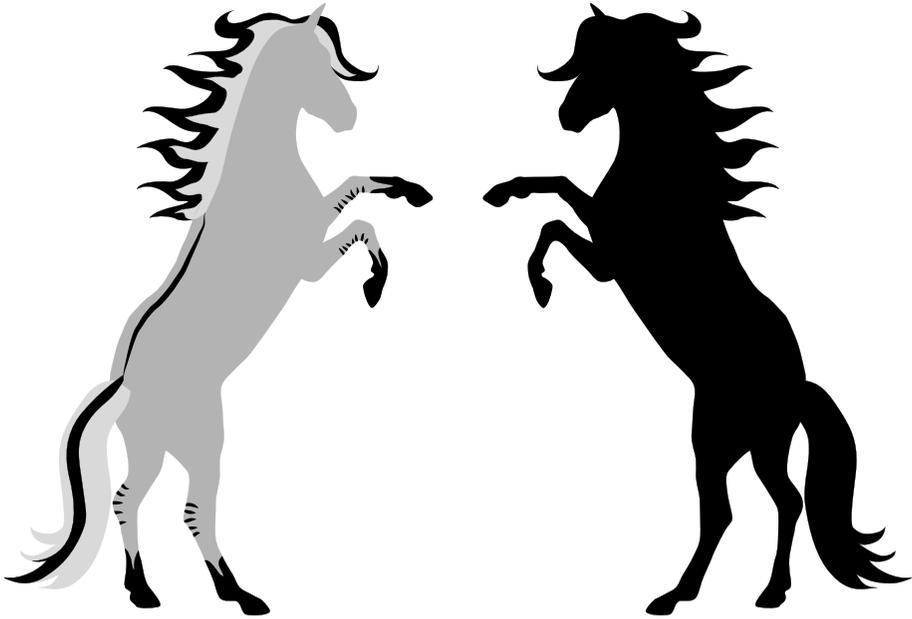


Figure 2. Illustration of a Blue Dun and a Black horse. (Drawing: Freyja Imsland.)

For clarification of discussion from here on, a summary of genes with molecular association to pigmentary variants in the domestic horse can be found in Table 1 (page 47). Not all of these have been connected to a causative variant, even though they have been mapped to a gene. Known coat colour variants without published molecular associations are not included. Tables 2 and 3 present a summary of allelic notation and nomenclature for some equid colour variants (pages 47-48). In Table 4, colour nomenclature in English is further elucidated in Icelandic, Swedish and German (page 48). Readers are advised to refer to these tables when confronted with unfamiliar terms for horse coat colours in this work.

Introduction of papers

Paper I

The *Rose-comb* Mutation in Chickens Constitutes a Structural Rearrangement Causing Both Altered Comb Morphology and Defective Sperm Motility

The aim of this investigation was to discover the causative variant behind the Rose-comb phenotype. The task was begun with pedigree mapping. Considering the fertility issues homozygous Rose-combed roosters can have, and the implications this can have for the zygosity of a Rose-combed population⁷⁵, a pedigree generation strategy that avoided these issues was chosen. An F₁ mapping population was generated by crossing heterozygous Rose-combed sires and homozygous wild-type dams, thereby generating about 50% Rose-combed, and 50% Single-combed offspring, entirely avoiding homozygotes for *Rose-comb*.

Mapping by microsatellite and SNP analysis revealed a large region with suppressed recombination in association with the Rose-comb trait, indicating an inversion. Second generation sequencing of a mate pair library from a pool of Rose-combed birds was brought into play in an effort to aid in finding the breakpoints of the suspected inversion. The birds were from the French breed Le Mans, which has been selected for a classical Rose-comb phenotype.

Analysis of mate pairs with aberrant orientation and mapping distance revealed two different alleles in the Rose-combed Le Mans pool. One of the alleles, *R1*, is the result of 7.38 Mb inversion. The other, *R2*, is the result of a recombination event between *R1* and a wild-type chromosome, whereby most of the chromosome is restored to a wild-type state, but retaining one of the original inversion breakpoints as part of a duplicated sequence, 91 kb in length. The *R1* inversion transfers the *MNR2* gene into a new genomic context, and severs the *CCDC108* gene between exons 3 and 4. The recombined *R2* allele restores an intact *CCDC108*, but retains the novel context of *MNR2*. The conformation of *wt*, *R1* and *R2* chromosomes, and the makeup of the breakpoints involved, is illustrated in Figure 3 of Paper I.

MNR2 codes for an Mnx-class homeodomain protein, a class of proteins known to act as transcriptional repressors involved in the specification of cell identity⁷⁶. Interestingly, hyaluronic acid, a major component of the extracellular matrix that is found in great concentrations in the comb of

chickens, shows strong accumulation around early *MNR2*-expressing neurons during embryonic development⁷⁷. PCR of cDNA derived from embryonic comb tissue revealed ectopic expression of *MNR2* in the region of the comb on the 9th day of embryonic development in Rose-combed embryos, but not in wild-type embryos, strongly indicating that this novel genomic context of *MNR2* is causative for the altered comb morphology observed in Rose-combed birds. Histological experiments lent this scenario further support.

On the other side of the inversion, the *CCDC108* gene is disrupted, with the first three exons located 7.38 Mb from their original context and in an inverted orientation, effectively deleting a part of the gene. As *CCDC108* is a gene implicated in testis development and flagellar movement it is an obvious candidate for the reduced fertility observed in homozygous Rose-combed roosters^{78,79}. Testing this hypothesis, a mating experiment was performed with roosters of various Rose-comb genotypes. Roosters homozygous for the *R2* allele revealed no depression of fertility when compared with roosters homozygous for a wild-type allele at the Rose-comb locus, whereas *R1/R1* roosters showed the expected subfertility.

Paper II

Copy number expansion of the *STX17* duplication in melanoma tissue from Grey horses

This paper reports two avenues of investigation into the intronic duplication in *STX17* that had previously been reported to be in association with Greying with Age in the domestic horse, a phenotype of gradual hair depigmentation associated with melanomas and vitiligo^{69,72}. One was to investigate whether expansion of the duplication could be observed in Grey horses, and if so, whether the expansion was correlated with incidence and aggressiveness of Grey-associated melanoma. The other was to attempt to investigate the region surrounding the duplication in an effort to examine whether other variants than the duplication could be causative for Greying with Age. Additionally, it was intended as proof of principle that a duplication could be detected by second generation sequencing of captured DNA.

Analysis of copy number variation for the duplication revealed that germline copy number of the duplication seems stable at two copies of the sequence per each *Grey* chromosome. Somatic expansion of the duplication in tumours is, however, more common, with aggressive tumours showing more expansion than benign tumours.

The genomic region enriched for sequencing with regard to the *Grey* locus was 352 kb in length, spanning the region containing the previously

mapped *Grey* haplotype⁶⁹. A single Lipizzaner horse, homozygous at the *Grey* locus, was chosen as a case. A second horse, a non-grey Arabian homozygous for a haplotype assumed to be closely related to the haplotype ancestral to the *Grey* mutation, was chosen as a control with the aim of identifying all genetic variants unique to the *Grey* haplotype. Thereby providing an opportunity to strengthen the argument that the previously observed duplication is indeed the causative variant. Five other non-grey horses were included in the experiment, an Icelandic, a Quarter horse, an Appaloosa, a Knabstrupper, and a Noriker. Three other genomic areas, relevant to other phenotypes, were also included in the study design, with the horses chosen for their *Grey* locus alleles serving as controls for the three other loci.

A custom glass-slide based probe array was designed and manufactured based on reference sequence information for the genomic areas of interest. After library preparation each sample was enriched by separate hybridisation on a probe array, enriching the samples for the target area sequences. Following post-enrichment amplification, sequencing was performed, yielding 35 bp reads.

Read alignment to a reference sequence revealed a striking read depth spike for the duplicated region in the single Grey horse, as illustrated in Figure 3 of Paper II, clearly indicating that this method of enrichment and sequencing is effective at detecting copy number polymorphisms like the *Grey* duplication. No other structural variants could be found in the Grey Lipizzaner horse with this method, although a note should be made that power to detect insertions, deletions, and inversions is very limited with the exact method employed. Paired reads and longer reads would be better suited to detecting such variants.

A considerable number of SNPs was found in the region, numbering in excess of 750 when the sequencing reads were mapped to a repeat masked reference sequence, averaging to about 2 SNPs per kb. However, the experimental design, with a single individual homozygous for the trait of interest contrasted with six individuals not carrying the trait, one of which had a haplotype ancestral to the *Grey* haplotype, enabled us to exclude all but 15 SNPs directly as being causative. This was done on the grounds that the genetic variant causative for the Grey phenotype had to be present in the Lipizzaner, most likely in a homozygous state, as well as not being found in any of the six non-grey horses.

Further analysis of possibly causative SNPs revealed that so far the 4.6 kb duplication is the only known polymorphism showing perfect association with the Greying with Age phenotype.

Paper III

Regulatory mutations in *TBX3* disrupt asymmetric hair pigmentation underlying Dun camouflage colour in horses

This manuscript details investigations into the genetic basis of Dun and non-dun phenotypes in the domestic horse and other equids. The *Dun* locus was explored by IBD mapping. This was accomplished by a number of methods, including microsatellites, single SNP assays, SNP panels, and second generation sequencing of both captured DNA and whole genomes of modern equids, as well as from ancient DNA samples. The *Dun* locus mapped to a genomic region containing the gene *T-box transcription factor 3 (TBX3)*. Loss of function mutations for *TBX3* are associated with ulnar mammary syndrome⁸⁰⁻⁸², but hitherto it has not been known to be involved in pigmentation.

A microsatellite mapping approach indicated a region on the p-arm of chromosome 8 in association with Dun. A screen using 27 SNPs further narrowed it down to a 200 kb region containing only *TBX3*. Those 200 kb were captured for second generation sequencing in the same experiment as described for Paper II, with the same seven horses used; an Icelandic and a Quarter Horse, both homozygous for *Dun*, as well as the aforementioned non-dun Lipizzaner, Arabian, Noriker, Appaloosa and Knabstrupper. Further sequences for the region were obtained from whole genome sequencing of numerous domestic horses, Przewalski's horses (*Equus ferus przewalskii*), as well as other modern equids. Sequences from ancient DNA from a museum sample of the extinct Quagga (*Equus quagga quagga*), and archaeological samples from two wild horses, ~4.400, and ~43.000 years old, were also included in the analysis.

SNPs discovered in the sequencing data were used to generate a SNP panel to assess the region in a larger number of horses. Two different causative variants for non-dun in horses were found from the combined sequencing data and SNP screen. One variant, termed *non-dun1 (d1)*, differed from the observed *Dun* alleles by two SNPs, termed *SNP1* and *SNP2*. The other variant, termed *non-dun2 (d2)*, had evolved from a *d1* allele, by the deletion of a 1.6 kb region 5 kb downstream of *TBX3*. This deleted region shows a high degree of sequence conservation to other mammals, and it also contains *SNP1*. As the horse genome assembly is based on a non-dun horse that is homozygous *d2/d2*, these 1.6 kb were missing from the horse reference genome.

Genotyping of over 1900 horses established that the *d2* allele is widespread among domestic breeds with European origins. It also revealed a rare *Dun* allele found in Estonian Native horses, which excluded *SNP2* as being required for a Dun phenotype. Genotyping of Przewalski's horses, kiangs (*E. kiang*), onagers (*E. hemionus*), an African wild ass (*E. africanus*),

plains zebras (*E. quagga*), mountain zebras (*E. zebra*), and Grévy's zebras (*E. grevyi*) revealed a homozygous *Dun* genotype for both *SNP1* and *SNP2* for all 23 individuals. The same applied for the museum sample from the Quagga. On the other hand, the ~43,000 years old sample had a *D/d1* genotype, and the ~4,400 years old sample had a *d1/d1* genotype, indicating that predomestic horses were not monomorphically Dun.

