Reproductive Genomics in Domestic Animals

COPYRIGHTED MATERIAL

Part I

Quantitative Genomics of Reproduction

1

Reproductive Genomics: Genome, Transcriptome, and Proteome Resources

Noelle E. Cockett

1.1 Introduction

Genomic resources, tools, and technologies that can be applied to studies in livestock species, including investigations related to reproduction, have been under development for the last decade. While many of the genomic approaches were originally developed for use in humans or laboratory model animals, they have been successfully applied to studies in livestock. There are now a myriad of resources specific to livestock species, such as well-characterized genome maps, high-resolution genome, and complementary DNA (cDNA) sequences, expression arrays, and high-density genetic marker chips. In addition, there is an explosion of high-throughput technology that will enhance these investigations, increasing the scope and accuracy of the results beyond anything that was imagined just 5 years ago. These technologies advance studies of single gene expression to full gene networks, from single gene sequences to whole genomes, and from hundreds of genetic markers to tens of thousands markers—all assayable in a few weeks to months as opposed to years.

These resources and technologies can be combined in innovative ways to advance two areas of research on reproductive traits, specifically the identification of genes or genetic regions influencing phenotypes and the characterization of expression of genes that are associated with traits.

1.2 Discovery of underlying genetic influences

The first area of interest for researchers studying reproductive traits is the characterization of genetic variation among animals or populations underlying a phenotypic trait, leading to the identification of the genetic cause of the phenotype. Two general approaches have been successfully used over the last 10–15 years, with a third approach now on the horizon. In the first approach,

polymorphisms in a candidate gene likely to be involved in the phenotype are tested for associations with different manifestations or phenotypes of the trait. The candidate genes are selected for analysis based on an understanding of trait physiology and/or because of their involvement in similar traits in other species. In the second approach, genetic markers are analyzed for linkage with the phenotype using pedigrees of animals segregating for the trait and the markers. This analysis identifies genetic regions that contain associated genes. By testing additional markers through the families, the interval is narrowed so candidate genes can be selected. The third approach, referred to as whole genome associations, will soon be possible for livestock species now that the development of high-density single nucleotide polymorphism (SNP) arrays are readily available. However, the application of whole genome associations requires very large numbers of phenotyped animals, which is a limitation for most research projects.

1.2.1 Candidate gene associations

As mentioned, the candidate gene approach uses information of the trait to determine likely candidates for the underlying gene(s). The choice of the gene is strengthened by its involvement in comparable traits in other species or its location in a region previously identified as containing a quantitative trait loci (QTL) with similar attributes.

In the past, polymorphisms in a candidate gene were routinely detected by polymerase chain reaction—restriction fragment length polymorphisms (PCR-RFLP), which involves steps of amplifying the gene, digesting the amplicon with a restriction enzyme, and then using gel electrophoresis to separate the resulting fragments. In the PCR-RFLP technique, gene sequence differences among animals are detected by whether or not a restriction enzyme cuts, resulting in different-sized fragments. The genetic differences are usually due to an SNP within the restriction enzyme recognition site, although there might be genetic differences due to insertions/deletions (in/del) in the gene, which will also result in fragment size differences, although there is no variation in the restriction enzyme recognition site. Animals are expected to have two alleles for every gene except those on the X and Y chromosomes in males, so that the presence of one fragment on the electrophoresis gel would indicate that an animal is homozygous for the PCR-RFLP allele whereas the presence of two different-sized fragments would suggest that an animal is heterozygous. However, an animal might be misclassified as a homozygote if there is a polymorphism in the PCR primer sequence, which prevents that allele from being amplified and therefore, not detected on the electrophoresis gel—referred to as a "null" allele. A null allele will often be detected when misparentages are routinely found for a marker system. An animal might also be misclassified if another, nonallelic form of the gene is amplified with the PCR primers and digestion with the restriction enzyme results in a different-sized fragment. A nonallelic form is revealed by sequencing the fragments contained within the electrophoretic bands, which is a recommended step when establishing any marker system.

However, new technologies have significantly advanced our ability to identify SNPs and then explore multiple candidate genes at one time at a much lower cost/polymorphism than the PCR-RFLP method. The identification of SNPs within a gene or genetic region is now relatively easy. To do this, the genomic DNA of key animals within a population is sequenced using high-

throughput automatic sequencing and then compared with other sequences within the population or to sequences in publically available databases. The later approach is referred to as *in silico* SNP detection. Regardless of the approach, confidence of the SNP is dependent on the quality of the sequence across the multiple sources of data.

Once an SNP is identified, the polymorphism can be detected by establishing a PCR-RFLP assay. However, allele-specific PCR using allele-specific oligonucleotides (ASOs) is an emerging technique for detecting genetic variation created by the SNP (Saiki et al. 1986). The 3′ ends of the primers used in the PCR amplification step of the ASO technique are designed to include the polymorphic site so that amplification of the animal's DNA is dependent on the absence or presence of the polymorphism within the primer sequence. Allele-specific primers can be combined into a single amplification reaction and the presence of the specific allele detected by the melting temperature of the alleles (Papp et al. 2003; Wang et al. 2005). Appropriate controls and design of the primers (e.g., Strerath et al. 2007) are critical in the allele-specific amplification assay so that absence of amplification is due to the polymorphism and not because of technical problems.

SNP arrays are an extension of the ASO method, but by spotting multiple ASOs onto a membrane or bead, multiple alleles or even multiple genetic markers can be assayed in a single run. Custom-built SNP chips specific to a trait are usually designed in a 92-, 384-, or 1534-SNP format. While the cost/SNP is lower for the SNP chip than with the PCR-RFLP or allele-specific amplification techniques, the initial setup for the chip is substantially higher. Thus, the number of SNPs that are tested and the number of animals included in the analysis will determine whether a custom-built SNP array is economical.

Emerging technology is now allowing the detection of differences in copy number variant (CNV) among animals. For some time, copy number variation has been associated with diseases (McCarroll 2008; Schaschi et al. 2009), while the ongoing analyses of livestock whole genome sequences has revealed the presence of CNV in multiple gene systems involved with innate immunity, including milk composition traits (Rijnkels et al. 2009; Tellam and Bovine Genome Sequencing and Analysis Consortium 2009). Detection of differences among animals for genes that are known to be present in the genome in multiple copies is now possible using microarray technology (Baumbusch et al. 2008), with higher copy number resulting in greater intensity for that spot on the array.

