Deciphering the genetic basis of animal domestication

Pamela Wiener* and Samantha Wilkinson

The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian EH25 9RG, UK

Genomic technologies for livestock and companion animal species have revolutionized the study of animal domestication, allowing an increasingly detailed description of the genetic changes accompanying domestication and breed development. This review describes important recent results derived from the application of population and quantitative genetic approaches to the study of genetic changes in the major domesticated species. These include findings of regions of the genome that show between-breed differentiation, evidence of selective sweeps within individual genomes and signatures of demographic events. Particular attention is focused on the study of the genetics of behavioural traits and the implications for domestication. Despite the operation of severe bottlenecks, high levels of inbreeding and intensive selection during the history of domestication, most domestic animal species are genetically diverse. Possible explanations for this phenomenon are discussed. The major insights from the surveyed studies are highlighted and directions for future study are suggested.

Keywords: animal domestication; breed differentiation; selective sweep; population bottleneck

1. INTRODUCTION

Understanding the history of domestication has been of interest to biologists at least since Darwin. He appreciated the wide variation within domesticated species, and throughout On the origin of species [1] (and later, in his two volumes of Variation under domestication [2]) he used them as examples of his theories. It is now well accepted that the process of animal domestication has involved a combination of human-imposed selection and non-selective forces, the latter including various forms of interference with the demography and mating programme of these species.

It is only recently, with advances in genetic and statistical technologies, that the genetic changes that have accompanied animal domestication and breed development can be characterized. A rapidly increasing number of species now have full-genome sequences. High-density, genome-wide single nucleotide polymorphism (SNP) panels have been produced for humans, as well as many other plant and animal species. A variety of statistical techniques have concurrently been developed to analyse this data. One of the key aspects of this analysis is to use genomic data in order to make inferences about the selective and demographic forces that have operated on individual species.

This review discusses the contribution that genetic data have made to our understanding of both selective and non-selective processes of evolutionary change in domesticated animal species, and the insights into the domestication process that have been revealed by these studies. Applications of various population genetic-based methods for the detection of genomic regions under selection are presented, as well as methods for elucidating non-selective processes; many of these, if not all, were first developed for human genetic analysis. This article does not attempt to review all the relevant literature, but rather to use specific examples to illustrate common themes. The examples presented are primarily taken from cattle, pigs, chickens and dogs, where the most advanced genetic resources are available.

2. SELECTIVE FORCES

The assumption underlying the detection of signatures of selection in the genome is that selection is locus-specific. By comparison, the effects of other evolutionary forces (random genetic drift, mutation and inbreeding) should be expressed genome-wide. Under this premise, the methods for detecting selected loci attempt to identify those at which allele frequencies have changed in a pattern consistent with positive selection. The methods differ in the information they use to find such loci (particularly as there are very few data on historical allele frequencies), which will be outlined further.

(a) Candidate gene studies

One approach adopted in domestication genetics is to examine patterns of diversity around candidate genes that, based on their function, are likely to have been targets of selection. Two such genes are growth differentiation factor 8 (GDF-8), associated with muscle conformation, and melanocortin 1 receptor (MC1R), associated with coat colour.

GDF-8 (myostatin) is a negative regulator of skeletal muscle growth, and naturally occurring mutations in this gene have been associated with increased levels of muscle conformation in cattle, dogs, sheep and humans. There is...
substantial diversity in GDF-8 across cattle breeds [3,4], including, at the extreme, two independent loss-of-function mutations that are associated with the ‘double muscling’ phenotype, in which animals have highly exaggerated muscle conformation. At the beginning of the twentieth century, the majority of Belgian Blue cattle had conventional conformation, and were used for both milk and beef production [5] (figure 1a). However, after less than a century of animal breeding, the double muscling phenotype is now nearly fixed in the breed, suggesting that there has been strong selection in favour of this trait, presumably owing to the increased amount of derived meat [8,9]. Analysis of microsatellite diversity in the region flanking the GDF-8 gene revealed a significant decrease in heterozygosity with increasing proximity to GDF-8 in three primarily double-muscled breeds, including the Belgian Blue, as well as in the sub-group of double-muscled South Devon cattle [6,7] (figure 1b), which was not seen in most non-double-muscled breeds. The pattern of heterozygosity in both Belgian Blue and South Devon cattle is consistent with strong selection on this gene. While evidence of a signature of selection near myostatin has not yet been published for other species, it is likely to exist as variation in this gene has been shown to influence traits of economic interest in breeds of dogs [10] and sheep [11].