Transverse sectioning of hairs from Dun horses showed that their dilute hairs are asymmetrically pigmented, demonstrating that Dun is not a variant that has an effect upon the melanogenic pathway. Sections of hairs from non-dun horses were much more uniformly pigmented, as illustrated in Figure 3. Dilute hairs from other equids, such as Przewalski's horse and African Wild Ass, also showed the same radially asymmetric pigment distribution as hairs from domestic Dun horses, cementing the ancestral state of the Dun phenotype within the equid lineage.



Figure 3. Illustration of what transverse sections of hairs from different horses might look like. They show typical pigment distribution patterns within hairs from areas dilute in Duns. From left to right; Dun, non-dun1, and non-dun2. (Drawing: Freyja Imsland.)

TBX3 codes for a transcription factor, and thus it does not by itself lead to a pigmentary variant. Investigations on skin biopsies taken from Dun and non-dun horses revealed that *TBX3* expression in the hair follicles of dilute hairs on Dun horses leads to altered expression of *KIT-ligand (KITLG)*, coding for a signalling protein with a well established role in the migration and survival of melanocytes^{83,84}. This expression of *TBX3* is radially asymmetric, resulting in inhibition of melanocyte migration and/or survival mediated by reduced *KITLG* expression. This extremely exact alteration of pigment cell distribution is disrupted by both the *d1* and *d2* variants, leading to more uniformly pigmented hairs than is exhibited in the ancestral Dun phenotype.

An interesting aspect of the biology of *TBX3* in relation to its role in horses, is that *TBX3* expression is essential to the patterning of digit III as has been demonstrated in chickens^{85,86}. Digit III happens to be the sole remaining digit in modern equids^{87,88}, intimating that *TBX3* ought to be under considerable selective constraint in horses.

The mechanism we have revealed behind the Dun phenotype, where radially asymmetric pigmentation of individual hairs is responsible for the patterning of the entire animal, is entirely novel to science, and so is the implication of *TBX3* being involved in pigment regulation.

Discussion

The rooster's Rose-comb

There were several interesting complications associated with this study, some with implications for experimental design, and some with biological implications.

One of the challenges was how to interpret the sequencing data generated from the aforementioned pool of 8 Rose-combed Le Mans roosters of unknown zygosity. Disentangling the mate-pair analysis proved less straightforward than expected, with results indicating that a fair number of *wild-type* chromosomes were present in the pool, as well as a few aberrant read pairs not consistent with the majority of the results. First thoughts turned to the sequenced birds having been heterozygous, as it is well documented that fixing a population for *Rose-comb* has been problematic due to the poor fertility of homozygous males. It turned out that assumptions about the zygosity of the sequenced birds proved both correct and incorrect.

Reads were found supporting suspicion of an inversion, 7.38 Mb in length, with breakpoints located approximately at 16.50 Mb and 23.88 Mb. Confirming the presence of the inversion by amplification and sequencing over the breakpoints proved problematic due to a GC-rich area adjacent to the breakpoints. A particularly robust and effective polymerase was necessary to accomplish characterisation of the breakpoints. The proximal inversion breakpoint is at 16.5 Mb, in a 203 bp interval between the genes *PLEKHA3* and *FKBP7*. The distal breakpoint is located at 23.88 Mb in the third intron of the gene *CCDC108*. The inverted region contains the entire *FKBP7* gene and extends over 7.38 Mb, all the way to the gene *MNR2* and the first three exons of *CCDC108* (see Figure 3 in paper I). The inversion was subsequently validated by a Fluorescence In Situ Hybridisation (FISH) experiment.

However, anomalies were discovered when typing birds for these two inversion breakpoints. Some Rose-combed birds seemed to be heterozygous for the 16.5 Mb breakpoint, yet negative for the 23.89 Mb wild-type breakpoint. The key to this proved to be the three aberrant read pairs from the second generation sequencing that connected the proximal breakpoint to an area around 91 kb proximal to the distal breakpoint (see Figures 2 & 3 in paper I). A plausible scenario for how this could have happened is an unequal recombination event resulting in a new allele at the *Rose-comb*

locus, one that reintroduced the entire wild-type sequence of the chromosome, but maintained the 16.5 Mb breakpoint found in the *Rose-comb* allele, with a translocated *MNR2* gene. This scenario explains the aberrant reads and breakpoint amplification results that the presence of wild-type chromosomes could not account for.

To investigate this proposed secondary rearrangement, a FISH experiment was performed, confirming the presence of three different alleles at the *Rose-comb* locus; one *wild-type* allele (*wt*), a widespread *Rose-comb* allele with both altered comb morphology and reduced fertility in homozygous roosters (*R1*), and a more rare allele with altered comb morphology (*R2*), but possibly no adverse effect on fertility, as the second derived allele has an intact version of the *CCDC108* gene. Mating experiments carried out are in support of the hypothesis that *R2/R2* roosters do not show subfertility as *R1/R1* roosters do.

Another complication encountered in this project was when birds that were presumed to be homozygous *wild-type* consistently typed positive for the original inversion allele *R1*. This held true even when different tissues that were collected and isolated on separate occasions were used, making sample contamination an unlikely explanation. This turned out to be due to a highly atypical *Rose-comb* phenotype, a phenotype very reminiscent of a *Single-comb*, but distinguishable if one is aware of what to look for. This, combined with genotyping of a few dozen Icelandic chickens, a landrace breed that has not been subject to particularly stringent selection, unlike most modern chicken breeds, revealed that diversity in comb shape among *Rose-combed* birds is quite extensive, contrary to the traditionally held view. An illustration of this diversity of the *Rose-comb* phenotype can be found in Supplementary Figure 1 of Paper I.

The phenotypic plasticity in the shape of the *Rose-comb* suggests multiple modifying factors affecting comb size and shape. A reminder that even for traits with discrete phenotypic extremes due to monogenic loci, there still is a polygenic influence upon those traits, which results in a more or less continuous spectrum of variance within the monogenic variant.

Without employing second generation sequencing it is unlikely that both chromosomal variants connected with the *Rose-comb* phenotype would have been discovered without much more effort. It was pure happenstance that the breed chosen for the resequencing effort happened to carry the second derived *Rose-comb* allele, which appears to be far more rare worldwide than the original allele, indicating that it arose only relatively recently. Similarly, a lone cockerel with a slightly odd comb and aberrant genotyping results indicated that the phenotype itself is not as set in stone as might otherwise have been assumed.

The strikingly Grey steed

Analysis

The analysis of the *Grey* region was potentially complicated by the fact that the horse upon which the genome reference sequence is based, the Thoroughbred mare Twilight, is heterozygous at the *Grey* locus. This necessitated that the sequence of other non-Grey individuals be used as a wild-type template when considering the plausibility of causation for a particular variant. Thus it would not be sufficient proof or disproof of association with Grey if a variant discovered in a Grey horse was found or not found in the reference sequence assembly.

Having taken that into account and compensated by inclusion of both non-greys of varied origins but unknown haplotype composition, and a non-grey known to have extensive haplotype similarity to the *Grey* haplotype, the data analysis could be tailored to the situation. Out of the more than 750 SNPs discovered in the seven sequenced individuals, 15 were found to have an allele only seen in the homozygous Lipizzaner. Upon further investigation with Sanger sequencing, five of these SNPs proved to be false positives, and 10 were true variants. Out of these 10 SNPs, seven came up with a unique allele in the Lipizzaner, for three it shared an allele with the reference genome, indicating that these three SNPs occurred in a region hailing from Twilight's *Grey* chromosome.

Being aware that Twilight is indeed a heterozygous *Grey* was not the only factor complicating the experiments detailed in Paper II. Studying genomic changes in melanomas can be a daunting undertaking. Complications common to studies of tumour DNA, such as limited amount of material, genetic heterogeneity of tumours, and so forth are to be expected. In addition, melanomas produce melanins, strong PCR inhibitors. As most current laboratory methods in genetics rely on PCR in some capacity, this can prove a steep hurdle. Special attention must be paid to removal of melanins from samples during the isolation process. If, as one might assume, aggressive melanomas in Grey horses produce more melanin than non-aggressive Grey melanomas, this could have an impact upon the effectiveness of copy number detection in melanoma samples, and lead to underestimation of copy number expansion.

Rate and mode of Greying

As stated in Paper II, studies have found that Grey Lipizzaner horses have largely depigmented hair by the age of 6-8 years, depending upon their zygosity for *Grey*⁶⁹. However, observing horses from other breeds that have a more variable genetic background, in particular as regards colour, will

reveal that this is not a universal rate of depigmentation. An example of this is presented in Paper II, of a Connemara horse that shows slow progress of Greying, being still largely darkly pigmented at the age of 14. Other related Connemaras show similarly slow Greying, but it is unknown whether this slower rate of depigmentation is in linkage with the *Grey* locus, independently inherited, or even possibly polygenic.

Personal observations of Grey horses over the years indicate that multiple factors affect the rate of hair depigmentation, as well as the visual presentation and development of the Grey phenotype. Some of these factors are demonstrably unconnected to the *Grey* allele itself, and others are of uncertain origin.

One commonly variable and well-known factor that affects the rate of hair depigmentation is the colour a horse is born. Variants leading to decreased effectiveness of melanogenesis often cause horses to become white more speedily. For instance, a Palomino born Grey horse will be lighter in shade more quickly than a Chestnut born Grey will, presumably because increased melanogenesis driven by *Grey* is insufficient to overcome the already present defects in melanogenesis.



Figure 4. Icelandic horses depicting various phenotypic aspects of the Greying with Age variant. See text for details. (Photographs: Freyja Imsland and Páll Imsland.)

Figure 4 illustrates some aspects of the effects Greying with Age has on the coat colour of various Icelandic horses. Panels A and B show two four winter old stallions of unknown birth colour, who are going Grey at a similar pace, but in a different manner. Stallion A is Iron Grey, with a fairly even

admixture of pigmented and depigmented hairs, with a cline towards greater depigmentation at the hindquarters of the horse. Stallion B, however, is Dappled Grey, where the depigmentation is concentrated in flecks of white hairs on the generally dark pelage, with the greatest depigmentation seen on the head. Panel C shows a seven winters old Dappled Grey stallion born Chestnut, who has become a bit lighter than Stallion B, showing more pronounced Greying of the forequarters, but in contrast to Stallion A, the hindquarters are more pigmented.

In panels D and E two mares are depicted. They are closely related, mare D is the daughter of mare E's half-brother. Both have inherited their *Grey* allele from the heterozygous *Grey* dam of mare E. Both mares are heterozygous *Grey*. The images are taken when the mares were 8 winters old. As can be seen, mare D has greyed noticeably on the head, but is in general a very dark Dappled Grey at this stage. Mare E, on the other hand, has lost nearly all pigment, except in the tail and around leg joints. Moreover, she already shows considerable freckling on the head, where somatic loss or deactivation of the *Grey* duplication reinstates pigmentation in small flecks. As she aged, she showed progressively more freckling, leading to a Flea-bitten Grey phenotype. Mare D's foal in the image is non-grey, but all of her Grey foals have shown a relatively slow rate of Greying, similar to their dam. Mare E's foal in the image is born Bay, and is already very depigmented in the first coat of hair grown after birth. This relatively rapid depigmentation has been true of all her other Grey foals. Observing the greying rate of the mares' foals and not knowing of their relatedness, one might assume that the rate of greying is in linkage with *Grey*. But given that their *Grey* chromosomes are only separated by 3 meioses, two explanations come to mind. The less likely explanation would be a novel linked modifier acquired by one of the *Grey* alleles, and the more likely an unlinked modifier in either one of the mares. An unlinked modifier might either be present in a homozygous state, or have by chance segregated in a fashion reminiscent of linkage.

Panels F and G show two yearling fillies. Filly F is born Red Dun and is already nearly completely white. Those hairs that remain pigmented appear black in colour, either because the *Grey* melanogenic drivers result in a pigmentation switch into eumelanogenesis, or the phaeomelanin produced is in locally sufficient quantity to appear black. Filly G is born Chestnut with a blaze, and is Rose Grey at this age, showing some darkening of the mane, tail and legs, and an admixture of depigmented hairs and hairs pigmented by phaeomelanin on the body. Rose Grey horses generally develop darker pigment in pigmented hairs as they age, whilst the depigmented hairs increase in number, giving them an Iron Grey cast before turning white. This mirrors the lack of phaeomelanin hue seen in filly F.

