8

Paul Denny

8.1 Introduction: Key Experimental Advantages

The earliest written record of an interest in mice is found in a Chinese lexicon, published in about 1100 BC, in which there was a special word for a spotted mouse. In more recent times, directed breeding of unusual or "fancy" mice dates back to at least the eighteenth century in Japan (Keeler and Fuji 1937). As a small social mammal with a short reproductive cycle, that breeds well in communal housing all year round, the mouse is an excellent laboratory animal. Mice are used extensively in many areas of biomedical research and are the most commonly used genetically modified mammal. Mice were the first mammals shown to follow Mendel's laws of genetics (Cuenot 1902) and in the early twentieth century were one of the first organisms in which genetic linkage was demonstrated (Haldane et al. 1915). More recently, one of the reasons for the utility of the mouse in genetic studies is the relative ease with which embryonic stem cells can be derived, chosen genes modified, and the genetic change transmitted to offspring. This technological feat was recognized in the award of the 2007 Nobel Prize (Mak 2007). The mouse was the second mammal (after humans) for which a draft genome sequence was determined (Waterston et al. 2002) and indeed there is now a "finished" genome (Church et al. 2009) of higher quality than that available for humans. The availability of the draft sequence further stimulated the utility of the mouse in genetic and postgenomic research.

The Mouse Genome Database (http://www.informatics.jax.org/) is the closest murine equivalent of the databases for *C. elegans* (WormBase— see Chap. 2) or *Drosophila* (FlyBase—Chap. 3), but there are numerous other important online or print resources describing the biology, genetics, and genomics of the mouse. The classical work on the anatomy of the adult mouse (Cook 1965) is also available in electronic form (http://www.informatics.jax.org/cookbook/) and the reference works on the developing mouse embryo and fetus (Kaufman 1992; Theiler 1989) underpin the Edinburgh Mouse Atlas Project (EMAP: http://www.emouseatlas.org/emap/home.html). Mouse genetics is well served by the classic works "Genetic Variants and Strains of the Laboratory Mouse" (Lyon et al. 1996) and "Mouse Genetics: Concepts and Applications" (Silver 1995). The primary source for the mouse reference genome sequence is the assembly maintained by the NCBI (currently at version 38); this underpins the genome annotations in the main genome browsers at Ensembl, the University of California at Santa Cruz (UCSC) and NCBI (see Table 8.1).

In the wild, mice have a wide distribution throughout Europe and Asia and were most likely introduced into the Americas and Australia

P. Denny (✉)

Mammalian Genetics Unit, Medical Research Council
Harwell, Oxfordshire OX11 0RD, UK
e-mail: paul.denny3@gmail.com

Table 8.1 Genome sequence resources

Genome resource	Citation/URL	Description
Ensembl genome browser	http://www.ensembl.org/Mus_musculus/Info/Index	Annotations of the public C57BL/6J reference sequence assembly.
NCBI genome browser	http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=10090	Public genome assembly for the C57BL/6J reference sequence and annotations. BAC-based partial assemblies from other mouse strains, together with annotations. Alternate assembly and annotation for the Celera mouse genome.
UCSC genome browser	http://genome.ucsc.edu/cgi-bin/hgGateway?hgsid=238431269&clade=mammal&org=Mouse&db=0	Annotations of the public C57BL/6J sequence assembly.
Genome Reference Consortium	http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/mouse/index.shtml	Documents changes to the C57BL/6J reference sequence assembly as gaps are filled and errors corrected, in an ongoing effort to improve quality.
Vega	http://vega.sanger.ac.uk/Mus_musculus/Info/Index	Manual annotation of mouse genes—contrast with NCBI, Ensembl & UCSC, which are all based on automated annotation "pipelines".
WTSI mouse genomes project	http://www.sanger.ac.uk/resources/mouse/genomes/	Genome sequences from 17 additional key mouse strains.

by man. Mice are omnivores and can be maintained easily on a standard diet from commercial suppliers. The generation time of mice is short, varying somewhat between inbred strain; typically 8–12 weeks from egg to egg. They are prolific breeders, with litter sizes of more than ten not uncommon in outbred stocks, but usually smaller numbers in inbred mice. Despite their prolific nature, some mouse stocks are maintained using assisted reproductive technologies such as in vitro fertilization (Takahashi et al. 2010). Embryos from mice can be stored in the gas phase of liquid nitrogen for at least 35 years and spermatozoa for at least 22 years (Nakagata 2000). This technology reduces the cost of maintaining large numbers of different mutant lines, because they need not be kept as breeding animals, minimizes spontaneous genetic drift in mouse colonies, simplifies shipping of mice between facilities and also can help improve the health status of animals. The reproductive technology of cloning has also been used

successfully in mice (Thuan et al. 2010), but, unlike IVF, is not in widespread use.

The origins of laboratory mice are complex and it seems that most strains derive from the interbreeding of three mouse subspecies: *Mus musculus musculus*, *Mus musculus domesticus*, and *Mus musculus castaneus* (Guenet and Bonhomme 2003). One of the critical factors in the success of the mouse as a laboratory animal has been the fact that mice can be inbred, unlike many other species. Over the last 100 years, many inbred and outbred mouse genetic resources have been developed, as given in Table 8.2.