Once the polymorphism is detected within a population, the genotypes are usually analyzed for association with the trait by comparing the trait means among the marker genotypes (Rocha et al. 1992). Appropriate statistical models are needed in order to account for additive, dominant, and epistatic effects. In addition, the selection of animals used in the analysis must be sufficiently broad; otherwise, the marker alleles will merely serve as a trace of unique families, particularly when one of the alleles is at a very low frequency in the population and present only in one family in the analysis. This situation can result in a spurious significant association, simply because the family differs for the trait and not because the allele itself is associated.

The choice of the candidate gene(s) can be strengthened by its association with similar traits in the same or other species. Possible candidate genes can be found through literature searches using key words based on

Table 1.1 Websites containing genomics information in livestock species.

Species	Website	Information
Cattle	www.animalgenome.org/QTLdb/cattle.html	QTL
	www.vetsci.usyd.edu.au/reprogen/QTL_Map/	QTL
	www.hgsc.bcm.tmc.edu/projects/bovine/	Genome sequence
	bovinegenome.org	Genome project
Goat	dga.jouy.inra.fr/cgi-bin/lgbc/main.pl?BASE=goat	Genome project
Horse	www.uky.edu/Ag/Horsemap/welcome.html	Genome project
	www.broad.mit.edu/mammals/horse	Genome sequence
Sheep	rubens.its.unimelb.edu.au/~jillm/jill.htm	Primary Web source
	www.livestockgenomics.csiro.au/perl/gbrowse.cgi/vsheep2/	Virtual sheep genome
	www.ncbi.nlm.nih.gov/genome/guide/sheep/index.html	NCBI resources
	www.sheephapmap.org/	International Sheep Genome Consortium
Pig	www.animalgenome.org/QTLdb/pig.html	QTL
	www.sanger.ac.uk/Projects/S_scrofa/	Genome sequence
	www.piggenome.org/index.php	Genome project

the trait physiology or through searches of databases devoted to genetic abnormalities. One such database for livestock traits is called Online Mendelian Inheritance in Animals (OMIA; www.omia.angis.org.au/). The OMIA database contains details on genes, inherited disorders, and traits for a large range of animals species, similar to what is found within Online Mendelian in Man (OMIM; www.ncbi.nlm.nih.gov/sites/entrez?db=omim).

There are also databases that describe the location of QTLs for traits of interest in livestock species (Table 1.1). Additional candidate genes can be identified by searching genetic sequences that lie within QTL intervals and have involvement in the physiology of the trait, providing not only functional evidence but also positional evidence for inclusion in the candidate gene analysis. These genes are therefore referred to as "positional candidate genes" (see below).

1.2.2 Analysis of genetic variation

The second approach for detecting genes or, more commonly, genetic regions involved in traits is based on identifying and characterizing genetic variation that is found in pedigrees of animals. This approach has most commonly been done using linkage analysis, which examines the segregation of marker alleles through animal families with known phenotypes (Nejati-Javaremi and Smith 1995; Knott and Haley 2000; de Koning et al. 2003) and subsequent refinement of the genetic interval containing the trait locus (Riquet et al. 1999; Farnir et al. 2002). The data are analyzed to determine the coinheritance of marker alleles with the causative genetic mutation, presumably because they are closely located within the genome.

Linkage mapping requires pedigrees with specific family structures; these pedigrees are most commonly reciprocal backcrosses or F2 crosses developed from lines or breeds of animals that significantly differ for the trait. The analysis can include families within a single breed or line but the key parents must be heterozygous for both the markers and the trait in order for linkage to be detected. As with the association analyses, appropriate statistical models are needed

to detect genetic mutations that are controlled by complex gene actions, such as the imprinted callipyge (Cockett et al. 1994, 1996) and IGF2 (Van Laere et al. 2003) loci. The effects of these loci would not have been detected without the appropriate statistical model (see Sandor and Georges 2008).

To perform a screen of markers across the complete genome (i.e., a genome scan), markers are typically selected about one every 10–20 centimorgans (cM). Because the typical mammalian genome is about 3000 cM, around 150–300 markers are needed for a genome scan. The availability of genome-wide maps in livestock species provides the information needed to select markers at appropriate intervals, which is dependent on the number of informative offspring in the families and the genetic variability in the trait. Several reviews on conducting a genome scan and subsequent analysis are available, including Schwerin (2001), Rocha et al. (2002), Andersson and Georges (2004), and Georges (2007).

1.2.3 Whole genome sequence

Genetic markers for a genome scan are usually selected from a genome map. The most complete genome map for a species is produced from a whole genome sequence that has been assembled and annotated. Assembled whole genome sequences are now publically available for cattle, swine, and horses (Table 1.2). Millions of bases of sequences can be accessed for the analysis of genes, SNPs, regulatory features, and so on. Comparisons across species, including non-livestock species, are now possible using "landmark" loci that anchor segments of the genome from species to species. International consortiums of experts have been organized for annotation of the sequences; for example, "the Horse Genome Project is a cooperative international effort by over 100 scientists in 20 countries to define the genome, the DNA sequence, of the domestic horse" (www.uky.edu/Ag/Horsemap/welcome.html). A wealth of knowledge from the analysis of these sequences is now being released.

In addition, assembled whole genome sequences can serve as the "reference" for comparison of individual animal sequences generated with state-of-the-art high-throughput platforms such as ABI's SOLiD™ (Carlsbad, CA), Roche 454 FLX Titanium™ (Branford, CT), and Illumina's Solexa™ (San Diego, CA) systems. These technologies produce millions of reads of short sequences (50–400 bases) in a single run at relatively

Table 1.2 Whole genome sequence assemblies in livestock species.[1]

Species	Reference or website	Sequenced animal	Method	Coverage
Cattle[2]	Baylor School of Medicine	Hereford male L1 Domino 99375	WGS	7.1X
Horse[3]	Broad Institute	Thoroughbred female Twilight	WGS	6.8X
Pig[4]	Sanger Institute	Duroc female	BAC by BAC tile path	4X

[1] As of February 1, 2009.
[2] www.hgsc.bcm.tmc.edu/projects/bovine/.
[3] www.broad.mit.edu/node/318.
[4] www.sanger.ac.uk/Projects/S_scrofa/.
WGS, whole genome shotgun.

Table 1.3 Most recent published linkage maps in livestock species that do not have a whole genome sequence.

Species	Population	No. of loci	Reference
Goat	INRA	307	Schibler et al. (1998a)
Deer	Red deer × Pere David's deer	621	Slate et al. (2002)
Sheep	IMF	1062	Maddox et al. (2001)

low cost ($10,000/10 Gb) from either single or pooled DNA samples. The sequences for each run can then be compared back to the reference genome sequences, allowing detection of genetic differences across animals.