Panel H shows the face of a Black born Grey filly half-way through shedding the birth coat, illustrating how rapid the depigmentation of hair can

be, whilst both skin and eyes remain fully pigmented. Panel I shows a Bay born five winter old mare, greying at a slow rate, her forelock being the clearest indication of her Grey colour. She also shows the same lack of phaeomelanin colouration seen in filly F.

These descriptions are at best anecdotal, being based on few individuals, but they illustrate the principles of and possibilities for variation in rate, shade, and distribution of depigmentation in Grey horses, based on personal observations.

The disappearing Dun

The Dun phenotype

The Dun coat colour pattern in equids is a very striking phenotype, juxtaposing a pale background pelage with darker striping and mottling. These intensely pigmented primitive markings can be broken down into elements that can occur in different combinations. Some elements are fairly rare, but may be common in particular breeds. Other elements are nearly universal. The element that is probably the closest to being universal is the dorsal stripe, an intensely pigmented stripe extending from the forehead, along the dorsal aspect of the spine, all the way down to the tip of the horse's tail. A non-exhaustive list of other elements includes leg striping, dark legs, ear-tip markings, facial masks, patterning around the eyes, intensely pigmented areas around cranial vibrissae, cobwebbing on the forehead, shoulder crosses, shadows on the neck and/or shoulder, dorsal barbs, and netting on the neck, shoulder, and flank (Paper III - Extended Data Figure 1b). The particular combination of primitive markings in each Dun horse seems to be inherited in an allelic fashion⁸⁹. Thus combinations of markings can be population specific, and subject to both drift and direct selection.

The precise nature of the reduced pigment intensity in Dun is highly unusual; it can in essence be described as a microscopic spotting pattern, whereby the pigment distribution within each hair is radially asymmetric. If a single hair from a paler part of the body of a Dun horse is examined in cross section, the outward facing side of the hair is most intensely pigmented, the side facing the body is less intensely pigmented, and the sides of the hair where these two areas meet are even more sparsely pigmented (Figure 3). The extent of this depigmentation is variable from hair to hair, thus is the patterning of the primitive markings formed, with the darkest zones bearing fully pigmented hair and the paler zones bearing hair with less pigment in proportion to their paleness. In addition, individual Dun horses also show a variation in how palely pigmented they are, with some horses showing a very stark contrast between the dilute areas and the

primitive markings due to the extent of dilution, whereas other Duns may have reduced contrast due to their overall intense pigmentation.

Non-dun horses can also show primitive markings, although they are less prominent than in Dun horses. The most prominent primitive markings in non-duns tend to be present in horses with a bay base colour, where the inherent intensity differences between eu- and phaeomelanins give rise to a striking visual difference without the presence of dilution. Other dilutions can then exaggerate this even further, and give rise to Dun mimics. Buckskin horses with primitive markings can in particular be convincing mimics of Dun colouration. A trained observer can, however, distinguish these when examining a horse, as the shade of the horse's hairs is quite different when the dilution present is due to defects in the melanogenic pathway compared to the intact melanogenesis of Dun horses.

Primitive markings and the non-dun horse

The question of why primitive markings are always found in Duns, but only sometimes in non-duns has gone unanswered up until now. It has been suggested that a separate locus governing primitive markings exists in linkage to the *Dun* locus⁹⁰. That would explain why some non-dun horses show primitive markings and some not, but that begs the question why we do not find any Dun horses that lack primitive markings, horses that are uniformly dilute from muzzle to tail-tip.

The allelic series of *Dun*, *non-dun1* and *non-dun2* provides a good explanation for this seemingly inconsistent presentation. Genotyping results for the three alleles confirmed that the primitive markings seen on a strongly pigmented background in some non-dun horses are associated with *d1*, and that horses homozygous *d2/d2* do not in general show primitive markings, indicating that the *d2* deletion leads to a stronger phenotypic effect than the SNP(s) associated with *d1*.

This can be elucidated further. Consistent with previous reports, we found *Dun* to be completely dominant over *non-dun*⁹¹⁻⁹⁴, both *d1* and *d2* alleles. Examining the *d1* allele's association with primitive markings revealed that *d1* seems to be incompletely dominant to *d2*, with the primitive markings of *d1/d1* individuals generally being more pronounced than those of *d1/d2* individuals.

When hairs from non-dun horses that show primitive markings are examined in a transverse section, some asymmetry can be observed in the pigment distribution. This asymmetry is far less pronounced than in the dilute hairs of Dun horses, with a smaller proportion of each hair having reduced pigmentation, and also less reduction of pigmentation in those areas (Figure 3).

To add to the complexity, primitive markings in non-duns have a variable penetrance. Both base colour, and generalized pigment intensity dependent

upon polygenic effectiveness of melanogenesis affect the manifestation of the *d1* allele. The generally less intense nature of phaeomelanistic pigmentation makes primitive markings easier to observe in Chestnut *d1/-* individuals. In Bay horses, where both eumelanins and phaeomelanins are present, *d1* manifestation is variable, dependent upon the extent of eumelanin distribution. A horse that is largely pigmented by eumelanins, such as a Dark Bay or Seal Brown *d1/-* horse, will show less pronounced primitive markings than a Blood Bay that is mostly pigmented by phaeomelanins. Many Black horses, which are eumelanistic, show little to no visual signs of primitive markings. However, when croup hairs from such horses that carry a *d1* allele are examined in cross-section, a faint asymmetry can be observed, but macroscopically this effect is masked by the sheer amount of pigment. Spectrophotometric evaluation might be capable of visualizing primitive markings in those individuals where human eyesight fails to discern a difference.

Some *d2/d2* horses can show very faint traces of primitive markings under certain conditions, often depending upon the length of the hairs, the angle and intensity of light, and other similar factors, suggesting that the deletion present in *d2* does not entirely abolish the pigmentation regulation underlying the patterning mediated by *TBX3* expression.

The effect of the *d2* allele is in particular noticeable on a Black background, as can be seen in Figure 1 of Paper III. This striking colouration of the Black *d2/d2* homozygote may possibly be one of the drivers behind the great spread of the *d2* allele. For instance, in the Arabian horse Black is a rather rare colour, and they also have the *d1* allele present at much higher frequency than we see in most European derived breeds. Perhaps, if Black were more common in Arabian horses, *d2* might be more frequent in the breed than it is.

Differences and similarities

As was mentioned before, domestic animals provide a particularly rich variety of polymorphic monogenic traits, where intermediate phenotypes between the extremes are more rare than the extremes. Dun is certainly one such example, but it also provides an interesting look at the concept. The ancient wild horse samples establish that the polymorphic phenotype predates domestication, and thus it is unusual compared to most colour variants in domestic animals, which are thought to have arisen during or after domestication. It is also interesting in the sense that one can observe a continuum of phenotypes from the palest of Duns, through darker Duns, paler non-duns with primitive markings, darker non-duns with primitive markings, and to the darkest of non-duns, where all pigmentation is intense and primitive markings are not discernible. Polygenic enhancers and reducers of melanogenesis can bring Dun and non-dun phenotypes close

enough for considerable overlap in pigmentation intensity, leading to a continuous spectrum with a bimodal distribution.

This overlap means that determining what colour a foal will be when an adult is not always obvious, as the birth coat colour is very often drastically different in appearance from the subsequent coat colour, leading to a marked change in appearance around the first shedding. The exception to this is that Dun foals in general look more like their adult phenotype at birth. The fur on their legs is not yet fully pigmented, but the body, head, mane, and tail are close to what they will be like in the adult. *d1/d1* foals can be born looking quite similar to Dun foals, pale in colour with a prominent dorsal stripe. Therefore they can be hard to distinguish, leading to incorrect colour registration in many cases in breeds where both Dun and non-dun occur. Thus it is in general advisable to reassess the phenotype of a foal after it has shed the birth coat, and again as a yearling.

Predomestic horse colour

The question remains why Dun equids are Dun, and not darker in colour. A common reason for mammalian colouration is crypsis, where the colour serves to help the animal blend in with the environment⁹⁵. If one lends thought to the environment where the ancestors of the domestic horse ranged, the great open grasslands of the Eurasian steppe, Dun is a quite plausible cryptic colouration. A Dun horse can be very similar in colour as yellowed grass and straw, whilst a non-dun is not, as can be seen in Extended Data Figure 1 of Paper III. Why then, after millions of years of Dun equids, did it suddenly become advantageous for some horses to be significantly darker than their ancestors?

One other colour locus in the horse is known to have been polymorphic prior to domestication; the *Leopard Complex* allele at *TRMP1* has been observed in ancient samples, both domestic and predomestic⁹⁶. This discovery was in many ways unexpected, as homozygosity for the *Leopard Complex* is associated with stationary night blindness in horses⁹⁷, a phenotype presumably highly detrimental to a prey species in predator rich pre-domestic environments. A potential environment where patchy pigmentation might be advantageous is in deciduous forests, where interplay between light and shadow is a pronounced environmental factor. Similarly, spotted animals in patchy snow cover can be more difficult to detect than solid coloured animals, even if the solid colour is pale. Heterozygote advantage could be in play, with the less extreme spotting phenotypes of heterozygous *Leopard Complex* individuals with unimpaired vision serving as disruptive camouflage, similar to the spotted phenotypes of many forest dwelling even-toed ungulates⁹⁵.

If one supposes that horses did indeed venture from the open grasslands and into grassy forests, one could imagine that a non-dun1 horse could

potentially have had a selective advantage in a densely forested setting, where darker pigmentation is less conspicuous than a light pelage. Such phenotype distribution has been observed in the Eurasian hamster (*Cricetus cricetus*). There two colour morphs exist, dark and pale, and they have different distributions. The dark form is found nearly exclusively in populations living on a wood-steppe, whereas the pale morph is predominant on the open steppe⁹⁸, the same habitats predomestic horses roamed.

Looking at the cave paintings of horses sporting a *Leopard Complex* like spotting⁹⁹, one does wonder if the darkness of their spots could be a representation of a non-dun horse. It is an intriguing idea whether the *Leopard Complex* co-occurred with predomestic non-dun, whether the bolder colouring of contrasting white with intense pigmentation served to increase the effectiveness of camouflage, conferring a selective advantage in a forested environment to *non-dun1 Leopard Complex* individuals over dilute Duns.

How does this potential scenario for the distribution of colour morphs in predomestic horses then fit together with what knowledge we have of horse domestication? The current view is that a limited number of stallions were domesticated¹⁰⁰, and that their harem herds of domestic mares were supplemented with wild mares of diverse origin¹⁰¹⁻¹⁰⁴, whether through intentional action of humans or the natural behaviour of a fit harem stallion in adding mares to their herd. The origin of horse domestication is thought to be in the western parts of the Eurasian steppe^{105,106}, potentially close to the great forests that bordered the steppe.

Recalling that in investigated domestic horses haplotype diversity is great among *non-dun1* alleles, and that domestic *Dun* alleles show limited haplotype diversity, the assumption would have to be that the largest contribution to domestication came from *non-dun1* horses, with only a few *Dun* horses contributing. The question remains how this meshes with current ideas of how and where the horse was domesticated, and what light research into the historic and predomestic distribution of *Dun* locus alleles can shed on those matters.

Assuming a distribution of *non-dun1* in and around forested areas, and *Dun* on the open grasslands of the steppe, that would indicate that the greatest contribution to domestication was made from forested regions. There is however no concrete evidence for such a distribution of colour morphs amongst predomestic horses, so for the time being, these speculations remain unproven.

Lessons learnt

Allelic evolution

Allelic evolution is a phenomenon geneticists are becoming ever more aware of, as improved sequencing methods enable a greater number of, and more detailed, studies on various animals. Allelic series can not only come about when multiple alleles of different origin are found at one locus, but also when an allele with a variant causing a phenotype subsequently gains a second variant at the same locus that modifies the original phenotype. This process of allelic modification can then repeat itself, with loci diversifying into variably complex allelic series. Humans' interest in the unusual when it comes to their domestic animals can lead to maintenance of multiple variants with differing effects at one locus. This can enable investigations into not only the end product of allelic evolution, but also the genetic configurations that have occurred along the way, thereby aiding the elucidation of the precise effects of the genetic changes, and giving further insights into the biology underlying the phenotypes.