Coat colour and pattern are key traits in the development of livestock and companion animal breeds, as they were under selection well before breed development [12]. A number of genes have been associated with coat colour in mammals, including the MC1R gene. MC1R influences the relative levels of eumelanin (black/brown) and phaeomelanin (yellow/red) pigments, and appears to have been a target for selection in pigs and other domesticated animal species. For example, there has been independent evolution of black coat colour in Asian and European pigs, and selection for this phenotype appears to have been particularly strong in Chinese pigs, where animals with the black coat were used preferentially in animal sacrifice rituals during the Neolithic period, because they were considered sacred [13]. Asian wild boar (the closest relative to the domestic pig) show extensive nucleotide variation at MC1R; however, nearly all European and Asian wild boar genotyped so far express the same MC1R protein, with genotypes differing primarily by synonymous substitutions [13,14]. This wild-type protein allows complete expression of both eumelanin and phaeomelanin pigments, and produces a coat in variable shades of brown [15]. In contrast, in domestic pigs, there is reduced synonymous variation relative to wild boar but at least nine different MC1R proteins in addition to the wild-type [14], which are associated with coat colour phenotypes ranging from red to black and including a variety of spotting patterns (white coat and spotting are determined by a different gene, KIT) (figure 2). Therefore, it appears that wild boar have been subject to purifying selection for a camouflaged coat, whereas a relaxation of this form of natural selection in combination with human-mediated selection for distinctive coat patterns has occurred in domestic pigs [14].

(b) Differentiation-based approaches

Changes within breeds have occurred on an evolutionarily short timescale compared with natural animal populations; however, there is considerable phenotypic
variation between domesticated animal breeds, particularly in dogs. Recent studies in various species have applied an approach where markers with strong evidence of genetic differentiation (e.g. high levels of Wright’s $F_{ST}$, a measure of genetic differentiation between populations, or allele-frequency differences) are taken as signals of differential selection across populations. This approach originated in the days when genetic markers were limited and sparse, and the focus was on specific markers [16,17], but in the current environment of dense, genome-wide markers for many species, genome scans of differentiation have become a viable strategy to identify selected genes or genomic regions using the tails of the genome-wide $F_{ST}$ distribution to define the significance threshold [18].

For this and other approaches, it has been recognized that instead of using single-locus statistical values, a sliding window analysis removes the stochastic variation between loci, and thus better highlights regions with signals of selection [19,20]. Although the population-differentiation approach was developed originally for analysis of human data (and is still used in this context [18,21]), this technique is possibly even better suited to studies of domesticated animals because breeds are in general genetically similar entities and the differences that do exist may reflect the relatively recent selection for breed-specific characteristics.

Akey et al. [22] conducted an $F_{ST}$ scan of the genome for 10 dog breeds and identified outliers, which they argued were candidates for targets of selection. This interpretation of the results was supported by the fact that five genes that had previously been mapped through association with ‘hallmark’ breed traits were among the 155 outlier SNPs (including the insulin-like growth factor 1 gene—$IGF1$, associated with body size—and several coat colour genes). Regarding the other outlier SNPs identified in their study, one of the highest $F_{ST}$ values was only found in the Shar-Pei breed, which is characterized by its distinctive skin-folding phenotype. The region where the high $F_{ST}$ signal was found contains several genes, including HAS2, the expression of which had previously been associated with skin wrinkling in this breed [23]. A recently discovered duplication upstream of this gene appears to be responsible for the wrinkling phenotype [24]. A separate study looking at genetic differentiation between 79 domestic dog breeds found that the top 11 $F_{ST}$ values measured across all breeds were found in genomic regions associated with morphological traits, including body size, skull and snout shape, coat characteristics and ear type [25].