1.2.4 Linkage or genetic maps

For those species without a whole genome sequence, linkage and physical maps are critical for the orientation of loci as well as comparisons across species. Linkage maps usually contain a preponderance of highly polymorphic anonymous markers, primarily microsatellites, and relatively few expressed genes, which have very limited genetic variability. Also, multiple linkage maps may exist for a species because different reference families were used to create the linkage maps. The various maps are often combined into a consensus linkage map, which is anchored by common markers genotyped in the different reference families (Table 1.3). The distances between loci on linkage maps are given in centimorgans (cM), with 1 cM representing 1% recombination between two loci.

1.2.5 Physical map assignments

In addition to the linkage maps, physical maps exist for each species. These maps are created by direct assignment of a gene or marker to an intact chromosome or chromosomal fragment. Physical mapping is usually done by *in situ* hybridization, somatic cell hybrid analysis, or radiation hybrid (RH) mapping. Because a genetic variant within the locus is not necessary for physical mapping, these maps contain a relatively large number of expressed genes.

Table 1.4 Physical map in livestock species that do not have a whole genome sequence.

Species	No. of loci	Reference
River buffalo	388	Di Meo et al. (2008)
Goat	202	Schibler et al. (1998a)
Deer	59	Bonnet et al. (2001)
Sheep	452	Di Meo et al. (2007)

One of the first reports assigning genes to physical locations within the genome was performed by hybridizing a radioactively labeled probe to a spread of metaphase chromosomes in a technique referred to as *in situ* hybridization. A significant adaptation of this method entailed labeling the probe with fluorophores, leading to the moniker of fluorescent *in situ* hybridization (FISH). Hundreds of genes and genetic markers have now been assigned to specific chromosomes in livestock species using *in situ* hybridization techniques (Table 1.4).

Chromosome painting is an approach for evaluating the conservation of chromosomal segments across species. In this technique, chromosomes of one species are fluorescently labeled and hybridized to metaphase

chromosomes of another species. Reciprocal chromosome painting has been performed between humans and farm animal species including pigs, cattle, sheep, and horses (Chowdhary et al. 1996; Chowdhary and Raudsepp 2001). These studies have defined the borders of conserved syntenies among the species, but because of insufficient resolution, they do not allow the study of gene order.

Somatic cell hybrid panels were used frequently in the 1970s–1990s to assign genes to specific chromosomes in livestock, but this method is now used much less frequently than other physical mapping approaches that have a much better resolution. A somatic cell hybrid panel is generated by fusing cells of the target species with cells of a rodent species, such as hamsters. The rodent cells randomly eject chromosomes of the target species, until at some point the cells are immortalized, leaving a complement of target chromosomes that is the signature of each somatic cell clone. DNA is harvested from each of the clones in the panel (usually around 30 clones) and then amplified with primers specific to a gene or genetic marker. Those clones that contain a piece of the chromosome harboring the gene amplify with the primers, as well as other "concordant" or linked genes. The somatic cell hybrid approach can be used to identify genes found within long segments of the chromosome, but the order of the genes along the chromosomal segment cannot be determined.

RH mapping provides a higher level of resolution of gene location and gene order than those produced by *in situ* hybridization and somatic cell hybrids. This technique is based on detecting the presence/absence of loci that are contained on fragments of DNA maintained in a panel of hybrid clones, similar to the somatic cell hybrid approach, but the rodent cells are fused with target cells that have been irradiated. Varying the radiation dose on the target species cells will create different-sized fragments and therefore, vary the resolution between two loci. The higher the radiation dose, the smaller the fragments and the better resolution between tightly linked loci. Thus, high rad panels are suitable when fine-mapping markers within a specific region, but large numbers of random markers must be screened in order to detect linkage of loci across the genome. Whole genome maps, which do not require a saturation of markers, are best done with a lower rad RH panel. RH panels have been generated for several livestock species (Table 1.5) and used for generating chromosome and whole genome RH maps.

Distances on an RH map are measured in centiRay (cR), with a distance of 1 cR_{rad} between two markers corresponding to a 1% frequency of breakage between these two markers after exposure to a specific radiation (rad) dose. Statistical programs have been developed to analyze the RH panel data to give the most likely order based on the least number of break points (e.g., Boehnke

Table 1.5 Radiation hybrid maps in livestock species that do not have a whole genome sequence.

Species	RH panel	Rad	No. of loci[1]	Reference
River buffalo	BBURH5000	5,000	3,990	Amaral et al. (2009)
Sheep	USUoRH5000	5,000	2,300	Wu et al. (2007, 2008, 2009)
	INRA	12,000	67	Laurent et al. (2007)

[1] As of February 1, 2009.

1992; Lange et al. 1995; Lunetta et al. 1996). A measure of relative likelihood of one order versus another is given for each map developed with the RH data.

Distance between loci on an RH map is directly proportional to physical distance, measured as the frequency of retention of a given pair of markers. The more times two loci are retained together, the closer they are found on a chromosome. Retention frequency is calculated as the percentage of clones that retain a given marker and is usually between 18% and 30% for whole genome RH panels.

1.2.6 Positional candidate genes

Once the location of a trait within the genome is determined because of linkage to previously mapped genetic markers, possible candidate genes controlling the trait can be inferred because of their proximity to the linked markers. A typical genome scan usually assigns the trait locus or QTL to a ~20-cM interval, which can contain hundreds of genes. However, it is not necessary to have map locations of all possible genes in a single livestock species. Rather, a subset of genes that are mapped in well-studied species, such as humans and mice, are also mapped in farm animals; these genes serve as "anchors" across the comparative maps and allow inference of the locations of other genes within a region, based on what is known within the well-mapped species (Burt 2002).

Positional candidate genes can be identified for traits mapped by linkage analysis once markers used in the linkage analysis are located on the comparative map, either by direct mapping or because a gene linked to the marker is placed on the comparative map. Several online comparative map databases have been established, which allow comparisons of genes contained within common genetic regions. These comparative maps can also be used to localize a single gene across multiple species.

Numerous causative mutations for single gene traits in livestock have been identified. In contrast, although numerous QTL have been identified for economically important traits in livestock (see Table 1.1), very few of the causative mutations for QTL (referred to as quantitative trait nucleotides or QTNs) have been characterized. There are numerous challenges in identifying the mutation for a quantitative trait, including a limitation on animals and/or families suitable for narrowing the QTL interval, an often unwieldy number of candidate genes and mutations within the genetic region, difficulty in estimating the interactions of other QTL on the trait, and technological and biological limitations when establishing the functionality of the candidate mutations. However, step-by-step approaches for establishing the causality of mutations involved in QTL have been proposed (Grisart et al. 2001, 2004; Andersson and Georges 2004; de Koning et al. 2007; Georges 2007; Ron and Weller 2007; Sellner et al. 2007).

