

GENETIC DISSECTION OF PHENOTYPIC DIVERSITY IN FARM ANIMALS

Leif Andersson

Farm animal populations harbour rich collections of mutations with phenotypic effects that have been purposefully enriched by breeding. Most of these mutations do not have pathological phenotypic consequences, in contrast to the collections of deleterious mutations in model organisms or those causing inherited disorders in humans. Farm animals are of particular interest for identifying genes that control growth, energy metabolism, development, appetite, reproduction and behaviour, as well as other traits that have been manipulated by breeding. Genome research in farm animals will add to our basic understanding of the genetic control of these traits and the results will be applied in breeding programmes to reduce the incidence of disease and to improve product quality and production efficiency.

PURIFYING SELECTION
Selection against a deleterious allele.

Several of the main farm animals (cattle, sheep, goat and pig) were domesticated 9,000–11,000 years ago, whereas the dog was domesticated earlier, ~14,000 years ago, and the chicken ~4,500 years ago. Whereas humans seem to have expanded enormously from a small population present roughly 100,000 years ago¹, most domestic animals have a much broader genetic basis. Molecular studies have shown that the two main forms of domestic cattle, European–African (*Bos taurus taurus*) and Asian (*Bos taurus indicus*), originate from two different subspecies of the wild ancestor^{2,3}. Similarly, both a European and an Asian subspecies of the wild boar have contributed to European and Asian breeds of domestic pig⁴. Phenotypic selection has created a wide diversity of breeds of domestic animal that are adapted to different climatic conditions and purposes. The phenotypic variation that is observed within and among breeds of domestic animals is overwhelming compared with that observed in natural populations (FIG. 1). Charles Darwin was the first to recognize that phenotypic diversity in crops and domestic animals that occurs because of breeding mimics evolution in natural populations that occurs because of natural selection⁵. In fact, phenotypic change under domestica-

tion was one of Darwin's strongest arguments for the evolution of new species by natural selection. Genomics now provides more and more powerful tools for unravelling the molecular basis of phenotypic diversity in domestic animals.

Genome research in farm animals differs in several respects from that in humans or experimental organisms. Notwithstanding the general interest in multifactorial genetics, the identification of simple monogenic disease loci in farm animals is less important, because animals with inherited disorders (and their parents) tend to be eliminated from breeding. Most traits of interest, such as growth, milk production and meat quality, have a multifactorial background and are controlled by an unknown number of quantitative trait loci (QTL; BOX 1). Mutations that modify gene function or gene expression dominate over mutations with pathological consequences because the latter tend to be eliminated by PURIFYING SELECTION. Selective breeding has been going on for thousands of years and with increasing intensity during recent centuries. Such selection, over many generations and in large populations, has driven the accumulation of new mutations with favourable phenotypic effects, as well as the development of alleles

Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Box 597, S-751 24 Uppsala, Sweden. e-mail: Leif.Andersson@bmc.uu.se



Figure 1 | **Phenotypic diversity of domestic animals.** The illustration shows the phenotypic diversity of different breeds of domestic chicken in comparison with the wild ancestor, the Red Jungle Fowl (a male and female are shown at the bottom). (Painting by Staffan Ullström.)

and haplotypes that differ by multiple functionally significant substitutions.

The main goal of genome research in farm animals is to map and characterize trait loci that control various phenotypic characters. The aim of this review is to show why farm animals provide unique resources for studying genotype–phenotype relationships. The strategies and prospects for mapping and cloning trait loci are reviewed and selected examples of interesting monogenic and multifactorial trait loci are given.

Genome resources in farm animals

The mapping and identification of trait loci, in particular QTL, require powerful genome resources. To a large extent, the strategies and resources used for genome research in farm animals have followed the way paved by the human genome project. Dense microsatellite maps comprising greater than 1,000 markers have been developed for all the main farm animals (status reports are available at the [Roslin Institute](#) and the [US Livestock](#)

[Genome Mapping Projects](#) web sites). The numbers of markers are sufficient for finding trait loci by linkage analysis. However, once a trait locus has been mapped, additional polymorphic markers often need to be developed for positional cloning. Single nucleotide polymorphisms (SNPs) are now being developed by several laboratories, but no major initiative like the [Human SNP Consortium](#) has yet been taken.

Large-insert yeast artificial chromosome (YAC) and bacterial artificial chromosome (BAC) libraries have been established for all the main farm animals. These libraries are now extensively used for cloning and characterizing trait loci. Radiation hybrid (RH) panels are also available for pig⁶, cattle⁷ and chicken⁸, and are used for developing high-resolution maps and, in particular, comparative maps⁹. The main strength of the RH strategy is that markers and genes can be ordered along the chromosome with a high resolution without the need for polymorphisms and extensive pedigree materials, unlike linkage mapping.

The development of a [preliminary transcript map](#) using expressed sequence tags (ESTs) has been an important component in the human genome project. The numbers of ESTs reported in farm animal species are small compared with that in humans (TABLE 1) but the numbers for pig and cattle have increased recently owing to a large scale initiative at the [US Meat Animal Research Center](#) (MARC). Large collections of sequenced cDNA clones are needed to allow animal researchers to take full advantage of the potential of transcript profiling using array technologies¹⁰. Existing arrays from other species (for example, humans or mice) might be useful but will give a lower signal-to-noise ratio than species-specific arrays. The development of species-specific arrays is therefore justified for all the main domestic animals.

The number of genes mapped in farm animals is still small compared with the situation in humans and mice. Comparative mapping to exploit the wealth of map information from other species is therefore an important component of farm animal genome programmes. [CHROMOSOME PAINTING](#) using human chromosome-specific probes has been successfully used to establish regions of conserved synteny between mammalian farm animals and humans¹¹ (FIG. 2), whereas linkage mapping has been used for comparative mapping in chicken¹². It seems that the genome organization is better conserved between humans and all farm animals, including chicken, than between humans and mouse or rat. This finding is unexpected considering the evolutionary relationship of these species, but is consistent with the well-established fact that the rate of chromosomal evolution differs considerably between phylogenetic lineages^{12,13} and that the genomes of rodents are more rearranged than other mammals. We find large regions of conserved synteny (up to whole chromosomes or chromosome arms) in comparisons between domestic animals and humans, but the gene order can be rearranged^{9,14}. Comparative mapping will allow extensive use of the near complete human transcript map.

CHROMOSOME PAINTING
Fluorescent *in situ* hybridization (FISH) to chromosomes using a probe that represents a whole chromosome or a part of a chromosome.

BREEDING VALUE

The genetic merit of an individual estimated using the phenotypic deviation of its offspring from the population mean.