For cattle, the genetic-differentiation approach has highlighted genomic regions that include genes encoding coat features or body size/conformation. Several studies have identified high levels of between-breed genetic differentiation near coat colour loci, including $MC1R$ (see §2a) and the Charolais dilution factor (Dc locus), indicating that these genes have been important in the establishment of cattle breeds [26,27]. Another gene that has been implicated as a possible target of selection based on allele-frequency differences between cattle breeds is the growth hormone receptor ($GHR$) gene [26–28].

Although it is clear that large qualitative effects have been detected using these methods, there are known to be limitations to $F_{ST}$-based methods for detecting genes with small or moderate effects. Wiener et al. [27] found the overall correlation between $F_{ST}$ and the statistical signal from linkage mapping analysis (see §2a) to be low in a study of two cattle breeds. While genes associated with coat colour could be detected as regions of large allele-frequency differences, the signals for loci associated with quantitative traits were generally weaker.

(c) **Frequency spectrum-based approaches**

A common approach to test for selection in human and wild plant and animal populations is to use ‘frequency spectrum’ tests in which empirical allele distributions are compared with those predicted under a neutral model. One set of methods involves searching the genome for regions with allele-frequency patterns that differ either from background (genome-wide) patterns or from those predicted by a neutral model [29,30]. These methods involve calculation of a composite log likelihood (CLL) for sliding window sets of genotypic data and testing significance based on a likelihood ratio test [29,30] or by permutation testing [31]. This approach has recently been applied to genome-wide SNP data for the 19 cattle breeds characterized by the Bovine HapMap Consortium [32]. In a follow-up analysis of this dataset, Stella et al. [31] calculated the difference for each SNP between the major allele frequency for a group of breeds defined by phenotype and the overall frequency across all breeds. For
black-coated breeds, there was a very strong signature of selection on BTA18 for windows that include the MC1R coat colour locus (see §2a). A signature of selection was also observed for polled (hornless) breeds on BTA1 within a region previously associated with presence/absence of horns. For dairy breeds, 699 putative signatures of selection were identified across the genome, with the highest (negative) CLL value on BTA6 near the KIT gene, which is associated with the level of white coat spotting in cattle. To make sense of the large number of significant results, the authors looked for cases where genes from the same gene family were at the centre of the significant window (e.g. potassium channel genes, integrins and arginine-/serine-rich splicing factors), arguing that these gene families may have been under selection during dairy cattle breeding.

Difficulties in applying frequency-spectrum-based tests to SNP data have been raised because of the bias towards high-frequency alleles inherent in SNP ascertainment, and thus interpretation of results can be problematic. While a number of solutions have been proposed to deal with this issue [33], in the long term the best remedy will involve use of full-genome sequence data in place of SNP data. Developments in next-generation sequencing are now making this a reality for many species (see §2d).

(d) Extended homozygosity approaches
Another population-genetic approach for the detection of selective sweeps has been to look for extended homozygous genomic regions. This approach is based on ‘hitchhiking’ theory [34], in which neutral variants increase in frequency owing to linkage disequilibrium (LD, the statistical association between allele frequencies at different loci) with alleles at a selected locus, resulting in reduced diversity across the region.

One particularly convincing example of reduced diversity near a selected locus relates to chondrodysplasia (shortened limbs) in dogs. A genome-wide SNP analysis revealed a 24 kb region of reduced heterozygosity on chromosome 18 in chondrodysplastic breeds (e.g. Dachshunds) relative to non-chondrodysplastic breeds [35]. This region includes an insertion of a retrogene encoding fibroblast growth factor 4 (FGF4) in the chondrodysplastic dogs, the expression of which may result in altered activation of one or more fibroblast growth factor receptors. A similar pattern of reduced heterozygosity near the IGF1 gene was observed in small dogs [36].