1.2.7 Analysis of genetic fragments

Several large insert libraries, including bacterial artificial chromosome (BAC), yeast artificial chromosome (YAC), and fosmid vectors, exist for each livestock species, with the vast majority being BAC libraries (Table 1.6). Most of the BAC libraries for livestock species have been prepared by Pieter de Jong's group at BACPAC Resources Center (bacpac.chori.org/) and contain inserts with an average size of 90–200 Mb. These libraries can be screened by PCR amplification of plate, row, and column pools or by probe hybridization of high-density filters. The

Table 1.6 Large insert libraries in livestock species.

Species	Library	Genome coverage	Reference for BAC map
Cattle	CHORI-240	10.7X	Snelling et al. (2007)
	RPCI-42	10X	
Horse	CHORI-241	11.8X	Gustafson et al. (2003)
Sheep	CHORI-243	12X	Dalrymple et al. (2007)
Goat	6:15 translocation	3.3X	Schibler et al. (1998b)
Pig	CHORI-242	11.4X	Humphray et al. (2007)
	RPCI-44	10.2X	

Table 1.7 High-density SNP chips in livestock species.

Species	SNP chip	Reference
Cattle	Affymetrix 25K GeneChip	Khatkar et al. (2007)
	Illumina 50K BeadChip	Van Tassell et al. (2008)
Horse	Illumina 50K BeadChip	Chowdhary and Raudsepp (2008)
Sheep	Illumina 50K BeadChip	Kijas J. et al. (unpublished)
Pig	Illumina 60K BeadChip	Schook L. et al. (unpublished)

screening provides an exact clone address that contains the DNA sequence or gene of interest, and the clones can be purchased for about $20 from the BACPAC Resources Center.

Large overlapping segments of DNA called contigs can be generated by chromosome walking. To do this, the ends of isolated clones are sequenced, and then primers designed from the new sequence and used to screen the library in subsequent rounds. Sequential screenings will provide overlapping clones that can be pieced together into a single continuous fragment.

1.2.8 Whole genome association

In contrast to genetic linkage methods that use pedigrees segregating for the trait, genome-wide association (GWA) mapping is an approach that tests for allelic association at the population level through an analysis of linkage disequilibrium (LD), using large samples of unrelated individuals (Meuwissen et al. 2001; Amos 2007; McCarthy et al. 2008). Assuming that the trait allele of interest has descended from one or a few ancestral chromosomes, animals displaying the trait of interest will possess the ancestral haplotype that contains the allele of interest surrounded by closely linked marker alleles. These haplotypes may be fixed in a population or breed because of natural or artificial selection for the favorable allele, which is known as a "selective sweep" (Kim and Nielsen 2004; Pollinger et al. 2005; Voight et al. 2006; McVean 2007). The trait haplotype will be identified from all the wild-type haplotypes through the GWA analysis. Because very large numbers of markers are used in the analysis (around one marker every 100–500 kb), the resulting LD maps are typically of higher resolution than genetic linkage scans, which helps to limit the size of the interval that contains positional candidate genes.

High-density SNP chips are now under construction for all major livestock species (Table 1.7). These chips usually include between 30,000 and 60,000 SNPs, suitable for

large-scale genotyping applications such as the GWA analysis. The availability of whole genome high-density SNP chips at a relatively low price per animal ($150–$300) means that GWA analyses in livestock become a more common approach for localizing a trait locus within the genome. An application of GWA has recently been illustrated by the fine mapping of five recessive disorders in cattle using the high-density bovine BeadChip (Charlier et al. 2008) with less than 20 affected animals and 20 controls. However, the number of unrelated animals needed for mapping quantitative traits is predicted to be greater than 1000 (McCarthy et al. 2008; Orr and Chanock 2008; Tian et al. 2008). Unfortunately, very few livestock populations of that size currently exist.

1.3 Characterization of gene expression

Researchers are often interested in characterizing the expression of genes in a specific tissue at a specific time or under a specific set of circumstances. The expression of these genes can be "captured" by examining the messenger RNA (mRNA) within the tissue sample. The abundance of a particular mRNA within a tissue can now be measured with relative ease using techniques developed within the last decade, such as Northern blots, and newly developed techniques, such as real-time PCR. It is also possible to determine a "profile" of mRNAs within a tissue using an analysis system such as expression microarrays or serial analysis of gene expression (SAGE).

1.3.1 Synthesis and analysis of cDNA

The array of mRNAs found within a tissue is often converted into a cDNA library, which can then be examined in a variety of ways. Because RNases, enzymes that break down RNA, are quite ubiquitous and difficult to degrade, care must be taken to preserve the tissue without delay after collection, such as snap freezing the tissue in liquid nitrogen or immersing the tissue in a preservative/RNase inhibitor such as RNALater™ (QIAGEN, Inc., Valencia, CA). After extracting the RNA from the tissue, another enzyme, reverse transcriptase, converts the RNA into the first strand of cDNA followed by second strand synthesis using DNA polymerase. The cDNA does not directly correspond to genomic DNA because intronic segments have been spliced out when the RNA molecule was produced and only the exonic sequences are contained with the cDNA strand. The cDNA mixture can be analyzed for transcript content using various techniques or used in the creation of a cDNA library by cloning into vectors, usually plasmids, which are then transformed into competent *Escherichia coli* cells. Replication of the host cells in the cDNA library results in the replication of the plasmid as well as the unique cDNA sequence contained within the plasmid.

Numerous kits for synthesizing cDNA from mRNA followed by analysis are now commercially available. Kits for the construction of cDNA libraries are also available, or library construction can be contracted for a relatively modest price or a library made from a particular tissue can be purchased from a commercial company. The bacterial library can be gridded onto filters and then screened for a particular gene by hybridization with a probe.

1.3.2 Analysis of gene expression

Quantitative real-time reverse transcription PCR (qRT-PCR; see Logan et al. 2009) is a

method for measuring levels of specific mRNA transcripts within a sample. After the RNA sample is treated with reverse transcriptase, the resulting cDNA is amplified in a PCR reaction using primers specific to a transcript and the amount of transcript quantified in "real time" after each amplification cycle.

Detection of the transcripts is usually done with fluorescent dyes that intercalate with double-stranded DNA, although nonspecific binding can occur, which decreases the accuracy of the quantification. Another method of detection in qRT-PCR uses DNA oligonucleotide probes specific to the transcript that fluoresce when hybridized with a cDNA molecule. This method is more accurate than the double-stranded dyes, but the synthesis of fluorescent reporter probes is expensive.