GENETIC DRIFT

The random fluctuation in allele frequencies as genes are transmitted from one generation to the next.

POPULATION STRATIFICATION

The population is divided into sub-populations. In farm animals, it is very common that some breeding animals are used more frequently in some herds than in others. This generates allele frequency differences between sub-populations and linkage disequilibrium in the population.

Box 1 | Quantitative trait loci

Quantitative traits, such as weight and length, show a continuous distribution of phenotypic values rather than the discrete values observed for a qualitative trait. Quantitative traits are usually controlled by multiple genes and influenced by environmental factors. A quantitative trait locus (QTL) is defined as a region of the genome that harbours one or more genes affecting a quantitative trait⁶⁹. The QTL concept is also used for traits with discrete distributions (such as sick or healthy) but where one assumes that multiple loci control the phenotype. The total number of QTL that control a given trait is not absolute, but depends on an arbitrarily chosen threshold, below which the effect of a locus is deemed too small to be considered as a QTL. Classical quantitative genetics theory assumes that quantitative traits are controlled by an infinite number of genes, each with an infinitesimal effect. This is of course only a statistical model and gene mapping studies have revealed that QTL with substantial effects exist. The presence of QTL is detected by gene mapping studies that show significant differences in phenotypic traits between individuals that have inherited different QTL alleles from their parents. The molecular basis for QTL is still largely unknown, but QTL mapping in *Drosophila* suggests that QTL are often associated with sequence variation in non-coding DNA³⁰. The cloning of a QTL is a challenge for several reasons. A major hurdle is the poor precision in the location of QTL. This is because the relationship between the genotype and the phenotype is more complex than it is for a monogenic trait and therefore one cannot directly identify recombinants between markers and trait loci. However, it is possible to determine the genotype at a QTL indirectly by progeny testing, which means that the segregation at the QTL is deduced using data on genetic markers and phenotypic variation among the progeny (see also FIG. 3).

Strategies for finding trait loci

There has been a long tradition of collecting and analysing data on phenotypic traits for breeding purposes in farm animals, and the most common strategy for finding trait loci is to use existing pedigrees. This approach is made easier in farm animals because of the large family sizes; for example, half-sib families comprising more than 1,000 progeny from a single male can be collected in species such as cattle, in which artificial insemination is practised. Moreover, it is possible to collect extended pedigrees that comprise multiple generations. A useful strategy is to increase the statistical power in QTL mapping by using BREEDING VALUES based on phenotypic data from progeny (FIG. 3). This strategy is used extensively to identify QTL for milk production traits, as well as for other traits of interest in dairy cattle^{15–17}.

There are two main strategies for finding trait loci — association tests using candidate genes and genome scans based on linkage mapping with anonymous DNA markers. The candidate gene approach can be very powerful and can detect loci even with small effects, provided that the candidate gene represents a true causative gene. However, there are often many candidate genes for the trait of interest and it might be more time-consuming to evaluate all of those than it is to do a genome scan.

Furthermore, the candidate gene approach might fail to identify a major trait locus simply because of the gaps in our knowledge about gene function. Candidate gene tests must also be interpreted with caution because spurious results can occur because of linkage disequilibrium to linked or non-linked causative genes (see below) or because the significance thresholds have not been adjusted properly when testing multiple candidate genes. A genome scan will always find the map location of a trait locus with a major effect, provided that an accurate genetic model has been postulated, a reasonable sample size has been used and that the marker set provides full genome coverage. However, a genome scan will fail to detect trait loci with smaller effects if they do not reach the stringent significance thresholds that must be applied when doing a large number of tests in a full genome scan.

High-resolution mapping of trait loci can be carried out by identity-by-descent (IBD) mapping^{18,19}. IBD mapping involves the collection of DNA samples from individuals that have inherited a certain allele at a trait locus from a common ancestor. The samples are screened with genetic markers to detect the minimum chromosomal region identical-by-descent among the individuals. Linkage disequilibrium is also more frequently observed in farm animals than in large outbred populations such as the human population. A recent study of dairy cattle populations revealed fairly strong linkage disequilibrium for loosely linked loci and even weak linkage disequilibrium between loci on different chromosomes²⁰. Linkage disequilibrium in farm animal populations is generated by GENETIC DRIFT, the effects of which are more marked because of the limited number of breeding animals, migration and POPULATION STRATIFICATION. The presence of linkage disequilibrium increases the probability of finding true associations between genetic markers and trait loci but also increases the risk of spurious associations in candidate gene studies.

A powerful approach for mapping trait loci is to use intercrosses between divergent populations. The F₁ animals show a high heterozygosity at marker loci and, in particular, at those loci that account for phenotypic dif-

Table 1 | Number of ESTs in public databases*

| Species | Number of expressed sequence tags |
|---|-----------------------------------|
| <i>Homo sapiens</i> (human) | 2,786,452 |
| <i>Mus musculus</i> and <i>domesticus</i> (mouse) | 1,841,172 |
| <i>Rattus sp.</i> (rat) | 233,529 |
| <i>Bos taurus</i> (cattle) | 141,567 |
| <i>Danio rerio</i> (zebrafish) | 76,310 |
| <i>Sus scrofa</i> (pig) | 48,790 |
| <i>Gallus gallus</i> (chicken) | 12,995 |

*dbEST summary by organism, 8 December 2000 (see links).

ZOO-FISH

Fluorescent *in situ* hybridization (FISH) of a chromosome-specific probe from one species to chromosomes from another species.

ADVANCED INTERCROSS LINES (AIL)

Subsequent generations (F_3 , F_4 and so on) of an intercross pedigree, maintained to allow high-resolution mapping of trait loci.

RESOURCE POPULATION

A population generated for particular research purposes, such as an intercross between two divergent breeds of farm animal or a population containing particularly interesting phenotypic data.

INTROGRESSION

Transfer of genetic material from one population to another by repeated backcrossing.

GERM PLASM

The physical basis of heredity, and therefore the genetic material used for breeding.

MARKER-ASSISTED SELECTION

The use of genetic markers to predict the inheritance of alleles at a closely linked trait locus.

INTERSEXES

Individuals that have a mixture of male and female characters.