Examples of this have been discovered in a number of domestic species. Alleles that are the product of allelic evolution can for instance be found segregating at loci for white spotting in pigs³⁹ and dogs⁴⁴, colour-sidedness in cattle¹⁰⁷, and dominant-white/smoky plumage colour in chickens¹⁰⁸.

The evolution of alleles can be observed as a component of all three studies discussed in this thesis, albeit in different ways. The Rose-comb study shows how structural rearrangements can lead to considerable phenotypic change, providing raw material for evolutionary selection, and how pleiotropic effects can be an issue when selection, be it natural or artificial, favours certain traits. The rearrangement of the *R1* allele into the *R2* allele is a good illustration of how a non-tandem duplication can occur. It also enables experimental validation of separate causality for both altered comb shape and reduced rooster fertility.

The *Grey* paper touches upon a different kind of allelic evolution; how somatic changes to a structural rearrangement can have varying effects on the fate of both cells and the individual. In Paper II, somatic expansion of the *Grey* duplication is explored, revealing that it can lead to aggressiveness in melanomas, and thereby reduced survival of the individual. Another change, not yet thoroughly explored, is how somatic loss or inactivation of the *Grey* duplication results in reinstatement of normal pigmentation. This effect is commonly seen in the freckling of Flea-bitten Greys, and less commonly in larger patches that are established during embryogenesis. These latter patches are often known as blood marks, from their appearance in Grey Arabians, where the base colour of the horse is generally either Bay or Chestnut, resulting in a largely phaeomelaninic patch.

Another aspect of potential allelic evolution with regards to *Grey* is if the slowness of Greying in the Slow Greying Connemaras is indeed in linkage with *Grey*, then a causative variant could potentially be discovered. The question is, is the slow rate of Greying an acquired feature for the phenotype, or is this rate of Greying the original, with faster Greying having subsequently been selected for?

The evolution of the *Dun* locus has a long history. First the rise of the radially asymmetric microscopic patterning of pigment deposition mediated by *TBX3* at some point in time prior to the radiation of extant Equids, potentially even as far back as the split between odd- and even-toed ungulates. Subsequently the *Dun* locus underwent further diversification, into the *Dun* and *non-dun1* alleles, prior to domestication of the horse. The *non-dun1* allele was then modified further, presumably around the time of domestication. This gave rise to the *non-dun2* allele, resulting in horses of a deeper and darker colour than before, something that seems to have been favoured by humans, if the distribution of the *d2* allele in breeds of European ancestry is anything to go by.

Exploring variation

Rose-comb, Greying with Age and Dun are all phenotypes caused by cis-acting regulatory variants. Such variants, in particular ones such as the *Rose-comb* variant where a novel genomic context of a transcription factor results in ectopic expression, can readily contribute to fast acting evolution. Structural changes of this kind contribute to phenotypic variation in more diverse and unpredictable ways than most coding mutations. Their capacity to give established genes new and additional roles provides raw material for the diversification, evolution, and speciation of natural populations.

Inversions with their implications for recombination rates in heterozygotes can for instance facilitate speciation if an inversion either creates or captures variation that leads to reproductive isolation and/or improved habitat adaptation. An example of this can be seen in the yellow monkeyflower (*Mimulus guttatus*), where an inversion is associated with annuality/perenniality, flowering time and habitat in a manner that seems indicative of ongoing speciation¹⁰⁹.

Exploring regulatory variation poses problems, and it can be much harder to elucidate the effects of regulatory variants than of variants in coding sequences. One problem may be recognising that a variant is potentially functional. If the function that the variant modifies is recent on an evolutionary timescale, the sequences associated with that function may not be conserved in comparison to other species. Such a variant, in particular if it is a SNP or small structural variant, is often likely to go unnoticed, as there may not be anything indicating that it could be of interest. A haplotype

closely related to the variant haplotype is perhaps the best enabler of detection for such scenarios, but often such a haplotype is not available.

A structural variant noticeable enough to be considered a strong candidate for a phenotype because of association with a phenotypic variant in a species may still run up against the same problem, that the elements affected by the variant are not conserved, and thus elucidation of function may be problematic.

One such example is how muscular hypertrophy in some breeds of sheep has been associated with a SNP in an unconserved part of the 3' untranslated region of the gene *Myostatin*. This variant allele creates a binding site for two microRNAs, resulting in translational inhibition of *Myostatin*, and thus increased muscle mass¹¹⁰.

Even if a variant is considered potentially causative due to conservation, there is still the possibility that the function of the conserved element is different in the species under investigation than in species with which it shares conservation. This is in particular likely to happen when an established gene gains a new function.

Sometimes, a variant escapes detection due to experimental limitations, the genome can resist yielding up the causative variant, even if it is known which gene is involved¹¹¹. The reasons for this can for instance involve high GC content, which makes DNA difficult to amplify. Novel PCR-free sequencing methods could hold the key to overcoming these issues of amplification, and enable study of DNA sequences which have up until now been a veritable mystery.

Not only do these drawbacks make it harder to elucidate which variants are likely to be of interest, they can also hinder the use of functional investigations into the mechanisms by which genetic variants have their phenotypic effects, commonly done by experiments on model organisms such as mice, thereby leaving geneticists with fewer tools at their disposal.

Dun is for instance a good example of a phenotype that would be hard to study in the most commonly used mammalian model organism, the mouse. The ancestral diluted phenotype with radial hair asymmetry seen in Duns is not known to occur in rodents, in spite of the pigmentary system of mice being more thoroughly studied than that of any other mammal. Yet the region in which the *non-dun* variants are found shows conservation in mammals. Looking at the base of SNP1, it is highly conserved for the *Dun* allele in mammals, including in mice. This implies that whatever the original function of this element is, down-regulation of pigmentation does not have to be the main function. Thus it is hard to predict what the phenotypic effect of introducing the sequences of the *non-dun* alleles into a mouse would be. Given the lack of asymmetric pigmentation in mice, it is highly unlikely to be similar to that of *non-dun*.

The sequences involved in the different alleles at the *Dun* locus are only a portion of the machinery that has evolved to give equids their pale and

patterned pelage. Some other factor or factors must be responsible for the radial patterning of the growing hair, and *TBX3* expression in the hair follicle is dependent upon that patterning. Without this other unknown element required for a Dun dilution, modelling this system in a mouse is not likely to be fruitful.

These limitations to functional studies are in particular vexing when investigating developmental variants in species for which embryonic experiments are either hard to accomplish, or ethically unacceptable, such as wild animals, larger domestic animals, and humans.

Transient gene expression is a confounding factor of many regulatory variants. Often enough the expression occurs in only a few cells in a localised area, as is exemplified by both *Rose-comb* and *Dun*. This necessitates precise timing of tissue collection, as well as knowledge of the tissues to be sampled. The accessibility of chicken embryos makes developmental variants in the chicken far easier to study functionally than mammalian developmental variants. This has led to extensive knowledge about the embryonic development of chickens, making good timing of sampling relatively feasible. Investigating the Dun dilution could have been fraught with much more difficulty than it proved to be. The regulatory mechanism behind the dilution turned out to be active in each hair-follicular cycle, throughout the lifetime of the horse, something not commonly seen in phenotypes involving melanocyte migration and survival.

Given the aforementioned issues inherent to investigations of regulatory variants, and that the large majority of associations found in genome wide association studies in humans occur in intergenic regions¹¹², something likely to hold true for other species of interest, it is clear that the field of genetics is far from depletion of interesting discoveries, and challenges to be overcome. They also stress the utility of studying the genetics of various species, both domestic and wild, to gain a more comprehensive understanding of the biology of not only our current biosphere, but also the history of life on Earth, and potential genetic innovations yet ahead.

Materials

Large and varied cohorts of phenotyped animals are an asset to studies aiming for phenotype-genotype association. They can be of use both for statistical validation of a hypothesis, and for elucidation of observed data. A large sample set is useful for excluding potentially causative variants, narrowing the interval of interest down as much as possible to what actually is causative. Such a set can also be of aid to explain phenotypic variability that may be found in connection to the locus, and to disentangle complex interactions between loci. For that to be possible, accurate phenotyping of individuals is of prime essence.

Carefully considered record keeping can minimise potential misclassifications, and aid in elucidation of potential allelic series. The importance of record keeping and phenotyping is clearly demonstrated by all three loci discussed herein, *Rose-comb*, *Dun* and *Grey*.

The *Rose-comb* locus has the allelic series of *R1* & *R2* to contend with. Whilst not different in external phenotype, the fact that only *R1* has a marked negative effect on fertility has implications for any potential studies on fertility in Rose-combed chickens. This stresses how individual phenotyping for non-visual traits, with good records and descriptions pertaining to the phenotype, are no less important than for externally observable phenotypes.

The discovery of highly atypical Rose-combs that barely deviate from a Single-comb when given a cursory glance also illustrates how important it is not to make assumptions about perfect penetrance of traits, even if they are well documented in the literature. This is in particular vital when the trait is under examination in a population with non-homogeneous background. It also serves as a reminder of how vital the importance of accurate phenotyping of controls is, something most medical researchers are undoubtedly frequently reminded of in their genetic research.

As the Grey horse changes in colour as it ages, inadequate record keeping can easily lead to misclassified individuals. Grey horses registered as foals may be registered as non-grey in their birth colour. Such misclassifications could confound investigations into the Grey phenotype. Conversely, horses that exhibit slow Greying, as the Connemaras discussed in Paper II do, can be misclassified as non-grey by investigators not familiar with the variability in speed of depigmentation.

For the *Dun* locus, another allelic series of *D*, *d1*, & *d2* was discovered, connected to phenotypic variation that had received little attention, and could easily have been overlooked if it had not been for extensive phenotyping experience, combined with good record keeping and knowledge about the individuals in the sample set. The elucidation of the phenotype-genotype association for these three alleles, coupled with sequences from ancient DNA samples threw light on phenotypic variation in predomestic horses, and may aid in answering where and how the horse was domesticated.

Phenotyping

Accurate phenotyping is not always a straightforward venture; it takes patience, practice, and skill, as well as a keen eye for detail. Defining what precisely it is that marks one rooster's comb as a Rose-comb, and not a Walnut-comb, or why a horse is Bay Dun rather than Bay or Buckskin can be hard to explain. One can draw an analogy to the phenotyping that we all do on a daily basis without even thinking about it. How can we tell a dog apart from a cat? How do we know that the person walking some distance

away in a loose T-shirt and jeans is male or female? Our brains have become so accustomed to classifying according to these categories that we have become blind to what indicators cement our categorisation. We just know. Thus the challenge of the scientific phenotyper is not only to have become so familiar with the categories that they enjoy a high accuracy in classification, but also to dissect their own knowledge and subject it to scrutiny in search for answers.

Rose-comb is far from as uniform as looking at some fancy breed chickens could lead one to think, where Rose-combs are generally neatly shaped, with pearl-like nodules on the top and a straight, sharp spike at the back. Examining the phenotypic variety in the shape and size of Rose-comb in less stringently selected populations reveals a large variety in shapes. Combs that are a riot of large papillae with a curved blunt posterior (Paper I, Fig. S1 - W,X), combs that are large and bulbous with smoothly undulating valleys and hills (Paper I, Fig. S1 - F,Q), tall and thin combs with long spikes (Paper I, Fig. S1 - A), or even combs that are hard to distinguish from a Single-comb (Paper I, Fig. S1, Y,Z).

What is it then that makes a Rose-comb a Rose-comb? Is the genotype alone enough, even though the trait does not penetrate with enough force to shape the comb into the standardised shape found in fancy breeds? It could be argued that as each of the combs depicted deviates from a Single-comb in some way, that is enough to classify them as Rose-combs. Another argument could be made that the more severely deviated combs should be considered separate phenotypes, derived from Rose-comb. Thus the modifiers that give them their unique shapes could be hunted down, further increasing our understanding of the development of chicken combs, whether these modifiers turn out to be genetic or environmental.