A number of statistical methods aim to distinguish the length of homozygous segments generated by selection from those generated by neutral processes, which extends the analysis beyond the heterozygosity of individual markers. One of the first methods introduced to exploit the hitch-hiking phenomenon in the context of high-density genotypic data was the long-range haplotype (LRH) test [37]. In this method, the age of each core haplotype in a genomic region is assessed using the length of extended haplotype homozygosity (EHH). Unusually, high EHH values suggest a mutation that increased more quickly than expected under a neutral model. In an alternative approach, the logarithm of the ratio of EHH for an ancestral allele to that for a derived allele (iHS) is used as the test statistic [38], such that large negative (positive) values of iHS indicate selection for the derived (ancestral) allele.

The extended haplotype-based methods have been applied mainly to human genetic data, but they have also been implemented for several cattle datasets. Studies by Hayes et al. [28,39] found high values of iHS for SNPs in several regions of bovine chromosome 6, including one region with the ABCG2 gene, associated with several dairy traits. The Bovine HapMap Consortium [32] also applied the iHS test across the genomes of 19 breeds and found high iHS values in one or more breeds on most chromosomes; these included regions on BTA2 near GDF-8, on BTA6 near ABCG2, and on BTA14 near a region associated with intramuscular fat. There were many other regions where a specific gene could not be implicated as a selection target. More recently, Qanbari et al. [40] applied the LRH test to denser (50 K SNP) data from Holstein dairy cattle. Although there were significant or nearly significant signals of selection for SNPs associated with some dairy-related candidate genes (e.g. the casein gene cluster encoding milk proteins and the DGAT1 gene associated with milk fat percentage), of the SNPs with greatest significance levels, none were found near these candidates.

The advent of whole-genome sequencing opens up new possibilities for the detection of selection signatures. Rubin et al. [41] sequenced whole genomes of eight pools of chickens representing commercial lines, experimental lines and breeds selected for specific traits. The genome was searched for regions of low diversity by calculating a normalized pooled heterozygosity measure in sliding windows. One of the lowest statistics (suggestive of positive selection) was found in the region of the beta-carotene dioxygenase 2 (BCDO2) gene, which is associated with skin colour in chickens. One or more regulatory mutations that inhibit expression of the BCDO2 gene appear to be responsible for the yellow skin phenotype [42]. Most chickens used for commercial egg and meat production in industrialized countries (as well as many local breeds worldwide) have the yellow skin phenotype and are homozygous for the recessive yellow skin allele locus, whereas other local chicken breeds have white skin and carry the dominant wild-type allele. The yellow skin allele appears to have been derived from a different ancestral species (possibly the grey junglefowl) than most of the commercial chicken genome (for which the red junglefowl is the presumed wild ancestor), suggesting a hybrid origin of commercial chickens (see §3b) [42].

The region with the lowest heterozygosity score across all domestic lines included the locus-encoding thyroid stimulating hormone receptor (TSHR) gene [41], which is involved in metabolic regulation and reproduction. This region was almost completely fixed over a 40 kb segment. Further analysis of this locus in domestic chickens from a number of countries revealed that each animal carried at least one copy of the derived haplotype (264/271 were homozygous). The role of TSHR in the domestication of chickens is still unknown; however, the authors suggest that it may be involved in the loss of seasonal reproduction present in non-domesticated relatives.

(e) Genotype-phenotype association analyses
A powerful approach for gene mapping in livestock species is linkage mapping, in which regions of the genome associated with particular traits (quantitative trait loci, QTL) are
identified. Populations generated by breed or line crosses have proved to be particularly useful for identifying the regions of the genome that distinguish the population founders. Although this technique generally identifies fairly wide intervals that include a large number of genes, in some cases it has led to the identification of individual genes that influence physical traits related to domestication, breed development or breed improvement (e.g. IGF2 in pigs [43], DGAT1 [44] and GHR [45] in cattle).