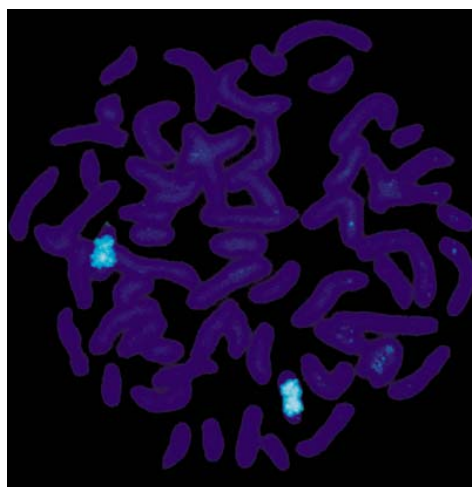


Figure 2 | Conserved genome organization among vertebrates. Results of a ZOO-FISH experiment in which a human chromosome-17-specific probe shows hybridization to a single cattle metaphase chromosome (number 19). The result shows that most of the conserved sequences on human chromosome 17 are present on cattle chromosome 19. (Picture provided by Dr Bhanu Chowdhary)

ferences between the two populations. Several intercrosses have been generated in farm animals, such as crosses between the European Wild Boar and Large White domestic pigs²¹ (FIG. 4a), Asian and European breeds of pig^{22,23} and *Bos taurus* and *Bos indicus* cattle²⁴. These studies have revealed several QTL in the pig with large effects on body composition. A combined analysis of almost 3,000 pigs from seven different intercross experiments revealed consistent QTL effects on growth and fatness for a region on pig chromosome 4 (REF. 25). High-resolution mapping of QTL can be obtained by backcross experiments using animals that carry recombinant chromosomes. The development of ADVANCED INTERCROSS LINES (AIL) is another useful approach for

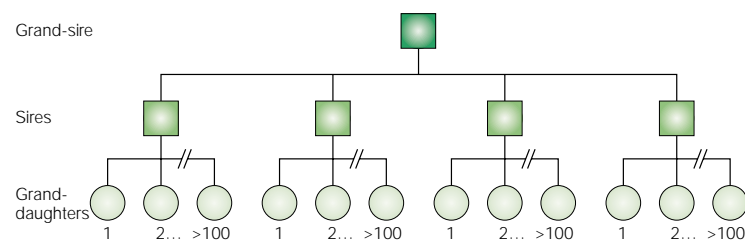


Figure 3 | Half-sib families and breeding value. The pedigree illustrates the grand-daughter design for mapping quantitative trait loci (QTL) using half-sib families⁷⁰. Each family comprises one grand-sire whose sons (20 or more) have been selected as sires. Each son has 100 or more daughters with phenotypic data, and these are used to estimate accurate breeding values for the sires, even for traits with a considerable environmental influence. Marker genotypes are only collected for the grand-sire and his sons, and QTL mapping is carried out by analysing the segregation of markers from the grand-sire to the sires in relation to differences in breeding values. If the grand-sire is heterozygous for a QTL with a major effect, the sons that have received the favourable allele will tend to have higher breeding values than those that received the unfavourable allele.

improving map resolution in intercross experiments^{26,27}. The generation of backcross and AIL populations is costly for the larger farm animals and this approach will only be used for particularly important RESOURCE POPULATIONS. However, cross-breeding is frequently used for genetic improvement and synthetic breeds are sometimes generated by intercrossing two or more breeds. It should be possible to use such existing populations for high-resolution mapping of trait loci. Giuffra *et al.*⁴ have even proposed that some European pig breeds could be used for this purpose, because of INTROGRESSION OF ASIAN GERM PLASM in the eighteenth and nineteenth centuries.

Once the chromosomal localization of a trait locus has been determined, this information can be applied in breeding programmes by using MARKER-ASSISTED SELECTION. However, the ultimate goal when mapping trait loci is the identification of the causative genes and causative mutations. Positional candidate cloning will continue to be the main strategy for this purpose. High-resolution mapping is a first step towards restricting the region of interest and thereby the number of potential candidate genes. Information on map location and gene function is then combined to identify positional candidate genes, which are subsequently evaluated by mutation screening and functional analysis. Positional candidate cloning in farm animals often relies heavily on the exploitation of comparative data and will become even more powerful with the completion of the human map and the generation of informative databases on gene function and gene expression patterns.

Pure positional cloning will only be used rarely in farm animals. So far, there is only one example of a successful positional cloning in a farm animal²⁸ (see below). Positional cloning will be used in those cases in which the human and mouse maps are incomplete, when there are map rearrangements between species or when a gene has evolved a new function in the farm animal. The positional cloning of the gene for polledness (lack of horns) in goats (a case in which the human transcript map might be of little help) should be accomplished in the near future because the gene has been precisely mapped and a large-insert contig has been constructed²⁹. The gene is particularly interesting because it is associated with the development of XY male INTERSEXES in homozygotes.

The poor precision in QTL mapping (BOX 1) implies that the molecular identification of these loci will primarily rely on positional candidate cloning because it is difficult to collect or generate sufficiently large pedigrees to allow the map resolution needed for a pure positional cloning approach with current technology. Furthermore, it might be difficult to prove a causal relationship because QTL mutations will often be in regulatory rather than in coding regions³⁰, and the phenotypic effect will usually be subtle compared with the simple loss-of-function mutations that cause inherited disorders.

Examples of major traits and trait loci
Having considered general aspects of variation in domestic animals, in this section I review examples of

trait loci for which the causative gene and mutation have been identified, or where this can be expected to happen in the near future. It should be noted that most of the examples I discuss are for traits in which a single major gene has provided the starting point for the analysis.

Coat colour variation. The coat colour of farm animals shows remarkable diversity, in contrast to the limited variation within natural populations (FIGS 1 and 4). This is partly explained by a release from natural

selection against individuals with an odd appearance such as would happen in prey–predator interaction and in mate selection. Furthermore, breeders have apparently selected for coat colour diversity: at least during the past 200 years, coat colour has been used as a trademark for different breeds. Coat colour is therefore surprisingly important in animal breeding, and several diagnostic DNA tests for coat colour variation are used by the industry at present. Molecular coat colour genetics in farm animals relies heavily on mouse coat colour genetics. As an example, mutations at *Mclr*, which encodes the melanocortin receptor 1, were first shown to control the distribution of red and black pigment in the mouse³¹ and subsequently in cattle³², horses³³, pigs³⁴, sheep³⁵, dogs³⁶ and chickens³⁷.