Phenotyping a Grey horse is made particularly problematic by virtue of the mutability of the colour. A Grey horse is born with one colour, generally in a darker shade than the average non-grey foal with the same colour. Then this foal begins to change, in some cases it immediately grows paler, but in others it grows darker still before blanching. Then after the horse has become white, some of them begin developing small speckles of pigment again. All of these changes and how differently they can occur in different individuals leads to Grey being confused with many other phenotypes.

A newborn Grey horse could be assumed to be non-grey in whatever colour it happened to be born. A Dappled Grey horse can be taken for a Black Silver, an Iron Grey horse could be confused with a Rabicano, Varnish Roan or a True Roan, a Grey with a blood mark could be thought of as having some spotting phenotype, and a fully depigmented Grey horse could be considered to be a Dominant White horse. The converse is also true, that all of these other coat colours can also erroneously be taken to be Grey. However, each of these have different tells. The darkly pigmented horses; the Black Silvers, Rabicanos, and newborn non-greys do for instance

not become completely white after a few years. A Grey horse does not have the unpigmented skin of a completely white Dominant White horse, even if a Grey horse may show vitiligo to some extent.

Dun phenotyping is no less fraught with pitfalls, as has been discussed at length herein. In particular, the distinction between Dun and non-dun1 is not always clear-cut.

An additional complication to determination of phenotype can be the terminology used. A Grey horse may be a multitude of colours, ranging all the way from pitch black through greys and browns to the purest of whites, yet the horse is always Grey, regardless of what it actually looks like.

When not discussed in terms of horse coat colour, the word ‘dun’ indicates a pale brown colour with a grey cast, a colour close to khaki and beige. The unspecific colour term ‘dun’ could therefore historically indicate any Bay or Chestnut based dilute horse colour. It is only more recently that Dun came to mean a horse with the ancestral equid dilution.

There is also considerable geographic and cultural variability in coat colour terminology. A Black based Dun is in this work called a Blue Dun, but the terms Mouse Dun, Grullo and Grulla are in use as well. Red Dun is also known by terms such as Fox Dun or Claybank Dun. Bay Duns may be called Wild Dun, Zebra Dun, Classic Dun, Yellow Dun, or even just Dun, and that is only scratching the surface of the variation in terminology employed in English. Add other languages to the equation, and figuring out a horse's coat colour from a registry can be less straightforward than one might assume.

Information about the phenotypes of the individuals upon which reference genomes are based is quite important, as the reference animal can carry non wild-type variants, as evidenced by both the *Grey* and *Dun* loci for Twilight, the horse upon which the EquCab assembly is based¹¹³. Good phenotypic information about the individuals upon which reference assemblies are based improves the utility of reference genomes to researchers. The ease with which genomes can now be sequenced could also be employed to develop addenda to reference genomes, annotating genetic variation within a species in a more comprehensive and easily accessible way than is currently done.

Conversations with the public

Search for knowledge, and elucidation of the fundamental workings of life, the universe (and everything), are the hallmarks of science. Knowledge shared adds to our collective comprehension. Knowledge that is not distributed is at best useless, and at worst harmful. Scientists have built up a system for dissemination of examined and acquired knowledge with the scientific community in the publication of peer-reviewed research. Academic publishing is only part of our duty to the society that has fostered scientific

advancement; the more often overlooked and underestimated service we can provide is to engage the public in scientific conversations.

Research on domestic animals is particularly well suited to conversations with the public, as farmers, pet owners, breeders and enthusiasts are often particularly engaged in matters concerning their species or breed of choice. With the amount of time spent on associating with and considering the animals, a great wealth of untested knowledge begging for scientific scrutiny is to be found in such communities.

The knowledge and insights that scientists have gained into the characteristics of domestic animals also need to be transmitted to those involved with the animals in question. The studies presented in this thesis are all good examples of some of the different ways that the knowledge gained by scientific inquiry can be of benefit to animal breeders.

As Rose-comb is a sought after trait by many a chicken-fancier, it could be of interest to chicken breeders to attempt to introduce the *R2* allele into their stocks, to counteract the fertility issues associated with the *R1* allele, and thereby aid their efforts to establish populations that breed true for Rose-comb.

Greying with Age is a well-known phenotype, and much sought after in certain breeds. It does, however, come at a cost, as homozygotes are more likely to develop melanomas, and their melanomas could potentially be more prone to become aggressive due to expansion of the *Grey* duplication. Breeder and owner awareness of this could lead to either altered breeding practices, or increased health monitoring of *Grey* homozygotes, hopefully resulting in improved wellbeing, as well as longer and healthier lifespans of Grey horses.

For certain breeds of horses the Dun colour is prized above other colours, and breeders will at times wish to assume that non-dun horses are Dun, in particular if a horse happens to be homozygous *dl/dl*, exhibiting primitive markings on a well pigmented background. Being able to test for *Dun* zygosity could enable breeders to more accurately plan their breeding, as well as price their horses accordingly when colour is considered an important factor. A homozygous *Dun* horse might for instance be considered a more valuable breeding animal in certain breeds than a heterozygous *Dun* that is in other respects comparable.

Future and reflections

It has been said that every question answered begets further questions, this certainly holds true for science, and is perhaps the *raison d'être* of the scientist. The discoveries discussed on the preceding pages do indeed sow seeds ripe for germination into new questions, some of which have already been discussed in this thesis.

Considering that the chicken comb is a sexual ornament, it is an interesting thought whether human selection is the only factor affecting the propagation of *Rose-comb* in chickens. It is well established that comb size and vibrancy of comb colour are deciding factors when it comes to hens' mate choice, and breeding success of roosters.^{6,114-116} As a *Rose-comb* is frequently more massive than a *Single-comb*, it is conceivable that hens in a promiscuous mating environment have themselves contributed to dissemination of the *Rose-comb* trait. Male behaviour could, however, also affect distribution negatively, as it has been shown that roosters that neither display dominant behaviour nor large and vibrantly coloured combs attract less aggression from dominant males with large and vibrant combs.¹¹⁷

This could be investigated with experiments on mate choice. Hens could for instance be presented with a choice of roosters, matched for fitness factors, but differing in comb phenotype. Comparison of offspring ratios between competitive mating scenarios and heterospermic experiments with *R2/R2* and *Single-combed* roosters could complement both mate choice studies and the studies previously performed on fertility in *Rose-combed* roosters.

The association of *CCDC108* with reduced fertility in homozygous *Rose-combed* roosters can also have value when fertility in other species is considered. The gene had not previously been implicated in fertility, although gene models predicted domains associated with sperm, and flagellar movement. It is conceivable that some instances of sperm motility disorders in humans could be connected to variants of *CCDC108*. The same applies to livestock, where sperm quality is often a very important factor, in particular for species and breeds where artificial insemination is used to a great extent, such as dairy cattle. Further characterisation of *CCDC108* in different species and model organisms, and inclusion as a candidate gene in individuals showing altered sperm motility could lead to insights into the biology of sperm.

Many questions remain unanswered with respects to Greying with Age. The experiments detailed in this thesis demonstrate that the duplication in *STX17* is nearly certainly the causative variant, but it remains to be elucidated precisely how the duplication causes Greying. Experiments on *Grey* melanomas have shown that the duplication is associated with elevated expression of not only *STX17*, but also the neighbouring gene, *NR4A3*⁶⁹.

The precise cause of Greying is not known either. One hypothesis is that the duplication drives melanocytic proliferation, thus explaining the intensely pigmented skin and proneness to development of melanomas seen in *Grey* horses. It could also explain the hair depigmentation if this mechanism winds up depleting the pool of melanocytic stem cells found in the bulge of the hair follicle. The melanocytes responsible for pigmenting the hair during each follicular cycle are recruited from this pool of stem cells²⁹, and if a greater number is recruited each follicular cycle they might eventually run out, leading to depigmented hair. This elevated rate of recruitment would also explain the greater intensity of hair pigmentation commonly seen in *Grey* horses prior to their depigmentation, compared to non-grey horses.

Further studies upon the mechanism of Greying could also potentially lead to treatment options for equine *Grey* melanomas, as well as increase our understanding of the regulation of melanocyte proliferation and recruitment. This in turn may aid research upon human melanomas.

The revelation that predomestic horses were indeed polymorphic at the *Dun* locus opens up a new avenue of investigation into the question of where and how the horse was domesticated. Our current report largely rests upon horses of West-European descent, with East-European and Asian individuals markedly underrepresented. Examining allelic diversity at the *Dun* locus in pre- and postdomestic archaeological remains, as well as in extant horse populations worldwide, and investigating the allelic distribution at the *Dun* locus could shed light on the history of horse domestication, in particular if paired with markers on the Y-chromosome, mitochondria, and other loci of interest, and particular attention is paid to horses from populations surrounding what is thought to have been the epicentre of horse domestication.

Taking a step back, and looking at the underlying theme of structural variation present in the projects discussed herein, begs the question of how and why structural variants arise. One can say that there is often a veritable black box in which the events that lead from one allele to the next take place. If one observes only the end product of a multi-step structural rearrangement it can appear as if something unfathomable and mysterious has happened to result in this variant allele. Elucidating how complex rearrangements take place, such as those reported for Duplex-comb in chickens¹², Colour-sidedness in cattle³² and Lavender in quails¹¹⁸, increases our understanding

of not only the traits associated with the rearrangements, but also the underpinnings of the generation of genetic variation.

Casting light on what happens in the interim between the alleles is something that can better be elucidated the more structural variants are known. Once a structural variant is present, further rearrangements and changes to the region may become more probable, but are there other factors that predispose a region to rearrangement? Taking the *Rose-comb* alleles as an example, the rearrangement events that led to the *R1* and subsequently the *R2* allele both introduced breakpoints, spaced less than 200 bp apart. Is there something about that particular short stretch that inclines it towards lesions?

These new questions will hopefully lead to further studies, and to more comprehensive knowledge and understanding. Essential to successful modern genetic research is the multidisciplinary team or collaboration, encompassing individuals with varied skill sets. State of the art genetic research requires intersecting variable combinations of phenotypic classification, pedigree generation, sample collection, pre- and post-PCR lab work, various functional studies involving cell cultures, protein analysis, and model organism biology, as well as bioinformatics of differing complexity levels, for instance dealing with the challenges of large quantities of data generated in the lab. No one researcher is capable of expertise in all of these areas, but a well balanced team with complementary skill sets can produce research of quality far surpassing the sum of what the individual team members could produce on their own.