Relative concentration of the transcript is determined in the qRT-PCR by plotting fluorescence (dependent on the number of copies of the transcript within the sample) against cycle number on a logarithmic scale. The quantity of a control, such as a housekeeping gene, is also measured on each sample so as to normalize for possible variation in the amount and quality of RNA between different samples, with the assumption that the expression of the control is similar across all samples.

"Global" expression of genes within a sample is commonly analyzed using expression microarrays, which allow simultaneous analysis of hundreds to thousands of genes. Probes spotted on the arrays were originally cDNAs but more recently, expression arrays contain oligonucleotides, usually in the range of 25–75 mers, designed from cDNA sequences. The longer the oligonucleotide, the more specific the detection, especially in cross-species experiments (Walker et al. 2006), but the shorter probes can be spotted on the array in higher density and are cheaper to synthesize. Oligonucleotide arrays are usually preferred to the cDNA arrays because of more uniform hybridization and ease of probe synthesis (Barrett and Kawasaki 2003; Hardiman 2004). Detection of a specific transcript within a sample is based on hybridization to the probes on the array. Annotation of the probes on the arrays is a key consideration for their usefulness. Statistical analysis of the data is challenging and requires appropriate controls, normalization of the signals, and adjustments for multiple comparisons. Significant results from an expression array experiment are often verified by qRT-PCR.

SAGE allows whole genome analysis of gene expression (i.e., mRNA) within a sample (Velculescu et al. 1995, 1997). Based on the concept that 10–14 bp of sequence provides "sufficient information to uniquely identify a transcript" within a sequence database, "quantification of the number of times a particular tag is observed provides the expression level of the corresponding transcript" (www.sagenet.org/findings/index.html). Previously unreported genes can also be detected through the generation of tags that are not contained within the databases. Subsequent adaptations of SAGE, such as SuperSAGE (Matsumura et al. 2005), allow precise annotation of existing and new genes because of an increased tag length of 25–27 bp. However, SAGE is relatively much more expensive than DNA microarrays, so large-scale projects are typically not performed with SAGE.

1.3.3 cDNA libraries and reproductive transcriptomes

To date, there are at least 270 publically available cDNA libraries that were derived from different reproductive tissues/organs in

Table 1.8 cDNA libraries and EST sequences for reproductive tissues/organs in livestock species.

Tissue/organ	Cattle		Swine		Sheep	
	No. of libraries	No. of ESTs	No. of libraries	No. of ESTs	No. of libraries	No. of ESTs
Embryo	27	62,951	14	89,916	—	—
Fetus	14	72,914	16	6,468	—	—
Mammary	56	65,227	3	16,656	1	2,309
Ovary	17	13,813	28	75,026	6	2,899
Oviduct	2	70	2	3,556	—	—
Pituitary	5	2,102	7	12,404	—	—
Placenta	10	23,665	6	21,307	—	—
Testes	7	15,033	11	42,494	—	—
Uterus	12	31,380	15	43,392	1	2,722
Total	150	287,155	102	311,219	8	7,930

cattle, swine, and sheep (Table 1.8). The library names, tissue/organ/cell line sources, physiological or reproductive stages, and contributors can be retrieved from either the GenBank database at NCBI (www.ncbi.nlm.nih.gov/) or the Gene Index database at Harvard University (compbio.dfci.harvard.edu/tgi/). As seen in Table 1.8, cattle have a slight edge over swine in the number of constructed libraries (105 and 102, respectively), but swine lead cattle in the number of expressed sequence tags (ESTs) that have been placed in the public databases (311,219 and 287,155, respectively). To date, only eight libraries have been established in sheep, and less than 8000 ovine ESTs for reproductive tissues/organs have been released.

These resources have been widely used in the survey of reproductive transcriptomes, identification of some breed- and developmental-stage-specific genes or gene clusters, and investigations of the genetic and physiological mechanisms underlying reproduction quantitative traits in livestock species. In addition, comparisons of livestock ESTs with sequences from other species have served as a valuable resource for comparative map development.

1.4 Resources for protein analysis

A unique complement of proteins is present in the cells of an organism at any one time under any one condition. This complement of proteins does not necessarily match to the complement of mRNA transcripts within the cells because of posttranslational modifications, splicing variants, and protein and RNA degradation. A recently defined area of research called proteomics encompasses large-scale studies of proteins, including their structure, function, and quantity (Anderson and Anderson 1998; Blackstock and Weir 1999). Two-dimensional (2D) gel electrophoresis is a well-established method commonly used to analyze proteins (Berth et al. 2007), although there are challenges in automatic analysis software. Technologies that allow high-throughput analysis of proteins within a tissue are now available, such as high-performance chromatography and mass spectrometry, but these approaches require highly specialized equipment.

Because of increasing emphasis on systems biology, databases have been created that present whole biological systems of interconnected proteins, with access to underlying genes, their sequences, and the background

studies with a click of a mouse (see www.biochemweb.org/systems.shtml and www.semantic-systems-biology.org/biogateway/querying).

1.5 Future research directions

Genomic resources are now available for all major livestock species. These resources will allow researchers to identify regions within the genome that influence reproductive traits with relative ease. While the pursuit of the causative mutation controlling a quantitative trait may be complicated, combining knowledge from several lines of investigations should lead to the successful identification of the responsible gene. There are also multiple approaches for estimating gene expression in livestock species at both the single and whole genome levels. However, resources for the study of proteins are much less developed in livestock species, and therefore, researchers will need to exploit available information from humans and biomedical animal models.

References

Amaral, M.E., Grant, J.R., Riggs, P.K., Stafuzza, N.B., Filho, E.A., Goldammer, T., Weikard, R., Brunner, R.M., Kochan, K.J., Greco, A.J., Jeong, J., Cai, Z., Lin, G., Prasad, A., Kumar, S., Saradhi, G.P., Mathew, B., Kumar, M.A., Miziara, M.N., Mariani, P., Caetano, A.R., Galvão, S.R., Tantia, M.S., Vijh, R.K., Mishra, B., Kumar, S.T., Pelai, V.A., Santana, A.M., Fornitano, L.C., Jones, B.C., Tonhati, H., Moore, S., Stothard, P., and Womack, J.E. 2009. A first generation whole genome RH map of the river buffalo with comparison to domestic cattle. *BMC Genomics* **9**: 631.

Amos, C.I. 2007. Successful design and conduct of genome-wide association studies. *Human Molecular Genetics* **16**: R220–R225.

Anderson, N.L. and Anderson, N.G. 1998. Proteome and proteomics: New technologies, new concepts, and new words. *Electrophoresis* **19**: 1853–1861.