The *Dominant white* allele in pigs is a particularly interesting example. Many domestic pigs are white because they lack melanocytes in the skin (FIG. 4a). Molecular studies have shown that the dominant white colour is determined by mutations at *KIT*, which encodes the mast/stem-cell growth factor receptor^{38,39}. *KIT* is crucial for the survival of migrating melanocytes during embryogenesis and for haematopoiesis and germ-cell development. Loss-of-function mutations in *KIT* cause *Dominant white spotting* in mice and the dominant *Piebald* trait in humans, but often have a severe or lethal phenotype in the homozygous condition. The *Dominant white* allele in pigs has a more drastic effect on pigmentation than any known mouse mutation, but homozygotes are still fully viable. This puzzle is solved by the observation that the *Dominant white* allele involves two mutations — a duplication and a splice mutation^{38,39} (FIG. 4b). The duplication is found on its own in the *Patch* allele, which causes large patches of unpigmented skin and hair, probably owing to overexpression or ectopic expression of *KIT*. The splice mutation causes skipping of exon 17 and is assumed to result in a receptor with normal ligand binding but no intracellular tyrosine kinase activity. The combination of overexpression and a structural mutation in the same allele is thought to explain the extreme effect on pigmentation in *Dominant white* heterozygotes. *KIT* also has a crucial role in the normal development of blood cells, and the presence of the duplicated but otherwise normal copy rescues homozygous *Dominant white* animals from lethality. This illustrates that alleles at trait loci under selection for many generations can differ by multiple mutations.

Body composition. Selection based on body composition, in particular the relative proportion of muscle to fat tissue, is very important in meat-producing animals. During the past 50 years, there has been an intensive selection for lean growth in several breeds. Several genes that influence body composition have already been identified or are close to being identified.

The *Halothane* locus in pigs was one of the first trait loci to be characterized at the molecular level⁴⁰. A recessive mutation at this locus causes susceptibility to **malignant hyperthermia**, which can be triggered by stress or exposure to the anaesthetic gas halothane. This mutation

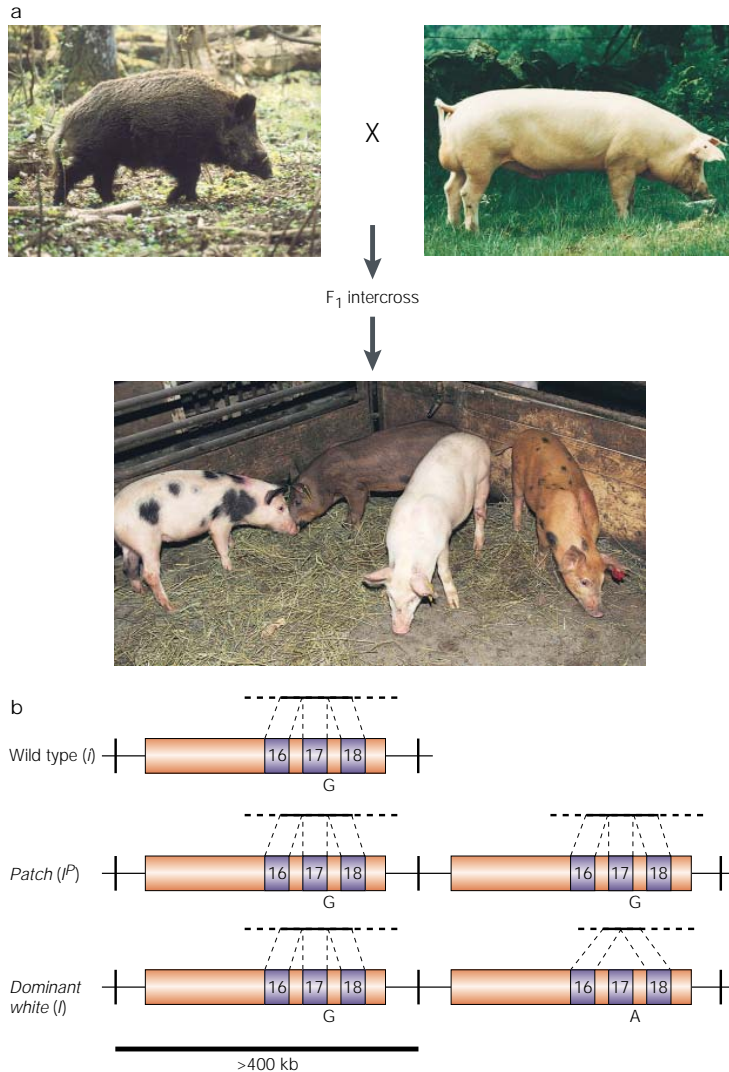


Figure 4 | The use of intercrosses between divergent populations. **a** | An intercross between the European Wild Boar (left) and Large White domestic pigs (right), and the segregation of coat colour diversity in the F₂ generation. The white founder and the white F₂ animals carry the *Dominant white* allele at the *KIT* locus, whereas the Wild Boar and the three coloured progeny are homozygous for the recessive wild-type allele. **b** | Three *KIT* alleles in the pig. The tandem duplication present in the *Patch* and *Dominant white* alleles is larger than 400 kb and it includes the entire *KIT* coding sequence (L. A. *et al.*, unpublished observations). G and A indicate the first nucleotide in intron 17. The splice mutation G→A causes skipping of exon 17. (Photographs of the Wild Boar, Large White and F₂ animals were provided by Bo Kristiansson, Quality Genetics AB and Mats Gerentz, respectively.)



Figure 5 | Double-muscling in cattle. Comparison of a Belgian blue bull (top) showing the double-muscling phenotype and a Charolais bull (bottom) without this phenotype; note the marked muscular hypertrophy of the hind leg in the Belgian blue animal. The double-muscling phenotype is caused by homozygosity for a loss-of-function mutation in the *myostatin* gene^{46–48}. The white colour in the Belgian blue animal is associated with a missense mutation in *MGF*, which encodes the KIT ligand⁷¹.

was also associated with a higher lean meat content and increased in frequency because of strong selection for this character. The gene causing this phenotype is *RYR1*, which encodes the ryanodine receptor 1 — an ion channel that regulates the release of Ca^{2+} in skeletal muscle. *RYR1* was identified as a positional candidate gene, and sequencing of *RYR1* cDNA from a homozygous normal and a homozygous mutant animal revealed a single missense mutation, R614C, in the mutant^{40,41}. Interestingly, the same missense mutation causes **malignant hyperthermia** in some human families.

The *RN⁻* allele is another example of a mutation that has probably increased in frequency because of selection for meat content in pigs. It occurs at a high frequency only in the Hampshire breed and increases glycogen content in muscle by ~70%. The mutation is associated with poor processing quality when producing ham and low pH in the meat because of post-mortem degradation of glycogen. Positional cloning recently showed that this phenotype is caused by a missense mutation (R200Q) in *PRKAG3*, which encodes an isoform of the regulatory γ -subunit of AMP-activated protein kinase (AMPK)²⁸. The distinct phenotypic effect of this mutation indicates that this AMPK isoform has an important role in energy metabolism in skeletal muscle. It might be involved in insulin-independent glucose uptake by

skeletal muscle and is therefore a potential drug target for treatment of type II diabetes in humans.