QTL-encoding physical traits may also be associated with behavioural traits. One such instance is the PMEL17 gene encoding plumage colour in chickens, which was identified from a cross between red junglefowl and the commercial White Leghorn. A 9 bp insertion in exon 10 acts in a dominant fashion, such that birds homozygous for the ancestral junglefowl allele (I) are black, whereas those carrying the White Leghorn allele (I) are white (heterozygotes sometimes have minor pigmentation). It has been demonstrated that there are substantial behavioural differences between birds carrying the junglefowl and White Leghorn alleles, such that i/i individuals birds are more vocal, have lower activity levels in a test measuring fear of humans, and are more aggressive, social and explorative [46–48]. This locus may also be associated with feather-pecking, a bullying behaviour that is particularly interesting in that it demonstrates the possibility of selection for correlated traits in domesticated animals. It is likely that the behavioural traits associated with PMEL17 were not the target traits in the development of the White Leghorn breed but may have been co-selected owing to selection for white plumage. Association between behavioural traits and coat colour appears to be a common phenomenon. Genes in the melanocortin system (including MC1R and the agouti gene) have been associated in mice and other vertebrate species (e.g. lions, lizards and birds) with both coat colour and various behavioural traits, including aggressiveness, sexual behaviour and learning behaviour [51]. Eumelanin-based coloration is generally associated with more aggressive behaviour. In her treatise on captive breeds, Felius [12] claims that the Romans and later Europeans also associated coat colour with cattle performance traits: a red coat (the most common phenotype) was associated with a ‘fiery’ and hard-working character, whereas the rare white coat was associated with a sluggish and lazy disposition. However, the genetic association between behavioural traits and coat colour is not universal, as demonstrated by a study on rats in which ‘tameness’ QTL (see below) were on different chromosomes from a QTL segregating for white coat spotting [52].

In some cases, correlated selection appears to go in the other direction, such that selection for behavioural traits may result in associated changes in more visible phenotypes, as has been seen in the well-described selection experiment involving the silver fox (the ‘farm-fox experiment’) [53]. Initiated in 1959 in Novosibirsk, Siberia, the original fox population showed continuous variation for tameness/aggressiveness. A breeding programme was established with 100 females and 30 males, from which foxes were selected for their tameness using severe selection criteria [54]. The resulting population of tame foxes behaved much like domestic dogs. Behavioural traits other than tameness also evolved (e.g. tail wagging, licking). Moreover, in addition to the changes in behaviour, other morphological changes also occurred, some of which are reminiscent of dog breeds. For example, traits such as floppy ears, curly tails and shortened snouts appeared in some foxes. Recent development of a linkage map for the silver fox [55] has allowed QTL analysis of backcross and intercross populations derived from the tame population and an unselected (aggressive) population. QTL for several tameness-related behavioural traits map to fox chromosome 12; however, it is still unclear whether these are associated with a single locus [56]. Furthermore, inconsistencies between results from different crosses suggest a complex inheritance pattern (e.g. strong epistatic interactions) for these traits.

The study of silver foxes suggests that laboratory selection for behavioural traits can emulate the process of domestication. Other researchers from Novosibirsk conducted an experiment selecting for reduced or enhanced aggression to humans in a population of wild-caught rats [57]. Like the silver fox, this population has recently been exploited using genetic techniques to map regions of the genome associated with ‘tameness’ (as referred to above), defined by a linear combination of a set of behavioural traits [52]. QTL analysis indicates that more than one region is involved in the evolution of tameness in these rats [52] and that individual QTL may comprise multiple sites [58].

Modification of behaviour is believed to have been one of the key aspects of animal domestication, including selection for ‘reduced fear, increased sociability and reduced anti-predator responses’ (p. 5 in [59]). As dog breeders and owners know well, behaviour is also associated with breed differences. In an investigation of four composite personality traits (playfulness, curiosity/fearlessness, sociability and aggressiveness) in 31 dog breeds, Svartberg [60] found large differences between breeds for all traits. For example, popular pet breeds tended to have higher sociability and playfulness scores than less popular breeds.

3. NON-SELECTIVE FORCES

While selection has clearly been an important force in the history of animal domestication, as with wild species, other non-selective mechanisms have strongly influenced evolutionary change in these species. There are various approaches that allow inferences about demographic and mating processes using genetic data.