Andersson, L. and Georges, M. 2004. Domestic-animal genomics: Deciphering the genetics of complex traits. *Nature Reviews in Genetics* **5**: 202–212.

Barrett, J.C. and Kawasaki, E.S. 2003. Microarrays: The use of oligonucleotides and cDNA for the analysis of gene expression. *Drug Discovery Today* **8**: 134–141.

Baumbusch, L.O., Aarøe, J., Johansen, F.E., Hicks, J., Sun, H., Bruhn, L., Gunderson, K., Naume, B., Kristensen, V.N., Liestøl, K., Børresen-Dale, A.L., and Lingjaerde, O.C. 2008 Comparison of the Agilent, ROMA/NimbleGen and Illumina platforms for classification of copy number alterations in human breast tumors. *BMC Genomics* **9**: 379.

Berth, M., Moser, F.M., and Kolbe, M. 2007. The state of the art in the analysis of two-dimensional gel electrophoresis images. *Applied Microbiology and Biotechnology* **76**: 1223–1243.

Blackstock, W.P. and Weir, M.P. 1999. Proteomics: Quantitative and physical mapping of cellular proteins. *Trends in Biotechnology* **17**: 121–127.

Boehnke, M. 1992. Multipoint analysis for radiation hybrid mapping. *Annuals in Medicine* **24**: 383–386.

Bonnet, A., Thevenon, S., Claro, F., Gautier, M., and Hayes, H. 2001. Cytogenetic comparison between Vietnamese sika deer and cattle: R-banded karyotypes and FISH mapping. *Chromosome Research* **9**: 673–687.

Burt, D.W. 2002. Comparative mapping in farm animals. *Briefings in Functional Genomics and Proteomics* **1**: 159–168.

Charlier, C., Coppieters, W., Rollin, F., Desmecht, D., Agerholm, J.S., Cambisano, N., Carta, E., Dardano, S., Dive, M., Fasquelle, C., Frennet, J.C., Hanset, R., Hubin, X., Jorgensen, C., Karim, L., Kent, M., Harvey, K., Pearce, B.R., Simon, P., Tama, N., Nie, H., Vandeputte, S., Lien, S., Longeri, M., Fredholm, M., Harvey, R.J., and Georges, M. 2008. Highly effective SNP-based association mapping and management of recessive defects in livestock. *Nature Genetics* **40**: 449–454.

Chowdhary, B.P., Fronicke, L., Gustavsson, I., and Scherthan, H. 1996. Comparative analysis of the cattle and human genomes: detection of ZOO-FISH and gene mapping-based chromosomal homologies. *Mammalian Genome* **7**: 297–302.

Chowdhary, B.P. and Raudsepp, T. 2001. Chromosome painting in farm, pet and wild animal species. *Methods in Cell Science* **23**: 37–55.

Chowdhary, B.P. and Raudsepp, T. 2008. The horse genome derby: Racing from map to whole genome sequence. *Chromosome Research* **16**: 109–127.

Cockett, N.E., Jackson, S.P., Shay, T.L., Nielsen, D., Steele, M.R., Green, R.D., and Georges, M. 1994. Chromosomal localization of the *callipyge* gene in sheep (*Ovis aries*) using bovine DNA markers. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 3019–3023.

Cockett, N.E, Jackson, S.P., Shay, T.L., Farnir, F., Berghmans, S., Snowder, G., Nielsen, D., and Georges, M. 1996. Polar overdominance at the ovine *callipyge* locus. *Science* **273**: 236–238.

Dalrymple, B.P., Kirkness, E.F., Nefodov, M., McWilliam, S., Ratnakumar, A., Barris, W., Zhao, S., Shetty, J., Maddox, J.F., O'Grady, M., Nicholas, F., Crawford, A.M., Smith, T., de Jong, P., McEwan, J.C., Oddy, V.H., and Cockett, N.E. 2007. Constructing the virtual sheep genome. *Genome Research* **8**: R152.

de Koning, D.J., Pong-Wong, R., Varona, L., Evans, G.J., Giuffra, E., Sanchez, A., Plastow, G., Noguera, J.L., Andersson, L., and Haley, C.S. 2003. Full pedigree quantitative trait locus analysis in commercial pigs using variance components. *Journal of Animal Science* **81**: 2155–2163.

de Koning, D.J., Archibald, A., and Haley, C.S. 2007. Livestock genomics: Bridging the gap between mice and men. *Trends in Biotechnology* **25**: 483–489.

Di Meo, G.P., Perucatti, A., Floriot, S., Hayes, H., Schibler, L., Incarnato, D., Di Berardino, D., Williams, J., Cribiu, E., Eggen, A., and Iannuzzi, L. 2008. An extended river buffalo (*Bubalus bubalis*, 2n = 50) cytogenetic map: Assignment of 68 autosomal loci by FISH-mapping and R-banding and comparison with human chromosomes. *Chromosome Research* **16**: 827–837.

Di Meo, G.P., Perucatti, A., Floriot, S., Hayes, H., Schibler, L., Rullo, R., Incarnato, D., Ferretti, L., Cockett, N., Cribuiu, E., Williams, J.L., Eggen, A., and Iannuzzi, L. 2007. An advanced sheep (*Ovis aries*, 2n = 54) cytogenetic map and assignment of 88 new autosomal loci by fluorescence in situ hybridization and R-banding. *Animal Genetics* **38**: 233–240.

Farnir, F., Grisart, B., Coppieters, W., Riquet, J., Berzi, P., Cambisano, N., Karim, L., Mni, M., Moisio, S., Simon, P., Wagenaar, D., Vilkki, J., and Georges, M. 2002. Simultaneous mining of linkage and linkage disequilibrium to fine map quantitative trait loci in outbred half-sib pedigrees: Revisiting the location of a

quantitative trait locus with major effect on milk production on bovine chromosome 14. *Genetics* **161**: 275–287.

Georges, M. 2007. Mapping, fine mapping, and molecular dissection of quantitative trait loci in domestic animals. *Annual Reviews of Genomics and Human Genetics* **8**: 131–162.

Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P., Spelman, R., Georges, M., and Snell, R. 2001. Positional candidate cloning of a QTL in dairy cattle: Identification of a missense mutation in the bovine *DGAT1* gene with major effect on milk yield and composition. *Genome Research* **12**: 222–231.

Grisart, B., Farnir, F., Karim, L., Cambisano, N., Kim, J.J., Kvasz, A., Mni, M., Simon, P., Frère, J.M., Coppieters, W., and Georges, M. 2004. Genetic and functional confirmation of the causality of the *DGAT1 K232A* quantitative trait nucleotide in affecting milk yield and composition. *Proceedings of the National Academy of Sciences in the United States of America* **101**: 2398–2403.