A third locus with an important effect on muscle development and body composition in the pig is a QTL located at the distal tip of pig chromosome 2. This locus was independently identified by Jeon *et al.*⁴² by analysing a Wild Boar–Large White intercross and Nezer *et al.*⁴³ by using a Pietrain–Large White intercross. The QTL has a large effect on meat content in the pig and shows clear paternal expression. It provides a valuable opportunity to study the molecular basis for a QTL because the effect is large, and it is possible to exclude all positional candidate genes that do not show paternal expression. The fact that the QTL maps to the same location as the paternally-expressed insulin-like growth factor II gene (*IGF2*) implicates *IGF2* as the main candidate gene.

Several beef breeds of cattle, including Belgian Blue, Charolais and Piedmontese, show a form of muscle hypertrophy often called double-muscling (FIG. 5). The phenotype is controlled by a recessive mutation at the *mh* locus on cattle chromosome 2 (REF. 44). *Myostatin* became the obvious candidate gene for this phenotype when it was shown that mice homozygous for a targeted deletion of this gene developed extreme muscularity⁴⁵. The *myostatin* gene product apparently acts as a negative regulator of skeletal muscle growth. Soon afterwards several groups were able to show that *myostatin* maps to the same region of cattle chromosome 2 as *mh*, and that double-muscled cattle (*mh/mh*) are homozygous for loss-of-function mutations in *myostatin*^{46–48}. Grobet *et al.*⁴⁹ reported five different loss-of-function alleles present in different breeds of beef cattle. Although the selection for muscularity has picked up several different alleles in cattle, no such allele has so far been reported in pig, despite a similarly strong selection pressure in this species. A possible explanation is that the prenatal overgrowth associated with these mutations in cattle is incompatible with the reproductive strategy of the pig, which produces 5–20 piglets in each litter.

The callipyge (*CLPG*) gene in sheep causes muscular hypertrophy. The phenotype was first observed in 1983 in a single ram, and the mutation most probably arose spontaneously in one of his close ancestors. *CLPG* was mapped to sheep chromosome 18 (REF. 50) and the phenotype was shown to have a peculiar inheritance, denoted polar overdominance⁵¹ — the phenotype is only expressed in heterozygotes, and only when the *CLPG* allele has been transmitted from the sire. The causative gene for this phenotype has not yet been reported but a possible explanation is that the phenotype is due to a defect in the control of genetic imprinting⁵¹.

Two cases have now been reported in which genetically imprinted loci affect body composition in farm animals — *CLPG* in sheep and the *IGF2*-linked QTL in pig. A recent study in pig indicates that this might be a common phenomenon, as four out of five detected QTL for body composition were reported to show genetic imprinting⁵². The result is surprising because only a minority of genes show genetic imprinting⁵³ and the observation needs to be verified.

Fertility traits. Fertility is a very important production trait but is difficult to study, partly because phenotypic records are primarily available on breeding animals. Furthermore, fertility traits have a low heritability because environmental factors such as health and nutritional status have a large impact.

Two different alleles at the X-linked *Inverdale* locus that affect ovulation rate in sheep were recently reported to be due to mutations in *BMP15*, which encodes a growth factor expressed in the oocyte⁵⁴. Heterozygotes for these mutations have a high ovulation rate, whereas the homozygotes are infertile. *Booroola* is another example of a gene with a large effect on female reproduction — it increases both ovulation rate and litter size. *Booroola* has been mapped to a region on sheep chromosome 3 that has conserved synteny with human chromosome 4 (REF. 55).

Polymorphism in an oestrogen receptor gene, a candidate gene for reproductive traits, has been reported to be associated with litter size in pigs⁵⁶. However, a causal relationship for the association remains to be shown. Several QTL for reproductive traits have been identified in resource populations generated by crossing European pig breeds and hyperprolific Chinese pigs^{57,58}.

Monogenic disorders. Several mutations that cause simple monogenic disorders in farm animals have been identified, and are catalogued in the Online Mendelian Inheritance In Animals (OMIA) database. Autosomal recessive disorders present a problem in animal breeding, because normal carriers can transmit the disease allele to a large number of progeny. A good example of this is **bovine leukocyte adhesion deficiency**, a severe immunodeficiency syndrome caused by a missense mutation in *ITGB2*, which encodes the integrin $\beta 2$ subunit⁵⁹. The mutation was widely spread in the US Holstein-Friesian cattle population and all affected animals were related to one famous bull with a huge number of progeny in the 1950s and 1960s. **Severe combined immunodeficiency in Arabian horses** is caused by a frameshift mutation in the gene for the catalytic subunit of DNA-dependent protein kinase (*PRKDC*)⁶⁰. In both these cases, as well as in most other disorders in which the molecular basis has been revealed, the identification of the causative gene took advantage of comparative information from the corresponding disorders in humans or mice. Diagnostic tests have been developed and are used by breeders to reduce the incidence of monogenic disorders. In some cases, such disorders will provide useful animal models for human disease and might be used to develop improved therapy. An interesting example of this is the report that the **narcolepsy** sleeping disorder in a breed of dogs is caused by a mutation in the gene that encodes the hypocretin (orexin) receptor 2 (REF. 61).

Markers for improved disease resistance. Selection for improved general disease resistance in farm animals is highly desirable for animal welfare, but also because reducing the losses due to disease is an excellent way of improving production efficiency. Selection for disease

resistance is difficult to include in traditional breeding programmes because of the lack of accurate records on disease incidence in the breeding herds and the strong environmental factors, such as the exposure to pathogens. As a result, there is a considerable interest in finding genetic markers or diagnostic tests that can be applied in breeding programmes. However, little progress has been made so far, primarily because of the difficulty of gaining access to pedigree data with informative disease records. A marked difference from human genetics is the lack of detailed clinical records. Animals that are highly susceptible to disease are rapidly eliminated from the commercial populations. Furthermore, the generation of dedicated resource pedigrees with high incidence of disease (for example, following a pathogen challenge experiment) is very expensive and might be unethical. This is an area where more research is needed and there is no doubt that the animal breeding industry will readily use the knowledge to reduce disease incidence. The collection of more accurate data on disease incidence and disease-related parameters as well as the exploitation of such data for genetic studies should be promoted.