(a) Human-mediated modifications to population size and structure

One important advance with the advent of dense markers is the ability to exploit the relationship between LD and effective population size (Ne, the number of individuals in an idealized population that would have the same rate of genetic drift as the actual population), such that Ne and r² (the correlation between allele frequencies
There is clear evidence of declining \( N_e \) in commercial animal breeds, and in some cases this has resulted in extremely low variability. A feral British breed, Chillingham cattle, was found to be homozygous at 24 out of 25 microsatellite loci [73], which is strikingly low when compared with other British cattle breeds [74]. The low levels of homozygosity in the Chillinghams presumably result from a very severe bottleneck and absence of immigration. Looking over a longer timescale, ancient \( B. \ taurus \) DNA has revealed a reduction in diversity at several cattle genes over the last 4000 years [75]. It is not yet clear whether this is a genome-wide or loci-specific pattern.

(b) Introgression
Another human-related phenomenon that is manifested in the architecture of genomes is that of introgression between breeds. Animal breeders may practice cross-breeding to introduce certain desirable traits for breed improvement. In the case of pig breeds, past human activity has influenced the genetic composition of European breeds. In the 18th and 19th centuries, Asian alleles were introduced into certain British pig breeds to promote traits such as fattening and earlier maturation [2]. Breeds that experienced genetic introgression included Berkshire and Middle White, and Asian morphological characteristics such as the squashed face of the Middle White are still evident (see figure 2). Molecular studies have since provided genetic evidence of the introgression from Asia to Europe. A study examining mitochondrial diversity in pigs revealed that a number of European commercial pig breeds carry Asian-like mitochondrial DNA (mtDNA) haplotypes [76]. The levels of Asian genetic introgression were highly variable, depending on the breed and commercial line, with an average of 29 per cent frequency of Asian mtDNA haplotypes across European breeds. Genetic introgression can also be non-human-mediated, such as gene flow from wild relatives into the domestic pool and vice versa. For example, a Chinese wild boar genotyped by Fang et al. [14] carried an \( \text{MC1R} \) allele common in European domestic pigs, which must have resulted from gene flow. It is not clear whether the introgression of grey junglefowl into the primarily red junglefowl background of commercial chickens, suggested by the presence of the yellow skin phenotype [42] (see §2d), was a human-mediated event.

4. LEVELS OF GENETIC DIVERSITY
One of the most interesting and somewhat surprising findings arising from genetic studies of domesticated animals is that despite the role of intensive selection, inbreeding and population bottlenecks, many domesticated animal species are characterized by a high degree of genetic diversity. Cattle, particularly \( B. \ indicus \) breeds, have substantial nucleotide diversity [32], indicating a large ancestral effective population size. There is also evidence from a number of individual genes that nucleotide variation is relatively high in domesticated pigs [77], where sustained gene flow with their wild boar relatives (see §3b) appears to play an important role [78]. Despite the extensive bottleneck and associated loss of alleles that accompanied the commercialization of broiler and layer lines [72], domestic chickens have extensive sequence diversity [79], again presumably owing to a very large ancestral population which had even
greater levels of diversity (as also seen in present-day red junglefowl [80]). These high levels of genetic diversity contribute to the continuing ability of breeders to select for production traits. Despite the very low effective population size of the Holstein, average milk yield has continued to increase [81]. Similarly, heritability for growth in broiler chickens has remained at a similar level despite intensive selection over the last 50 years [82].

Certain livestock breeds with particularly low population size (such as Chillingham cattle, discussed in §3a) and some purebred dogs appear to be exceptions to this pattern. Many dog breeds were established with very low initial sizes, resulting in highly inbred populations and a high prevalence of inherited diseases (e.g. syringomyelia in Cavalier King Charles Spaniels and atopic dermatitis in various breeds [83]), presumably owing to the high frequency of individuals homozygous for recessive alleles. This is reflected in the low level of nucleotide diversity seen in the dog, when compared with chickens and cattle (electronic supplementary material, table S1).