Tables

Table 1. *Loci with variants known to affect pigmentation in the domestic horse.*

Gene	Phenotypes	Effect of variant on hair
<i>MC1R</i>	Black/Chestnut ⁵³	Phaeomelanism
<i>ASIP</i>	Bay/Black ⁵⁷	Eumelanism
<i>TBX3</i>	Dun/non-dun ^a	Increased pigmentation
<i>SLC45A2/MATP</i>	Cream ⁵⁸	Dilution of phaeomelanins
<i>PMEL</i>	Silver ⁶⁰	Dilution of eumelanins
<i>SLC36A1</i>	Champagne ⁶²	Dilution of eu- and phaeomelanins
<i>STX17 & NR4A3</i>	Grey ⁶⁹	Gradual hair depigmentation
<i>KIT</i>	Tobiano ⁶⁷ , Sabino ⁵⁹ , Roan ¹¹⁹ , Dominant White ^{61,63,66,120} , White Markings ^{65,121}	Spotting, and other white hair phenotypes
<i>MITF</i>	Splashed White ⁶⁴ , Macchiato ⁶⁴ , White Markings ¹²¹	Spotting/White markings
<i>PAX3</i>	Splashed White ^{64,65} , White Markings ⁶⁵	Spotting/White markings
<i>EDNRB</i>	Frame Overo ⁵⁴⁻⁵⁶	Spotting/White markings
<i>TRPM1</i>	Leopard Spotting Complex ⁶⁸	Spotting/White hairs

^aPresent study

Table 2. *Allelic notation and effects on hair pigmentation for some colour loci of the domestic horse.*

Phenotypes	Basic allele notation	Dominance and effects on hair pigmentation
Black/Chestnut	<i>E e</i>	Dominant <i>E</i> permits eumelanin production
Bay/Black	<i>A a</i>	Dominant <i>A</i> spatially restricts eumelanin production
Dun/non-dun	<i>D d</i>	Dominant <i>D</i> allows patterned dilution
Cream	<i>C c</i>	Incompletely dominant <i>C</i> dilutes phaeomelanins in heterozygous state, and both eu- and phaeomelanins to a greater degree in homozygous state
Silver	<i>Z z</i>	Dominant <i>Z</i> dilutes eumelanins
Grey	<i>G g</i>	Dominant <i>G</i> leads to hair depigmentation with age
Tobiano	<i>To to</i>	Dominant <i>To</i> gives rise to spotting, with no effect on the colour of remaining pigment

Table 3. *English names for some equid colour variants. Note that this table is not exhaustive, and that other terms may be in use as well as these.*

Colour	Basic allele notation	Phenotype composition
Chestnut	<i>ee cc dd</i>	Phaeomelanins + non-dun
Bay	<i>E- A- cc dd zz</i>	Eu- and phaeomelanins + non-dun
Black	<i>E- aa -c dd zz</i>	Eumelanins + non-dun
Red Dun	<i>ee cc D-</i>	Chestnut + Dun
Bay Dun	<i>E- A- cc D- zz</i>	Bay + Dun
Blue Dun	<i>E- aa -c D- zz</i>	Black + Dun
Palomino	<i>ee Cc dd</i>	Chestnut + Cream + non-dun
Buckskin	<i>E- A- Cc dd zz</i>	Bay + Cream + non-dun
Double Cream	<i>CC</i>	Any colour + homoz. Cream
Cremello	<i>ee CC</i>	Chestnut + homoz. Cream
Perlino	<i>E- A- CC</i>	Bay + homoz. Cream
Smoky Cream	<i>E- aa CC</i>	Black + homoz. Cream
Red Dun Cream	<i>ee Cc D-</i>	Chestnut + Cream + Dun
Bay Dun Cream	<i>E- A- Cc D- zz</i>	Bay + Cream + Dun
Bay Silver	<i>E- A- cc Z-</i>	Bay + Silver + non-dun
Black Silver	<i>E- aa -c Z-</i>	Black + Silver + non-dun
Bay Dun Silver	<i>E- A- cc D- Z-</i>	Bay + Dun + Silver
Blue Dun Silver	<i>E- aa -c D- Z-</i>	Black + Dun + Silver
Bay Dun Cream Silver	<i>E- A- Cc D- Z-</i>	Bay + Dun + Cream + Silver
Grey	<i>G-</i>	Any colour + Grey
Tobiano	<i>To-</i>	Any colour + Tobiano

Table 4. *English, Icelandic, Swedish and German names for some colour variants in the domestic horse.*

English	Icelandic/íslenska	Swedish/svenska	German/Deutsch
Chestnut	Rautt	Fux	Fuchs
Bay	Jarpt	Brun	Braun
Black	Brúnt	Svart	Rappe
Red Dun	Bleikt	Rödblack	Fuchsfalbe
Bay Dun	Bleikálótt	Brunblack	Braunfalbe
Blue Dun	Móalótt	Musblack	Mausfalbe
Palomino	Leirljótt	Isabell	Isabell
Buckskin	Moldótt	Gulbrun	Erdfarbe
Double Cream/Albino	Fölt	Dubbelgul/Albino	Bleich/Albino
Cremello	Rauðfölt	Gulvit	Weissisabell/Cremello
Perlino	Jarpfölt	Pärlvit	Perlino
Smoky Cream	Brúnfölt	Rökvit	Smokey Cream
Dunalino	Bleikt leirljótt	Isabellblack	Isabellfalbe
Dunskin	Bleikálótt moldótt	Gulbrunblack/Vitblack	Erdfarbfalbe
Bay Silver	Jarpvindótt	Silverbrun	Braunwindfarbe
Black Silver	Móvindótt	Silversvart	Rappwindfarbe
Bay Dun Silver	Bleikálótt vindótt	Silverbrunblack	Braunwindfarbfalbe
Blue Dun Silver	Móalótt vindótt	Silvermusblack	Rappwindfarbfalbe
Dunskin Silver	Bleikálótt moldótt vindótt	Silvertitblack	Erdwindfarbfalbe
Grey	Grátt	(Avblekbar) Skimmel	(Echter) Schimmel
Tobiano	Skjótt	Skäck	Schecke

Acknowledgements

Now it comes to the only part most people will actually read. Perhaps those who read the entire thesis of their own volition ought to question their sanity. Or mine, your choice.

Leif, I am very grateful to have had you as a supervisor. When I was first considering contacting you I was told by an erstwhile colleague of yours that he thought it was a good idea. Not just because you are an outstanding scientist, but because you are a good human being, and that one can never be sure to find those two traits together in a person. I can safely say that he was right, and that your humanity and kind nature have been invaluable during this overly long and both physically and mentally taxing journey. You have of course also been scientifically excellent, and encouraged me to grow as a researcher, but other and greater people than I have attested to that ability of yours and will continue to do so. I feel it is my role to point out to you how important it is that you have a capacity for kindness, patience, and mercy, beyond what many other highly accomplished scientists do.

Gabriella, I think that your outstanding qualities are positivity and enthusiasm. You always bring along a fast paced and eager spirit. At the risk of sounding a bit odd, I have sometimes thought to myself that if you were not a human but rather one of the horses we have studied, you would be one of the lithe hot-blooded horses, a spirited Arabian, or a fast Thoroughbred. My generally even-keeled nature does not lend itself to the kind of fiery avidity you possess, and it has been a pleasure to partake in yours!

Kerstin, you are remarkable, as I am sure you have oft been told. There is a side to you that I wish more people had the opportunity to see, the alternately mischievous and serious girl, not just the extraordinarily clever and talented scientist. I remember having sat with you on the floor of your kitchen in Flogsta, talking about everything and nothing, a fond memory of a very inspiring woman. Please take care of yourself.

I have had the great luck to enjoy the company of excellent **colleagues** and **collaborators**, whom I would like to thank, **one and all**. Not only have you been splendid people, but you have also been very many. According to tradition, my memory at this stage is extremely suspect, so I would like to offer apologies in advance if I've left you out. Please refrain from haunting me like a ghost out of an Icelandic folk-tale. It would be most disturbing.

I have been fortunate to enjoy some very good collaborations and scientific contacts, among them our chicken colleagues at INRA, **Bertrand** and **Michelle, Chungang** during his time at CAU, as well as Virginia Tech's **Paul Siegel** with his penchant for entertainment.

Working with **Greg, Kelly** and their colleagues on the Dun story has been a privilege. So many others have contributed to that project, and helped me tell this story I have held so dear, among them **Cecilia Penedo, Claire Wade, Ludovic Orlando** and last but not least **Terje Raudsepp** and **Ingrid Randlaht** who provided the Estonian puzzle piece.

The good folks at Neuro, **Finn, Henrik B, Henrik R,** and **Shahrzad** have also been extremely helpful collaborators, and very generous with their time.

Talking about horses with my **colleagues at SLU**, as well as **Þorvaldur Árnason**, has been very engaging and I hope to be able to continue to do so in the future.

Elizabeth, Chris and **Ben**, I am profoundly grateful that you guys kept me afloat during your time here. Thank you for all the conversations, both the clever ones and the less than appropriate ones. I will be looking forward to visits from you all in the future.

Doreen, Fabiana, Jonas and **Marta**, you have listened to the silly things that I allow out of my mouth, and pretended they were actually semi-sensible, or joined me for a good and hearty laugh. You have also listened to my worries, and encouraged me with your friendship. You are all wonderful friends, and I feel fortunate to know you. We have had some very good times with **Agnese, Fan, Chao, Axel, Nima, Sangeet, Iris** and **all the others**. Played games, had delicious food, been ridiculous, been serious, and last but not least, had some great laughs!

Calle and **Marcin**, your office is always a good place to visit. Banter and science blend for a very pleasing result. It may be as innocent as a rabbit, but it may also be as serious as a rat infestation. It has been fantastic to be able to work with you Calle, I'm quite certain you will make a name for yourself if you put your mind to it.

Mike Zody also made a lot of bioinformatic magic happen on my behalf. Your incredible brain impresses me every time I communicate with you, not the least because of your sense of humour. That time you and **Michael Strömberg** giggled at my dinner table is always particularly memorable.

I have shared an office with many people in the lab, among them **Ulrika, Sus, Elisabeth, Gerli, Anders L, Sergey, Ulrika** (again?) and, for the longest time, **Ulla**. It should not be underestimated how important agreeable office mates are. We have had good conversations, and clever people willing to help each other with problems!

Ulla behöver faktiskt pratas om lite mer, eftersom hon är helt enastående i sin förmåga att veta allt som kan tänkas att någon behöver veta om labbet, eller hur man borde utföra sitt labbande. **Guðrun Wieslander, Barbro Lowisin, och Cecilia Johansson** var också starka klippor för oss meniga. Nu ser vår fina **Eva, Åsa** och **Jessica** till att allt fungerar, trots att vi andra försöker störa saker hela tiden. Kämpa på, utan er skulle allt gå åt skogen!

Yang and Sharda, you share my passion for food, and have both allowed me a glimpse into your delicious cuisines. Something I treasure, and I promise to keep paying it forward.

Lin, you were a wonderful flatmate to have, and a good friend. I hope that some day you will be able to come to see Iceland. I promise to have woollen socks and a pullover to keep you warm!

Whilst I always think everyone would have their day improved by riding a good Icelandic horse, three people in particular are very welcome to visit once my horses have been educated to acceptable levels. **Anna O, Doreen** and **Lisa**, I would very much like to enjoy Iceland from horseback with you.

I have also had the pleasure to guide people, both in their own projects, and on how they could maybe help me with mine. **Ranran, Karin, and Sara** have all worked with me, and helped me develop the skills needed for close collaborations.

Other members of the lab, past and present, you have also spiced life with your presence, all the good folks we have the opportunity for daily lunches with. The people whose company we had the pleasure of before, but have now flown to other locations aren't be forgotten either.

Saga, Dana, Erlendur, Höskuldur, Steinar, Snævar og allir hinir Íslendingarnir sem hafa dvalið með mér hér í Svíaríki, takk fyrir vináttuna, matinn, og tækifærin til að láta gamminn geysa á hinu ástkæra ylhýra!

Allir mínir gömlu og góðu hjartans vinir heima á Fróni (sem og fjarri Íslands ströndum), ykkur grunar kannski að ég sé enn til í henni veröld þrátt fyrir langa fjarveru. Þið megið telja þessa ritgerð sem sönnun þess að ég lifi enn og hrærist, þó kannski ekki nema í eigin hugarfylgsnum. Vinátta ykkar hefur verið mér ómetanleg stoð.

Ártúnahjúin, Ásgarðsbúar, Lágfellingar, Miðengjamenn, Röðulshjón, Sölvhyltingar, Tryggvi í Bár, og allir aðrir bændur og hestamenn sem ekki eru nefndir á nafn, en hafa lagt mér lið, spjallað margt og mikið, sýnt einstaka gestrisni, og leyft mér að hárreyta og trufla hross (og hænsn) í þágu vísindanna. Þessi ritsmíð væri með mjög ólíku sniði ef ykkar hefði ekki notið við, og kann ég ykkur öllum bestu þakkir fyrir!

Ég er alveg hreint einstaklega lánssöm að eiga þá fjölskyldu að sem ég á. Þið hafið stutt mig í því að fara þá leið sem mér hefur hugnast allt frá blautu barnsbeini, og ekki möglað þó ég hafi horfið af landi brott.

Það eru ekki allir svo vel í sveit settir að eiga foreldri sem hefur haft jafn mikið með doktorsnám afkvæmis að gera og faðir minn, **Páll**, hefur gert. Hvað þá að foreldri leggi jafn mikinn tíma og vinnu í það án þess að hafa eitthvað sérstakt á því að græða annað en að slaka þorsta eftir vitneskju og ástríðu fyrir vísindum. Hugsjónastarf föður míns varðandi þekkingarmiðlun til samferðamanna þykir mér aðdáunarvert. Ég hef fengið anga af þessarri áráttu frá honum, og vonandi endurspeglast vottur af henni í þessarri ritsmíð, þó fræðileg sé.