Numerous studies on the relationship between genetic variation and disease resistance have focused on major histocompatibility complex genes (reviewed in REF. 62). An interesting attempt to identify genes for resistance to a major disease involves a cross between trypanotolerant N'Dama cattle from West Africa and susceptible African Zebu cattle⁶³. Trypanosomiasis is a major cattle disease, and susceptible cattle do not survive in areas infected with tsetse flies, the vector for trypanosomes. Genome scans for detecting QTL are underway, and the results might be used to improve tolerance in susceptible cattle.

Two major loci affecting susceptibility to *E. coli* infections in pigs have been identified. Susceptibility to adhesion of F18 FIMBRIATED *E. coli* (ECF18), which causes **oedema disease**, seems to be controlled by one or two missense mutations in *FUT1*, which encodes fucosyltransferase 1 (REF. 64). Adhesion of K88 fimbriated *E. coli* strains is controlled by a locus on chromosome 13, but the causative gene has not yet been reported⁶⁵.

Marek's disease (MD) is a lymphoproliferative disease caused by the MD virus in chicken and has a large impact on production. A QTL scan using a challenge test in a resource population revealed at least two QTL with significant effects on resistance to disease⁶⁶.

Future perspectives

Genome research in farm animals has already led to several important practical applications, such as the diagnostic tests for the *Halothane* (*RYR1*) and *RN* (*PRKAG3*) mutations that are widely used in pig breeding. However, it is important to note that the identification of the causative gene for a trait locus is not a prerequisite for practical applications. Several cattle and pig breeding companies are now using marker-assisted selection with markers flanking QTL as a complement to phenotypic selection of breeding animals. It is likely that large-scale marker analysis will be used routinely, as

FIMBRIATED

Fimbria are protein molecules (often called adhesins or colonizing factors) that allow the bacteria to adhere to receptors on host cells, such as enterocytes in the intestine.

soon as the cost for genotyping has been reduced by a factor of around ten.

It is obvious that behavioural traits, such as feeding, aggression/docility, stress tolerance and general activity, have been modified considerably during domestication and animal breeding. The various breeds of dog probably provide the most striking examples of how much behavioural traits have been changed. Compare, for instance, the strong tendency for aggressive behaviour in guard dogs (for example, Dobermann pinscher, Rottweiler and Mastiff) with those selected for other purposes (for example, Pointer, Golden retriever and Poodle). Behavioural traits are difficult to study because of the strong environmental component and because it is difficult to collect objective and informative records, in particular on the number of animals needed for high-resolution mapping. However, behavioural genetics in domestic animals is an exciting field for future research.

Linkage disequilibrium mapping will be a very powerful approach for mapping and finding trait loci in domestic animals once dense SNP maps become available and the cost for genotyping has been reduced such that genome scans using thousands of SNPs can be done. There is also a need for the development of improved statistical methods for analysing QTL data. For instance, the importance of epistatic interaction between QTL is well established in experimental organisms such as *Drosophila*³⁰. This has hardly been studied in outbred organisms, although some theoretical work indicates how this can be accomplished^{67,68}.

An increasing number of trait loci in farm animals have been characterized at the molecular level in recent years. The future prospects for cloning trait loci are bright, even for major QTL, provided that the QTL is due to one or more mutations in a single gene and not a haplotype effect. The reason for this optimism is the continuous development of better tools and new resources for genome research. Current initiatives to develop complete BAC contigs of farm animal genomes will provide researchers with a large-insert contig covering the region of interest as soon as a trait locus has been

mapped. Such large-insert contigs can then be used to build a preliminary transcript map of the region by high-resolution comparisons with the corresponding region in humans or mice.

It is also only a matter of time before initiatives will be taken to sequence the genomes of farm animals. This will most probably be carried out using a whole-genome shotgun approach giving 3–6-fold coverage. I would propose that this be carried out by using genomic DNA from two or more divergent populations, such as a beef and a dairy cattle breed, or from an improved breed and the wild ancestor (in those species in which the wild ancestor is not extinct). This will not reduce the efficiency in determining the genome sequence much but will detect a good proportion of fixed genetic differences between divergent populations. Bioinformatic analyses should provide scores for the likelihood that an observed substitution is functionally important, on the basis of the degree of phylogenetic conservation of the position and the possible functional consequences of the substitution (for example, a change of amino acid, or of a conserved promoter element). The functional consequences will then be evaluated by experimentation. By this approach, it should be possible to unravel the molecular basis for a variety of phenotypic traits of agricultural, biological and medical significance.

Links

DATABASE LINKS [Mc1r](#) | [KIT](#) | [dominant white coat colour in pigs](#) | [dominant white spotting in mice](#) | [piebald trait in humans](#) | [malignant hyperthermia in pigs](#) | [malignant hyperthermia in humans](#) | [myostatin](#) | [bovine leukocyte adhesion deficiency](#) | [severe combined immunodeficiency in Arabian horses](#) | [narcolepsy in dogs](#) | [oedema disease in pigs](#)

FURTHER INFORMATION [Roslin Institute](#) | [US Livestock Genome Mapping Projects](#) | [Human SNP Consortium](#) | [pig radiation hybrid maps](#) | [human transcript map](#) | [US Meat Animal Research Center](#) | [Online Mendelian Inheritance In Animals](#) | [dbEST summary by organism](#)