While there are many indicators to show that genetic variation is being lost in domesticated animals, this appears to be operating within an overall context of high levels of diversity in most cases, and therefore can be counteracted by informed breeding decisions. This is not to suggest that conservation and breed management is not required, but rather that animal breeding has not yet reached a point of no return.

5. CONCLUSIONS

(a) Preliminary insights from genomic analyses

Although identification of the genes important in animal domestication and breed development is still in its early stages, some common themes have emerged. One is that there are clearly strong signatures of selection near a number of genes associated with coat colour and pattern (e.g. MC1R, KIT). This should not be surprising in that these visible phenotypes provide a clear-cut mechanism for farmers and breeders to distinguish their animals from others, and in some cases have served a cultural role. Coat colour and pattern remain important features of breeds and are still under selection. For example, Red Angus cattle breeders have formed separate breed societies from Black Angus in a number of countries in part because the red coat (an MC1R variant) is thought to be more heat- and sun-tolerant than black.

There are also genomic indicators that suggest selection on genes related to growth and body composition. There is clear evidence in several cattle breeds for selection on the myostatin gene, associated with muscle composition, and several studies also suggest that there may have been selection on the GHR gene, associated with growth rate and various production traits. In dogs, there is also evidence of strong selection on a number of genes associated with growth (e.g. IGF1) and skeletal traits, many of which are related to breed-specific characteristics. The genomic picture of selection for dairy-related traits is somewhat cloudier than that seen for other cattle production characteristics. There is some indication of selection signals near the ABCG2 and DGA1 genes, which have been associated with milk-production traits, but this is not consistent across studies.

Although these studies have indicated several genes that appear to have been under selection and have highlighted demographic events, they also suggest difficulties in fully characterizing the history of animal domestication using genetic data because of the concurrent action of multiple factors. Both selective and non-selective forces have clearly played key roles in the history of most domesticated species, and it may be difficult to separate these factors. For example, extended homozygosity and increased LD can derive from population contraction and/or inbreeding as well as strong selection, leading to problems distinguishing between these causes [84].

(b) Directions for further study

Improvement and further development of statistical methods for identification of selection signals is an active area of investigation. In addition to the need for better ways of distinguishing between demographic and selection processes, new approaches may be required to adequately investigate the role of selection on quantitative traits such as milk yield. Low power to detect selection on quantitative traits [27] may help to explain the inconsistent picture of selection signals seen in dairy cattle.

Another important area of further research is the identification of the genes that have been selected for their impact on tameness and other domestication-related behavioural traits. While progress is currently being made in this direction, the study of the genetic basis of these traits is still in its infancy. Long-term experimental selection for tameness in the silver fox has provided valuable insight into the domestication process and promises to provide even greater understanding once genomic techniques are applied to this population. The loci underlying the rat and the fox tameness QTL do not map to orthologous regions [56], and thus these studies have already demonstrated that there are multiple genetic routes to evolving tameness.

As demonstrated by the silver fox and tame rat studies, experimental populations may provide great insight into the process of domestication. There have been several recent studies examining genetic changes over the course of experimental selection on Drosophila [85] and chicken [86] lines. More extensive analysis of this type of data, especially when genetic material is collected from different stages of the experiment, may allow inference of the processes of domestication that cannot be measured directly. A complementary and more direct approach involves the analysis of ancient genetic material from different historical periods. As techniques for working with these samples improve, they will increasingly provide insights into the genetic changes that have accompanied the domestication process [87].

The Roslin Institute is supported by a core strategic grant from the UK Biotechnology and Biological Sciences Research Council (BBSRC). S. Wilkinson is funded by a CASE studentship from the BBSRC and Rare Breeds Survival Trust.

REFERENCES


73 Visscher, P. M., Smith, D., Hall, S. J. G. & Williams, J. A. 2001 A viable herd of genetically uniform cattle: deleterious alleles seem to have been purged in a feral strain of inbred cows. *Nature* **409**, 303. (doi:10.1038/35053160)