Þá er ég einnig sérlega heppin með stóru systur mína, **Birnu** hina fjöltyngdu og afbragðs málfæru. Áreiðanlegar og vel studdar áminningar þegar maður hrasar um sjálfan sig í tjáningu hafa óneitanlega byggt upp sjálfsgagnrýni, og vanið mann á að hugsa sig kannski eilítið um áður en maður gloprar einhverju út úr sér. Án þíns sérdeilis fagmannlega prófarkalesturs hefði ritgerðin einnig verið stirðbusalegri og ókaraðri, þó ég hafi ekki haft tóm til að fara jafn djúpt í umbætur og best hefði verið. (Við látum það liggja milli hluta að þessi partur hennar hefði örugglega líka haft gott af prófarkalestri.)

Móðir mín, **Ragnheiður**, hefur ekki síður lagt mér góðan kost í skrínu á þessarri vegferð. Alltaf hefur hún haft trú á mér, og hvatt mig til dáða. Hjálpað mér að standa gegn straumnum og halda áfram þó á móti hafi blásið. Að því ógleymdu að hafa kennt mér snemma á tölvur og teikniforrit, nokkuð sem hefur ærið oft verið til mikils gagn við þessa vinnu.

At the very end I feel compelled to provide a translation of the quotation that preceded the main body of this work, the words of the Icelandic poet Einar Benediktsson. His wonderfully beautiful and descriptive poem about the Icelandic horse, *Fákar*, manages to capture the freedom, joy and fiery spirit a horse can bestow upon its rider. However, the brilliancy of the poem is not limited to that, for he also describes something utterly fundamental to the nature of human existence, and particularly applicable to the modern scientist.

A man alone is but half a man, with others he becomes more than merely himself.

- Einar Benediktsson

References

1. Lander, E.S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).
2. Bailey, J.A. *et al.* Recent segmental duplications in the human genome. *Science* **297**, 1003-7 (2002).
3. Mackie Ogilvie, C. & Scriven, P.N. Meiotic outcomes in reciprocal translocation carriers ascertained in 3-day human embryos. *Eur J Hum Genet* **10**, 801-6 (2002).
4. Kirkpatrick, M. How and why chromosome inversions evolve. *PLoS biology* **8**, 2040 (2010).
5. Eriksson, J. *et al.* Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. *PLoS Genet* **4**, e1000010 (2008).
6. Parker, T. & Ligon, J. Female mating preferences in red junglefowl: a meta-analysis. *Ethology Ecology & Evolution* **15**, 63-72 (2003).
7. Peterson, A.T. & Brisbin, I.L. Genetic endangerment of wild Red Junglefowl *Gallus gallus*? *Bird Conservation International* **8**, 387-394 (1998).
8. Bateson, W. Experiments with poultry. *Rep Evol Comm Roy Soc* **1**, 87-124 (1902).
9. Bateson, W. & Punnett, R. A suggestion as to the nature of the "Walnut" Comb in the fowls. *Proc. Cambridge Philosophical Soc.* **13**, 165-168 (1905).
10. Bateson, W. & Punnett, R. Experimental studies in the physiology of heredity. *Classic Papers in Genetics*. JA Peters, ed. Prentice-Hall, Englewood Cliffs, NJ, 42-59 (1908).
11. Wright, D. *et al.* Copy number variation in intron 1 of SOX5 causes the Pea-comb phenotype in chickens. *PLoS Genet* **5**, e1000512 (2009).
12. Dorshorst, B. *et al.* A genomic duplication is associated with ectopic eomesodermin expression in the embryonic chicken comb and two duplex-comb phenotypes. *PLoS Genet* **11**, e1004947 (2015).
13. Hindhaugh, W. Some observations on fertility in White Wyandottes. *Poultry J* **17**, 555-560 (1932).
14. Hutt, F. A relation between breed characteristics and poor reproduction in White Wyandotte fowls. *American Naturalist*, 148-156 (1940).
15. Crawford, R. & Merritt, E. The relationship between Rose Comb and reproduction in the domestic fowl. *Can. J. Genet. Cytol.* **5**, 89-95 (1963).
16. Crawford, R. & Smyth, J.R. Studies of the Relationship between fertility and the gene for Rose Comb in the domestic fowl - 1. The relationship between comb genotype and fertility. *Poultry Science* **43**, 1009-1017 (1964).
17. Crawford, R. & Smyth, J.R. Studies of the Relationship between fertility and the gene for Rose Comb in the domestic fowl - 2. The relationship between comb genotype and duration of fertility. *Poultry Science* **43**, 1018-1026 (1964).
18. Crawford, R. & Smyth, J.R. Semen quality and the gene for Rose Comb in the domestic fowl. *Poultry Science* **43**, 1551-1557 (1964).

19. Crawford, R. Comb dimorphism in Wyandotte domestic fowl. I. Sperm competition in relation to rose and single comb alleles. *Canadian Journal of Genetics and Cytology* **7**, 500-504 (1965).
20. Buckland, R.B. & Hawes, R.O. Comb type and reproduction in the male fowl segregation of the rose and pea comb genes. *Can J Genet Cytol* **10**, 395-400 (1968).
21. Crawford, R. Rose comb and fertility in Silver Spangled Hamburgs. *Poultry science* **50**, 867-869 (1971).
22. Petitjean, M. & Servouse, M. Comparative-Study of Some Characteristics of the Semen of Rr (Rose Comb) or rr (Single Comb) Cockerels. *Reproduction Nutrition Development* **21**, 1085-1093 (1981).
23. Kirby, J.D., Engel, H.N., Jr. & Froman, D.P. Analysis of subfertility associated with homozygosity of the rose comb allele in the male domestic fowl. *Poult Sci* **73**, 871-8 (1994).
24. McLean, D.J. & Froman, D.P. Identification of a sperm cell attribute responsible for subfertility of roosters homozygous for the rose comb allele. *Biol Reprod* **54**, 168-72 (1996).
25. McLean, D.J., Jones, L.G., Jr. & Froman, D.P. Reduced glucose transport in sperm from roosters (*Gallus domesticus*) with heritable subfertility. *Biol Reprod* **57**, 791-5 (1997).
26. Yu, M. *et al.* The developmental biology of feather follicles. *The International journal of developmental biology* **48**, 181 (2004).
27. Riley, P. Melanin. *The international journal of biochemistry & cell biology* **29**, 1235-1239 (1997).
28. Rawles, M.E. Origin of pigment cells from the neural crest in the mouse embryo. *Physiological zoology*, 248-266 (1947).
29. Lin, J.Y. & Fisher, D.E. Melanocyte biology and skin pigmentation. *Nature* **445**, 843-50 (2007).
30. Hayes, B.J., Pryce, J., Chamberlain, A.J., Bowman, P.J. & Goddard, M.E. Genetic architecture of complex traits and accuracy of genomic prediction: coat colour, milk-fat percentage, and type in Holstein cattle as contrasting model traits. *PLoS Genet* **6**, e1001139 (2010).
31. Philipp, U. *et al.* A MITF mutation associated with a dominant white phenotype and bilateral deafness in German Fleckvieh cattle. *PLoS One* **6**, e28857 (2011).
32. Durkin, K. *et al.* Serial translocation by means of circular intermediates underlies colour sidedness in cattle. *Nature* **482**, 81-4 (2012).
33. Fontanesi, L., Scotti, E. & Russo, V. Haplotype variability in the bovine MITF gene and association with piebaldism in Holstein and Simmental cattle breeds. *Animal genetics* **43**, 250-256 (2012).
34. Johansson Moller, M. *et al.* Pigs with the dominant white coat color phenotype carry a duplication of the KIT gene encoding the mast/stem cell growth factor receptor. *Mamm Genome* **7**, 822-30 (1996).
35. Giuffra, E. *et al.* The Belt mutation in pigs is an allele at the Dominant white (I/KIT) locus. *Mammalian Genome* **10**, 1132-1136 (1999).
36. Pielberg, G., Olsson, C., Syvänen, A.-C. & Andersson, L. Unexpectedly high allelic diversity at the KIT locus causing dominant white color in the domestic pig. *Genetics* **160**, 305-311 (2002).
37. Giuffra, E. *et al.* A large duplication associated with dominant white color in pigs originated by homologous recombination between LINE elements flanking KIT. *Mammalian Genome* **13**, 569-577 (2002).

38. Lim, H. *et al.* Novel alternative splicing by exon skipping in KIT associated with whole-body roan in an intercrossed population of Landrace and Korean Native pigs. *Animal genetics* **42**, 451-455 (2011).
39. Rubin, C.-J. *et al.* Strong signatures of selection in the domestic pig genome. *Proceedings of the National Academy of Sciences* **109**, 19529-19536 (2012).
40. Haase, B., Rieder, S. & Leeb, T. Two variants in the KIT gene as candidate causative mutations for a dominant white and a white spotting phenotype in the donkey. *Anim Genet* **46**, 321-4 (2015).
41. Karlsson, E.K. *et al.* Efficient mapping of mendelian traits in dogs through genome-wide association. *Nature genetics* **39**, 1321-1328 (2007).
42. Wong, A. *et al.* A de novo mutation in KIT causes white spotting in a subpopulation of German Shepherd dogs. *Animal genetics* **44**, 305-310 (2013).
43. Gerding, W., Akkad, D. & Epplen, J. Spotted Weimaraner dog due to de novo KIT mutation. *Animal genetics* **44**, 605-606 (2013).
44. Körberg, I.B. *et al.* A simple repeat polymorphism in the MITF-M promoter is a key regulator of white spotting in dogs. (2014).
45. Lyons, L.A. Feline genetics: clinical applications and genetic testing. *Topics in companion animal medicine* **25**, 203-212 (2010).
46. David, V.A. *et al.* Endogenous retrovirus insertion in the KIT oncogene determines white and white spotting in domestic cats. *G3 (Bethesda)* **4**, 1881-91 (2014).
47. Steingrimsón, E., Copeland, N.G. & Jenkins, N.A. Mouse coat color mutations: from fancy mice to functional genomics. *Developmental dynamics* **235**, 2401-2411 (2006).
48. Mountjoy, K.G., Robbins, L.S., Mortrud, M.T. & Cone, R.D. The cloning of a family of genes that encode the melanocortin receptors. *Science* **257**, 1248-1251 (1992).
49. Bultman, S.J., Michaud, E.J. & Woychik, R.P. Molecular characterization of the mouse agouti locus. *Cell* **71**, 1195-1204 (1992).
50. Miller, M. *et al.* Cloning of the mouse agouti gene predicts a secreted protein ubiquitously expressed in mice carrying the lethal yellow mutation. *Genes & development* **7**, 454-467 (1993).
51. Geschwind, I., Huseby, R. & Nishioka, R. The effect of melanocyte-stimulating hormone on coat color in the mouse. *Recent progress in hormone research* **28**, 91-130 (1971).
52. Cone, R.D. *et al.* The melanocortin receptors: agonists, antagonists, and the hormonal control of pigmentation. *Recent progress in hormone research* **51**, 287-317; discussion 318 (1995).
53. Marklund, L., Moller, M.J., Sandberg, K. & Andersson, L. A missense mutation in the gene for melanocyte-stimulating hormone receptor (*MC1R*) is associated with the chestnut coat color in horses. *Mammalian Genome* **7**, 895-899 (1996).
54. Metallinos, D.L., Bowling, A.T. & Rine, J. A missense mutation in the endothelin-B receptor gene is associated with Lethal White Foal Syndrome: an equine version of Hirschspung Disease. *Mammalian Genome* **9**, 426-431 (1998).
55. Santschi, E.M. *et al.* Endothelin receptor B polymorphism associated with lethal white foal. *Mammalian Genome* **9**, 306-309 (1998).