- Harpending, H. C. *et al.* Genetic traces of ancient demography. *Proc. Natl Acad. Sci. USA* **95**, 1961–1967 (1998).
- Loftus, R. T., MacHugh, D. E., Bradley, D. G., Sharp, P. M. & Cunningham, P. Evidence for two independent domestications of cattle. *Proc. Natl Acad. Sci. USA* **91**, 2757–2761 (1994).
A demonstration that cattle have been domesticated from two distinct subspecies and that many breeds of African cattle are hybrids between the two owing to male-derived gene flow.
- MacHugh, D. E., Shriver, M. D., Loftus, R. T., Cunningham, P. & Bradley, D. G. Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (*Bos taurus* and *Bos indicus*). *Genetics* **146**, 1071–1086 (1997).
- Giuffra, E. *et al.* The origin of the domestic pig: independent domestication and subsequent introgression. *Genetics* **154**, 1785–1791 (2000).
Evidence is presented for independent domestication of Wild Boar subspecies in Europe and Asia.
- Darwin, C. *On the origins of species by means of natural selection or the preservation of favored races in the struggle for life* (Murray, London, 1859).
- Hawken, R. J. *et al.* A first-generation porcine whole-genome radiation hybrid map. *Mamm. Genome* **10**, 824–830 (1999).
- Yang, Y. P. & Womack, J. E. Parallel radiation hybrid mapping: a powerful tool for high-resolution genomic comparison. *Genome Res.* **8**, 731–736 (1998).
- Kwok, C. *et al.* Characterization of whole genome radiation hybrid mapping resources for non-mammalian vertebrates. *Nucleic Acids Res.* **26**, 3562–3566 (1998).
- Band, M. R. *et al.* An ordered comparative map of the cattle and human genomes. *Genome Res.* **10**, 1359–1368 (2000).
- Lockhart, D. J. & Winzler, E. A. Genomics, gene expression and DNA arrays. *Nature* **405**, 827–836 (2000).
- Chowdhary, B. P., Raudsepp, T., Fronicke, L. & Scherthan, H. Emerging patterns of comparative genome organization in some mammalian species as revealed by Zoo-FISH. *Genome Res.* **8**, 577–589 (1998).
- Burt, D. W. *et al.* The dynamics of chromosome evolution in birds and mammals. *Nature* **402**, 411–413 (1999).
The analysis of comparative mapping data shows that the organization of the human genome is closer to that of the chicken than the mouse. This illustrates that the rate of chromosomal evolution varies considerably between species.
- Bush, G. L., Case, S. M., Wilson, A. C. & Patton, J. L. Rapid speciation and chromosomal evolution in mammals. *Proc. Natl Acad. Sci. USA* **74**, 3942–3946 (1977).
- Johansson, M., Ellegren, H. & Andersson, L. Comparative mapping reveals extensive linkage conservation—but with gene order rearrangements—between the pig and the human genomes. *Genomics* **25**, 682–690 (1995).
- Georges, M. *et al.* Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. *Genetics* **139**, 907–920 (1995).
- Zhang, Q. *et al.* Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. *Genetics* **149**, 1959–1973 (1998).
- Heyen, D. W. *et al.* A genome scan for QTL influencing milk production and health traits in dairy cattle. *Physiol. Genomics* **1**, 165–175 (1999).
- Charlier, C. *et al.* Identity-by-descent mapping of recessive traits in livestock: application to map the bovine syndactyly locus to chromosome 15. *Genome Res.* **6**, 580–589 (1996).
- Riquet, J. *et al.* Fine-mapping of quantitative trait loci by identity by descent in outbred populations: application to milk production in dairy cattle. *Proc. Natl Acad. Sci. USA* **96**, 9252–9257 (1999).
This paper shows how the identity-by-descent approach can be applied to high-resolution mapping of quantitative trait loci in farm animals by using extensive pedigree information.