Gustafson, A.L., Tallmadge, R.L., Ramlachan, N., Miller, D., Bird, H., Antczak, D.F., Raudsepp, T., Chowdhary, B.P., and Skow, L.C. 2003. An ordered BAC contig map of the equine major histocompatibility complex. *Cytogenetics and Genome Research* **102**: 189–195.

Hardiman, G. 2004. Microarray platforms—Comparisons and contrasts. *Pharmacogenomics* **5**: 487–502.

Humphray, S.J., Scott, C.E., Clark, R., Marron, B., Bender, C., Camm, N., Davis, J., Jenks, A., Noon, A., Patel, M., Sehra, H., Yang, F., Rogatcheva, M.B., Milan, D., Chardon, P., Rohrer, G., Nonneman, D., de Jong, P., Meyers, S.N., Archibald, A., Beever, J.E., Schook, L.B., and Rogers, J. 2007. A high utility integrated map of the pig genome. *Genome Biology* **8**: R139.

Khatkar, M.S., Zenger, K.R., Hobbs, M., Hawken, R.J., Cavanagh, J.A., Barris, W., McClintock, A.E., McClintock, S., Thomson, P.C., Tier, B., Nicholas, F.W., and Raadsma, H.W. 2007. A primary assembly of a bovine haplotype block map based on a 15,036-single-nucleotide polymorphism panel genotyped in Holstein-Friesian cattle. *Genetics* **176**: 763–772.

Kim, Y. and Nielsen, R. 2004. Linkage disequilibrium as a signature of selective sweeps. *Genetics* **167**: 1513–1524.

Knott, S.A. and Haley, C.S. 2000. Multitrait least squares for quantitative trait loci detection. *Genetics* **156**: 899–911.

Lange, K., Boehnke, M., Cox, D.R., and Lunetta, K.L. 1995. Statistical methods for polyploid radiation hybrid mapping. *Genome Research* **5**: 136–150.

Laurent, P., Schibler, L., Vaiman, A., Laubier, J., Delcros, C., Cosseddu, G., Vaiman, D., Cribiu, E.P., and Yerle, M. 2007. A 12,000-rad whole-genome radiation hybrid panel in sheep: Application to the study of the ovine chromosome 18 region containing a QTL for scrapie susceptibility. *Animal Genetics* **38**: 358–363.

Logan, J., Edwards, K., and Saunders, N. (eds.). 2009. *Real-Time PCR: Current Technology and Applications*. Wymondham, UK: Caister Academic Press.

Lunetta, K.L., Boehnke, M., Lange, L., and Cox, D.R. 1996. Selected locus and multiple panel models for radiation hybrid mapping. *American Journal of Human Genetics* **59**: 717–725.

Maddox, J.F., Davies, K.P., Crawford, A.M., Hulme, D.J., Vaiman, D., Cribiu, E.P., Freking, B.A., Beh, K.J., Cockett, N.E., Kang, N., Riffkin, C.D., Drinkwater, R., Moore, S.S., Dodds, K., Lumsden, J.K., Adelson, D., Birkin, H., Broom, J.E.,

Buitkamp, J., Cambridge, E., Cushwa, W.T., Gerard, G., Galloway, S., Harrison, B., Hawken, R.J., Hiendleder, S., Henry, H., Medrano, J., Paterson, K., Phua, S.H., Schibler, L., Stone, R.T., and van Hest, B. 2001. An enhanced linkage map of the sheep genome comprising more than 1000 loci. *Genome Research* **11**: 1275–1289.

Matsumura, H., Ito, A., Saitoh, H., Winter, P., Kahl G., Reuter, M., Kruger, D.H., and Terauchi, R. 2005. SuperSAGE. *Cellular Microbiology* **7**: 11–18.

McCarroll, S.A. 2008. Extending genome-wide associations studies to copy-number variation. *Human Molecular Genetics* **17**: R135–R142.

McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J., Ioannidis, J.P., and Hirschhorn, J.N. 2008. Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. *Nature Reviews Genetics* **9**: 356–369.

McVean, G. 2007. The structure of linkage disequilibrium around a selective sweep. *Genetics* **175**: 1395–1406.

Meuwissen, T.H., Hayes, B.J., and Goddard, M.E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**: 1819–1829.

Nejati-Javaremi, A. and Smith, C. 1995. Assigning linkage haplotypes from parent and progeny genotypes. *Genetics* **142**: 1363–1367.

Orr, N. and Chanock, S. 2008. Common genetic variation and human diseases. *Advances in Genetics* **62**: 1–32.

Papp, A.C., Pinsonneault, J.K., Cooke, G., and Sadée, W. 2003. Single nucleotide polymorphism genotyping using allele-specific PCR and fluorescence melting curves. *Biotechniques* **34**: 1068–1072.

Pollinger, J.P., Bustamante, C.D., Fiedel-Alon, A., Schmutz, S., Gray, M.M., and Wayne, R.K. 2005. Selective sweep mapping of genes with large phenotypic effects. *Genome Research* **15**: 1809–1819.

Rijnkels, M., Lemay, D.G., Barris, W.C., Casey, T.M., German, J.B., Hinrichs, A.S., Kriventseva, E.V., Lynn, D.J., Martin, W.F., Maqbool, N.J., Medrano, J.F., Molenaar, A.J., Neville, M.C., Pollard, K.S., Rincon, G., Zdobnov, E.M., Tellam, R.L., and Bovine Genome Sequencing and Analysis Consortium. 2009. Milking the bovine genome: insights in role of milk and variation of milk composition. *Plant and Animal Genomes XVII Conference*, W088, January 10–14, San Diego, CA.

Riquet, J., Coppieters, W., Cambisano, N., Arranz, J.J., Berzi, P., Davis, S.K., Grisart, B., Farnir, F., Karim, L., Mni, M., Simon, P., Taylor, J.F., Vanmanshoven, P., Wagenaar, D., Womack, J.E., and Georges, M. 1999. Fine-mapping of quantitative trait loci by identify by descent in outbred populations: Application to milk production in dairy cattle. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 9252–9257.

Rocha, J.L., Baker, J.F., Womack, J.E., Sanders, J.O., and Taylor, J.F. 1992. Statistical associations between restriction fragment length polymorphisms and quantitative traits in beef cattle. *Journal of Animal Science* **70**: 3360–3370.

Rocha, J.L., Pomp, D., and Van Vleck, L.D. 2002. QTL analysis in livestock. *Methods in Molecular Biology* **195**: 311–346.