56. Yang, G.C. *et al.* A Dinucleotide mutation in the endothelin-B receptor gene is associated with lethal white foal syndrome (LWFS); a horse variant of Hirschsprung Disease (HSCR). *Human Molecular Genetics* **7**, 1047-1052 (1998).
57. Rieder, S., Taourit, S., Mariat, D., Langlois, B. & Guérin, G. Mutations in the agouti (*ASIP*), the extension (*MC1R*), and the brown (*TYRP1*) loci and their association to coat color phenotypes in horse (*Equus caballus*). *Mammalian Genome* **12**, 450-455 (2001).
58. Mariat, D., Taourit, S. & Guérin, G. A mutation in the *MATP* gene causes the cream coat colour in the horse. *Genetics Selection Evolution* **35**, 119-133 (2003).
59. Brooks, S.A. & Bailey, E. Exon skipping in the *KIT* gene causes a Sabino spotting pattern in horses. *Mammalian Genome* **16**, 893-902 (2005).
60. Brunberg, E. *et al.* A missense mutation in *PMEL17* is associated with the Silver coat color in the horse. *BMC genetics* **7**(2006).
61. Haase, B. *et al.* Allelic Heterogeneity at the Equine *KIT* Locus in Dominant White (W) Horses. *PLoS Genetics* **3**, 1-8 (2007).
62. Cook, D., Brooks, S., Bellone, R. & Bailey, E. Missense mutation in exon 2 of *SLC36A1* responsible for champagne dilution in horses. *PLoS Genet* **4**, e1000195 (2008).
63. Haase, B. *et al.* Seven novel *KIT* mutations in horses with white coat colour phenotypes. *Anim Genet* **40**, 623-9 (2009).
64. Hauswirth, R. *et al.* Mutations in *MITF* and *PAX3* cause "splashed white" and other white spotting phenotypes in horses. *PLoS Genet* **8**, e1002653 (2012).
65. Hauswirth, R. *et al.* Novel variants in the *KIT* and *PAX3* genes in horses with white-spotted coat colour phenotypes. *Anim Genet* **44**, 763-5 (2013).
66. Haase, B., Jagannathan, V., Rieder, S. & Leeb, T. A novel *KIT* variant in an Icelandic horse with white-spotted coat colour. *Anim Genet* (2015).
67. Brooks, S.A., Lear, T.L., Adelson, D.L. & Bailey, E. A chromosome inversion near the *KIT* gene and the Tobiano spotting pattern in horses. *Cytogenetic and Genome Research* **119**, 225-230 (2007).
68. Bellone, R.R. *et al.* Evidence for a retroviral insertion in *TRPM1* as the cause of congenital stationary night blindness and leopard complex spotting in the horse. *PLoS One* **8**, e78280 (2013).
69. Rosengren Pielberg, G. *et al.* A cis-acting regulatory mutation causes premature hair graying and susceptibility to melanoma in the horse. *Nat Genet* **40**, 1004-9 (2008).
70. Fleury, C. *et al.* The study of cutaneous melanomas in Camargue-type gray-skinned horses (2): epidemiological survey. *Pigment Cell Res* **13**, 47-51 (2000).
71. Hadwen, S. The Melanomata of Grey and White Horses. *Can Med Assoc J* **25**, 519-30 (1931).
72. Lerner, A.B. & Cage, G.W. Melanomas in horses. *Yale J Biol Med* **46**, 646-9 (1973).
73. Fleury, C., Berard, F., Balme, B. & Thomas, L. The study of cutaneous melanomas in Camargue-type gray-skinned horses (1): clinical-pathological characterization. *Pigment Cell Res* **13**, 39-46 (2000).
74. Valentine, B.A. Equine melanocytic tumors: a retrospective study of 53 horses (1988 to 1991). *J Vet Intern Med* **9**, 291-7 (1995).
75. Wehrhahn, C. & Crawford, R. Comb dimorphism in Wyandotte domestic fowl. 2. Population genetics of the rose comb gene. *Canadian Journal of Genetics and Cytology* **7**, 651-657 (1965).

76. William, C.M., Tanabe, Y. & Jessell, T.M. Regulation of motor neuron subtype identity by repressor activity of Mnx class homeodomain proteins. *Development* **130**, 1523-36 (2003).
77. Meszar, Z. *et al.* Hyaluronan accumulates around differentiating neurons in spinal cord of chicken embryos. *Brain Res Bull* **75**, 414-8 (2008).
78. Shima, J.E., McLean, D.J., McCarrey, J.R. & Griswold, M.D. The murine testicular transcriptome: characterizing gene expression in the testis during the progression of spermatogenesis. *Biol Reprod* **71**, 319-30 (2004).
79. Pazour, G.J., Agrin, N., Leszyk, J. & Witman, G.B. Proteomic analysis of a eukaryotic cilium. *J Cell Biol* **170**, 103-13 (2005).
80. Bamshad, M. *et al.* Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome. *Nat Genet* **16**, 311-5 (1997).
81. Frank, D.U., Emechebe, U., Thomas, K.R. & Moon, A.M. Mouse TBX3 mutants suggest novel molecular mechanisms for Ulnar-mammary syndrome. *PLoS One* **8**, e67841 (2013).
82. Kumar, P.P. *et al.* TBX3 regulates splicing in vivo: a novel molecular mechanism for Ulnar-mammary syndrome. *PLoS Genet* **10**, e1004247 (2014).
83. Kunisada, T. *et al.* Transgene expression of steel factor in the basal layer of epidermis promotes survival, proliferation, differentiation and migration of melanocyte precursors. *Development* **125**, 2915-23 (1998).
84. Yoshida, H. *et al.* Review: melanocyte migration and survival controlled by SCF/c-kit expression. *J Invest Dermatol Symp Proc* **6**, 1-5 (2001).
85. Isaac, A. *et al.* Tbx genes and limb identity in chick embryo development. *Development* **125**, 1867-75 (1998).
86. Suzuki, T., Takeuchi, J., Koshiha-Takeuchi, K. & Ogura, T. Tbx Genes Specify Posterior Digit Identity through Shh and BMP Signaling. *Dev Cell* **6**, 43-53 (2004).
87. Marsh, O.C. Polydactyle horses, recent and extinct. *American Journal of Science*, 499-505 (1879).
88. Marsh, O.C. Recent polydactyle horses. *American Journal of Science*, 339-355 (1892).
89. Wriedt, C. Nedarvingen av kjakeflek med sort haar hos vestlandshester. *Norsk Veterinær-Tidsskrift*, 316-318 (1918).
90. Stachurska, A. *et al.* Colour variation in blue dun Polish Konik and Biłgoraj horses. *Livestock Production Science* **90**, 201-209 (2004).
91. Tuff, P. Genetiske undersøkelser over hestefarver. *Nordiska Veterinärmötet - Kongressberättelse (Helsingfors)*, 689-716 (1933).
92. Loen, J. Farge-Nedarvinga hos Vestlandshesten (Fjordhesten). in *Stambok over Vestlandshesten (Fjordhesten) - 11. Band - Hingstar fødd til og med året 1934*, Vol. 11 1-52 (Statens stambokkontor for hester, Oslo, 1939).
93. Aðalsteinsson, S. Inheritance of yellow dun and blue dun in the Icelandic toelter horse. *Journal of Heredity* **69**, 146-148 (1978).
94. Craig, L. & Vleck, L.D.v. Evidence for inheritance of the red dun dilution in the horse. *Journal of Heredity* **76**, 138-139 (1985).
95. Caro, T. The adaptive significance of coloration in mammals. *Bioscience* **55**, 125-136 (2005).
96. Ludwig, A. *et al.* Twenty-five thousand years of fluctuating selection on leopard complex spotting and congenital night blindness in horses. *Philos Trans R Soc Lond B Biol Sci* **370**, 20130386 (2015).

97. Sandmeyer, L.S., Breaux, C.B., Archer, S. & Grahn, B.H. Clinical and electroretinographic characteristics of congenital stationary night blindness in the Appaloosa and the association with the leopard complex. *Veterinary Ophthalmology* **10**, 368-375 (2007).
98. Huxley, J. Evolution. The Modern Synthesis. *Evolution. The Modern Synthesis*. (1942).
99. Pruvost, M. *et al.* Genotypes of predomestic horses match phenotypes painted in Paleolithic works of cave art. *Proceedings of the National Academy of Sciences* **108**, 18626-18630 (2011).
100. Lindgren, G. *et al.* Limited number of patriline in horse domestication. *Nature genetics* **36**, 335-336 (2004).
101. Lister, A., Kadwell, M., Kaagen, L., Richards, M.B. & Stanley, H. Ancient and modern DNA in a study of horse domestication. *Ancient Biomolecules* **2**, 267-280 (1998).
102. Jansen, T. *et al.* Mitochondrial DNA and the origins of the domestic horse. *Proceedings of the National Academy of Sciences* **99**, 10905-10910 (2002).
103. Vilà, C. *et al.* Widespread origins of domestic horse lineages. *Science* **291**, 474-477 (2001).
104. Achilli, A. *et al.* Mitochondrial genomes from modern horses reveal the major haplogroups that underwent domestication. *Proceedings of the National Academy of Sciences* **109**, 2449-2454 (2012).
105. Warmuth, V. *et al.* Reconstructing the origin and spread of horse domestication in the Eurasian steppe. *Proceedings of the National Academy of Sciences* **109**, 8202-8206 (2012).
106. Outram, A.K. *et al.* The earliest horse harnessing and milking. *Science* **323**, 1332-1335 (2009).
107. Durkin, K. *et al.* Serial translocation by means of circular intermediates underlies colour sidedness in cattle. *Nature* **482**, 81-84 (2012).
108. Kerje, S. *et al.* The Dominant white, Dun and Smoky color variants in chicken are associated with insertion/deletion polymorphisms in the PMEL17 gene. *Genetics* **168**, 1507-1518 (2004).
109. Lowry, D.B. & Willis, J.H. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS biology* **8**, 2227 (2010).
110. Clop, A. *et al.* A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nature genetics* **38**, 813-818 (2006).
111. Wang, Y. *et al.* The crest phenotype in chicken is associated with ectopic expression of HOXC8 in cranial skin. *PLoS One* **7**, e34012 (2012).
112. Altshuler, D., Daly, M.J. & Lander, E.S. Genetic mapping in human disease. *Science* **322**, 881-8 (2008).
113. Wade, C.M. *et al.* Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science* **326**, 865-7 (2009).
114. Ligon, J.D., Thornhill, R., Zuk, M. & Johnson, K. Male-male competition, ornamentation and the role of testosterone in sexual selection in red jungle fowl. *Animal Behaviour* **40**, 367-373 (1990).
115. Zuk, M., Johnsen, T.S. & Maclarty, T. Endocrine-immune interactions, ornaments and mate choice in red jungle fowl. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **260**, 205-210 (1995).
116. Ligon, J.D., Kimball, R. & Merola-Zwartjes, M. Mate choice by female red junglefowl: the issues of multiple ornaments and fluctuating asymmetry. *Animal Behaviour* **55**, 41-50 (1998).

117. Parker, T.H. & Ligon, D.J. Dominant male red junglefowl (*Gallus gallus*) test the dominance status of other males. *Behavioral Ecology and Sociobiology* **53**, 20-24 (2002).
118. Bed'hom, B. *et al.* The lavender plumage colour in Japanese quail is associated with a complex mutation in the region of MLPH that is related to differences in growth, feed consumption and body temperature. *BMC genomics* **13**, 442 (2012).
119. Marklund, S., Moller, M., Sandberg, K. & Andersson, L. Close association between sequence polymorphism in the *KIT* gene and the roan coat color in horses. *Mammalian Genome* **10**, 283-288 (1999).
120. Haase, B. *et al.* Five novel *KIT* mutations in horses with white coat colour phenotypes. *Anim Genet* **42**, 337-9 (2011).
121. Haase, B. *et al.* Accumulating mutations in series of haplotypes at the *KIT* and *MITF* loci are major determinants of white markings in Franches-Montagnes horses. *PLoS One* **8**, e75071 (2013).

Acta Universitatis Upsaliensis

*Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Medicine 1124*

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. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-259621



ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2015