20. Farnir, F. *et al.* Extensive genome-wide linkage disequilibrium in cattle. *Genome Res.* **10**, 220–227 (2000). **The pattern of linkage disequilibrium (LD) across the genome in dairy cattle is evaluated. The strong LD detected holds promise that LD mapping will be a powerful strategy for mapping quantitative trait loci in farm animals.**
21. Andersson, L. *et al.* Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science* **263**, 1771–1774 (1994). **The first paper to show the use of divergent intercrosses for mapping quantitative trait loci in outbred populations.**
22. Rohrer, G. A. & Keele, J. W. Identification of quantitative trait loci affecting carcass composition in swine: I. Fat deposition traits. *J. Anim. Sci.* **76**, 2247–2254 (1998).
23. de Koning, D. J. *et al.* Detection of quantitative trait loci for backfat thickness and intramuscular fat content in pigs (*Sus scrofa*). *Genetics* **152**, 1679–1690 (1999).
24. Breneman, R. A. *et al.* The polled locus maps to BTA1 in a *Bos indicus* × *Bos taurus* cross. *J. Hered.* **87**, 156–161 (1996).
25. Walling, G. A. *et al.* Combined analyses of data from quantitative trait loci mapping studies. Chromosome 4 effects on porcine growth and fatness. *Genetics* **155**, 1369–1378 (2000).
26. Darvasi, A. & Soller, M. Advanced intercross lines, an experimental population for fine genetic-mapping. *Genetics* **141**, 1199–1207 (1995).
27. Darvasi, A. Experimental strategies for the genetic dissection of complex traits in animal models. *Nature Genet.* **18**, 19–24 (1998).
28. Milan, D. *et al.* A mutation in *PRKAG3* associated with excess glycogen content in pig skeletal muscle. *Science* **288**, 1248–1251 (2000). **The first positional cloning of a trait locus in a farm animal. Linkage mapping, linkage disequilibrium mapping, radiation hybrid mapping, construction of a BAC contig, BAC sequencing and bioinformatic analysis eventually resulted in the identification of the causative missense mutation.**
29. Schibler, L., Cribiu, E. P., Oustry-Vaiman, A., Furet, J. P. & Vaiman, D. Fine mapping suggests that the goat Polled Intersex Syndrome and the human Blepharophimosis Ptosis Epicanthus Syndrome map to a 100-kb homologous region. *Genome Res.* **10**, 311–318 (2000).
30. Mackay, T. F. C. Quantitative trait loci in *Drosophila*. *Nature Rev. Genet.* **2**, 11–21 (2001).
31. Robbins, L. S. *et al.* Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. *Cell* **72**, 827–834 (1993).
32. Klungland, H., Vage, D. I., Gomez-Raya, L., Adalsteinsson, S. & Lien, S. The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination. *Mamm. Genome* **6**, 636–639 (1995).
33. Marklund, L., Moller, M. J., Sandberg, K. & Andersson, L. A missense mutation in the gene for melanocyte-stimulating hormone receptor (*MCT1R*) is associated with the chestnut coat color in horses. *Mamm. Genome* **7**, 895–899 (1996).
34. Kijas, J. M. H. *et al.* Melanocortin receptor 1 (*MCT1R*) mutations and coat color in pigs. *Genetics* **150**, 1177–1185 (1998).
35. Vage, D. I., Klungland, H., Lu, D. & Cone, R. D. Molecular and pharmacological characterization of dominant black coat color in sheep. *Mamm. Genome* **10**, 39–43 (1999).
36. Newton, J. M. *et al.* Melanocortin 1 receptor variation in the domestic dog. *Mamm. Genome* **11**, 24–30 (2000).
37. Takeuchi, S., Suzuki, H., Yabuuchi, M. & Takahashi, S. A possible involvement of melanocortin 1-receptor in regulating feather color pigmentation in the chicken. *Biochim. Biophys. Acta* **1308**, 164–168 (1996).
38. 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**, 822–830 (1996).
39. Marklund, S. *et al.* Molecular basis for the dominant white phenotype in the domestic pig. *Genome Res.* **8**, 826–833 (1998).
40. Fujii, J. *et al.* Identification of a mutation in the porcine ryanodine receptor that is associated with malignant hyperthermia. *Science* **253**, 448–451 (1991). **The first molecular description of a major trait locus in farm animals.**
41. MacLennan, D. H. & Phillips, M. S. Malignant hyperthermia. *Science* **256**, 789–794 (1992).
42. Jeon, J. -T. *et al.* A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the *IGF2* locus. *Nature Genet.* **21**, 157–158 (1999).
43. Nezer, C. *et al.* An imprinted QTL with major effect on muscle mass and fat deposition maps to the *IGF2* locus in pigs. *Nature Genet.* **21**, 155–156 (1999).
44. Charlier, C. *et al.* The *mh* gene causing double-muscling in cattle maps to bovine chromosome 2. *Mamm. Genome* **6**, 788–792 (1995).
45. McPherron, A. C., Lawler, A. M. & Lee, S. J. Regulation of skeletal muscle mass in mice by a new TGF- β superfamily member. *Nature* **387**, 83–90 (1997).
46. Grobet, L. *et al.* A deletion in the bovine myostatin gene causes the double-muscling phenotype in cattle. *Nature Genet.* **17**, 71–74 (1997).
47. Kambadur, R., Sharma, M., Smith, T. P. & Bass, J. J. Mutations in myostatin (*GDF8*) in double-muscling Belgian Blue and Piedmontese cattle. *Genome Res.* **7**, 910–916 (1997).
48. McPherron, A. C. & Lee, S. J. Double muscling in cattle due to mutations in the myostatin gene. *Proc. Natl Acad. Sci. USA* **94**, 12457–12461 (1997).
49. Grobet, L. *et al.* Molecular definition of an allelic series of mutations disrupting the myostatin function and causing double-muscling in cattle. *Mamm. Genome* **9**, 210–213 (1998). **References 46–49 show that selection for muscularity in beef cattle has increased the frequency of at least five independent loss-of-function mutations in myostatin.**
50. Cockett, N. E. *et al.* Chromosomal localization of the callipyge gene in sheep (*Ovis aries*) using bovine DNA markers. *Proc. Natl Acad. Sci. USA* **91**, 3019–3023 (1994).
51. Cockett, N. E. *et al.* Polar overdominance at the ovine callipyge locus. *Science* **273**, 236–238 (1996).
52. de Koning, D. J. *et al.* Genome-wide scan for body composition in pigs reveals important role of imprinting. *Proc. Natl Acad. Sci. USA* **97**, 7947–7950 (2000).
53. Reik, W. & Walter, J. Genomic imprinting: parental influence on the genome. *Nature Rev. Genet.* **2**, 21–32 (2001).
54. Galloway, S. M. *et al.* Mutations in an oocyte-derived growth factor gene (*BMP15*) cause increased ovulation rate and infertility in a dosage-sensitive manner. *Nature Genet.* **25**, 279–283 (2000). **Positional candidate cloning revealed two different causative mutations in *BMP15* affecting fertility in sheep.**
55. Montgomery, G. W. *et al.* The Booroola fecundity (*FecB*) gene maps to sheep chromosome 6. *Genomics* **22**, 148–153 (1994).
56. Rothschild, M. *et al.* The estrogen receptor locus is associated with a major gene influencing litter size in pigs. *Proc. Natl Acad. Sci. USA* **93**, 201–205 (1996).
57. Rohrer, G. A., Ford, J. J., Wise, T. H., Vallet, J. L. & Christenson, R. K. Identification of quantitative trait loci affecting female reproductive traits in a multigeneration Meishan-White composite swine population. *J. Anim. Sci.* **77**, 1385–1391 (1999).
58. Wilkie, P. J. *et al.* A genomic scan of porcine reproductive traits reveals possible quantitative trait loci (QTLs) for number of corpora lutea. *Mamm. Genome* **10**, 573–578 (1999).
59. Shuster, D. E., Kehrl, M. E., Ackermann, M. R. & Gilbert, R. O. Identification and prevalence of a genetic defect that causes leukocyte adhesion deficiency in Holstein cattle. *Proc. Natl Acad. Sci. USA* **89**, 9225–9229 (1992).
60. Shin, E. K., Perryman, L. E. & Meek, K. A kinase-negative mutation of DNA-PKCS in equine SCID results in defective coding and signal joint formation. *J. Immunol.* **158**, 3565–3569 (1997).
61. Lin, L. *et al.* The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* **98**, 365–376 (1999). **This paper shows how a well-established animal model for a human disorder was used to identify an autosomal recessive mutation responsible for the disorder.**
62. Schook, L. B. & Lamont, S. J. *The major histocompatibility complex region of domestic animal species* (CRC Press Inc., Boca Raton, 1996).
63. Teale, A. *et al.* Genetics of resistance to trypanosomiasis in mice and livestock. *Anim. Genet. (Suppl. 2)* **27**, 5 (1996).
64. Meijerink, E. *et al.* Two α (1,2) fucosyltransferase genes on porcine chromosome 6q11 are closely linked to the blood group inhibitor (S) and *Escherichia coli* F18 receptor (ECF18R) loci. *Mamm. Genome* **8**, 736–741 (1997).
65. Edfors-Lilja, I. *et al.* The porcine intestinal receptor for *Escherichia coli* K88ab, K88ac: regional localization on chromosome 13 and influence of IgG response to the K88 antigen. *Anim. Genet.* **26**, 237–242 (1995).
66. Vallejo, R. L. *et al.* Genetic mapping of quantitative trait loci affecting susceptibility to Marek's disease virus induced tumors in F_2 intercross chickens. *Genetics* **148**, 349–360 (1998).
67. Kao, C.-H., Zeng, Z.-B. & Teasdale, R. Multiple interval mapping for quantitative trait loci. *Genetics* **152**, 1203–1216 (1999).
68. Carlborg, O., Andersson, L. & Kinghorn, B. The use of a genetic algorithm for simultaneous mapping of multiple interacting quantitative trait loci. *Genetics* **155**, 2003–2010 (2000).
69. Geldermann, H. Investigations of quantitative characters in animals by gene markers. I. Methods. *Theor. Appl. Genet.* **46**, 319–330 (1975).
70. Weller, J., Kashi, Y. & Soller, M. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *J. Dairy Sci.* **73**, 2525–2537 (1990).
71. Seltz, J. J., Schmutz, S. M., Thue, T. D. & Buchanan, F. C. A missense mutation in the bovine *MGF* gene is associated with the roan phenotype in Belgian Blue and Shorthorn cattle. *Mamm. Genome* **10**, 710–712 (1999).

Acknowledgements

Sincere thanks are due to Erik Bongcam-Rudloff for expert assistance in preparing the illustrations. Work in the author's laboratory is primarily supported by the Swedish Research Council for Forestry and Agriculture.