Ron, M. and Weller, J.I. 2007. From QTL to QTN identification in livestock—Winning by points rather than knockout: A review. *Animal Genetics* **38**: 429–439.

Saiki, R.K., Bugawan, T.L., Horn, G.T., Mullis, K.B., and Erlich, H.E. 1986. Analysis of enzymatically amplified

beta-globin and HLA-DQ DNA with allele-specific oligonucleotide probes. *Nature* **324**: 163–166.

Sandor, C. and Georges, M. 2008. On the detection of imprinted quantitative trait loci in line crosses: Effect of linkage disequilibrium. *Genetics* **180**: 1167–1175.

Schaschi, H., Aitman, T.J., and Vyse, T.J. 2009. Copy number variation in the human genome and its implication in autoimmunity. *Clinical and Experimental Immunology*. February 11 [Epub ahead of print].

Schibler, L., Vaiman, D., Oustry, A., Giraud-Delville, C., and Cribiu, E.P. 1998a. Comparative gene mapping a fine-scale survey of chromosome rearrangements between ruminants and humans. *Genome Research* **8**: 901–915.

Schibler, L., Vaiman, D., Oustry, A., Guinec, N., Dangy-Caye, A.L., Billault, A., Cribiu, E.P. 1998b. Construction and extensive characterization of a goat bacterial artificial chromosome library with threefold genome coverage. *Mammalian Genome* **9**: 119–124.

Schwerin, M. 2001. Molecular genome analysis in livestock at the beginning of a new millennium. *Reproduction in Domestic Animals* **36**: 133–138.

Sellner, E.M., Kim, J.W., McClure, M.C., Taylor, K.H., Schnabel, R.D., and Taylor, J.F. 2007. Board-invited review: Applications of genomic information in livestock. *Journal of Animal Science* **85**: 3148–3158.

Slate, J., Van Stijn, T.C., Anderson, R.M., McEwan, K.M., Maqbool, N.J., Mathias, H.C., Bixley, J.J., Stevens, D.R., Molenaar, A.J., Beever, J.E., Galloway, S.M., and Tate, M.L. 2002. A deer (subfamily Cervinae) genetic linkage map and the evolution of ruminant genomes. *Genetics* **160**: 1587–1597.

Snelling, W.M., Chiu, R., Schein, J.E., Hobbs, M., Abbey, C.A., Adelson, D.L., Aerts, J., Bennett, G.L., Bosdet, I.E., Boussaha, M., Brauning, R., Caetano, A.R., Costa, M.M., Crawford, A.M., Dalrymple, B.P., Eggen, A., Everts-van der Wind, A., Floriot, S., Gautier, M., Gill, C.A., Green, R.D., Holt, R., Jann, O., Jones, S.J., Kappes, S.M., Keele, J.W., de Jong, P.J., Larkin, D.M., Lewin, H.A., McEwan, J.C., McKay, S., Marra, M.A., Mathewson, C.A., Matukumalli, L.K., Moore, S.S., Murdoch, B., Nicholas, F.W., Osoegawa, K., Roy, A., Salih, H., Schibler, L., Schnabel, R.D., Silveri, L., Skow L.C., Smith, T.P., Sonstegard, T.S., Taylor, J.F., Tellam, R., Van Tassell, C.P., Williams, J.L., Womack, J.E., Wye, N.H., Yang, G., Zhao, S., and the International Bovine BAC Mapping Consortium. 2007. A physical map of the bovine genome. *Genome Biology* **8**: R165.

Strerath, M., Detmer, I., Gaster, J., and Marx, A. 2007. Modified oligonucleotides as tools for allele-specific amplification. *Methods in Molecular Biology* **402**: 317–328.

Tellam, R.L. and the Bovine Genome Sequencing and Analysis Consortium. 2009. What does analysis of the bovine genome sequence say about innate immunity? *Plant and Animal Genome XVII Conference*, W087, January 10–14, San Diego, CA..

Tian, C., Gregersen, P.K., and Seldin, M.F. 2008. Accounting for ancestry: Population substructure and genome-wide association studies. *Human Molecular Genetics* **17**: R143–150.

Van Laere, A.S., Nguyen, M., Braunschweig, M., Nezer, C., Collette, C., Moreau, L., Archibald, A.L., Haley, C.S., Buys, N., Tally, M., Andersson, G., Georges, M., and Andersson, L. 2003. A regulatory mutation in IGF2 causes a major QTL

effect on muscle growth in the pig. *Nature* **425**: 832–836.

Van Tassell, C.P., Smith, T.P., Matukumalli, L.K., Taylor, J.F., Schnabel, R.D., Lawley, C.T., Haudenschild, C.D., Moore, S.S., Warren, W.C., and Sonstegard, T.S. 2008. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. *Nature Methods* **5**: 247–252.

Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. 1995. Serial analysis of gene expression. *Science* **270**: 484–487.

Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E., Hieter, P., Vogelstein, B., and Kinzler, K.W. 1997. Characterization of the yeast transcriptome. *Cell* **88**: 243–251.

Voight, B.F., Kudaravalli, S., Wen, X., and Pritchard, J.K. 2006. A map of recent positive selection in the human genome. *PLoS Biology* **4**: e72.

Walker, S.J., Wang, Y., Grant, K.A., Chan, F., and Hellmann, G.M. 2006. Long versus short oligonucleotide microarrays for the study of gene expression in nonhuman primates. *Journal of Neuroscience Methods* **152**: 179–189.

Wang, J., Chuang, K., Ahluwalia, M., Patel, S., Umblas, N., Mirel, D., Higuchi, R., Germer, S. 2005. High-throughput SNP genotyping by single-tube PCR with TM-shift primers. *Biotechniques* **39**: 885–893.

Wu, C.H., Nomura, K., Goldammer, T., Hadfield, T., Dalrymple, B.P., McWilliam, S., Maddox, J.F., Womack, J.E., and Cockett, N.E. 2008. A high-resolution comparative radiation hybrid map of ovine chromosomal regions that are homologous to human chromosome 6 (HSA6). *Animal Genetics* **39**: 459–467.

Wu, C.H., Nomura, K., Goldammer, T., Hadfield, T., Dalrymple, B.P., McWilliam, S., Maddox, J.F., Womack, J.E., and Cockett, N.E. 2009. A radiation hybrid comparative map of ovine chromosome 1 aligned to the virtual sheep genome. *Animal Genetics* **40**: 435–455.

Wu, C.H., Nomura, K., Goldammer, T., Hadfield, T., Womack, J.E., and Cockett, N.E. 2007. An ovine whole-genome radiation hybrid panel used to construct an RH map of ovine chromosome 9. *Animal Genetics* **38**: 534–536.