8.2 Genome Mapping

The earliest genome maps of the mouse were those constructed by linkage analysis of mutations that caused visible traits (Silver 1995). The first traits included on maps arose spontaneously and were noticed in the fancy mice bred in Europe and the USA. Later, mutations induced

Table 8.2 Genetic resources

Genetic resource	Number of stocks or strains
Inbred	>470
Recombinant inbred	>365
Congenic	>3,600
Recombinant congenic	>155
Consomic	>165
Wild-derived	>95
Conplastic	>10
Mutants	>4,200

Sources: IMSR, MGI (Burgio et al. 2007; Gregorova et al. 2008; Takada et al. 2008; Stassen et al. 1996)

by radiation or chemicals gave rise to many more traits that were mapped.

All mouse chromosomes are acrocentric, meaning that they have the centromere placed very close to one end, so that one chromosome arm is long and the other very short. The normal mouse karyotype is diploid, consisting of 40 chromosomes; 19 pairs of autosomes and two X chromosomes in females and a single X and a Y chromosome in males. The autosomes do not differ substantially in size and can only be identified reliably after dye staining. The first cytogenetic maps, showing the distinctive banding patterns of mouse chromosomes stained with Giemsa or quinacrine and more recently with DAPI, allowed the integration of genetic linkage and physical maps (Kouri et al. 1971). A representative Giemsa-banded mouse karyotype and an idealized version of the cytogenetic map are shown in Fig. 8.1.

Mouse genetic maps grew slowly over the first 80 years of the twentieth century, but with the development of molecular cloning methods and the discovery of highly polymorphic tandem repetitive elements ("microsatellites") that could be used to distinguish the parental origin of chromosome regions in the offspring of crosses between different inbred strains (Dietrich et al. 1992), this process accelerated. These microsatellite-based maps were used mainly to allow position-based cloning of genes underlying Mendelian (Zhang et al. 1994) and quantitative traits (Cormier et al. 1997).

Two main classes of physical maps have been developed, complementary to genetic maps, which do not depend on animal breeding and

ultimately relate the genome sequence to the underlying chromosomes. These are maps that give a direct view of genome features, such as cytogenetic maps (Kouri et al. 1971), and indirect maps, that involve fragmentation and an analytical process of reconstruction to infer genome structure, such as radiation hybrid (Hudson et al. 2001) and clone-based physical maps. Physical maps were needed for positional cloning of mutations, were useful in the process of hierarchical shotgun genome sequencing, and indeed have been used to independently validate the genome sequence map (Zhou et al. 2007).

The first, indirect, types of map derived from either panels of hamster cell lines carrying an intact subset of mouse chromosomes (Kozak et al. 1975) or random fragments of chromosomes produced by irradiation ("radiation hybrid" (RH) map) (McCarthy et al. 1997). These somatic maps have the advantage of being useful to locate any STS for which an amplicon can be designed that is either absent from the hamster genome or is a different size in hamster and mouse.

The other, molecular type of indirect physical maps is based on mainly yeast artificial chromosomes (YACs) or bacterial artificial chromosomes (BACs), using the same techniques pioneered in other organisms, e.g., BAC restriction digest "fingerprinting" in *C. elegans* (see Chap. 2). The BAC system generates clones that are more stable and more easily manipulated than YACs (Kim et al. 1996); nevertheless, YACs have been useful for studying long-range genomic structure and also in functional studies (Sharpe et al. 1999).

The current reference genetic map for the mouse is based on a massive effort to genotype a complex cross, with eight inbred strains as its starting point, with almost 10,200 single nucleotide polymorphism (SNP) markers (Cox et al. 2009). This level of map density has facilitated both haplotype mapping, which is useful for better localization of quantitative trait loci (QTL) and also a close examination of genome structure, e.g., to understand better the relationship between physical distance and variation in recombination frequency.

Evans EP (1996) Standard Idiogram. *In* Lyon MF, Rastan S & Brown SDM (eds.) *Genetic Variants and Strains of the Laboratory Mouse*, 3rd ed. Oxford University Press: pp 1446-1448

Courtesy of
Dr E P Evans

Fig. 8.1 (**a**) Standard idiogram for the Giemsa-banded mouse karyotype (Evans 1996). (**b**) Representative Giemsa-banded mouse karyotype (Courtesy of EP Evans and CV Beechey)

A number of reference maps of the mouse genome exist, based on physical or genetic mapping techniques (see Table 8.3), which have been used to underpin the ultimate physical map: the genome sequence.

8.3 Genomics

8.3.1 General Organization

The mouse genome is composed of 20 pairs of nuclear chromosomes, the largest being chromosome 1 (197 Mb) and the smallest the Y chromosome (15.9 Mb—NB only the euchromatic region has been assembled), and the 16.3 kb mitochondrial genome. On all mouse nuclear chromosomes, the relationship between recombination rate and physical size follows a slightly sigmoid shape, with recombination rates suppressed at the proximal, centromeric end and accentuated at the distal chromosome end (Cox et al. 2009).

There are two distinct mouse whole-genome sequence assemblies: one produced by Celera Genomics (Mural et al. 2002) and the other produced by the public genome project (Table 8.1). The Celera assembly is based on whole-genome shotgun sequence from four mouse inbred strains (A/J, DBA2/J, 129X1/SvJ, and 129S1/SvImJ), whereas the public sequence is a composite of sequences from finished BACs and from whole-genome shotgun sequence, both from the C57BL/6J (B6) inbred strain. The B6 genome sequence is now essentially finished, consisting of over 95 % finished BAC sequence, with only ~1,200 sequence gaps (Church et al. 2009). There is an ongoing endeavor to improve the public genome assembly, lead by the Genome Reference Consortium (Table 8.1).

8.3.2 Protein Coding Genes

Based on the Ensembl 65 annotations of the mouse genome, there are 21,879 "known" and 826 "novel" protein coding genes in the mouse genome. The existence of these genes is based on a combination of computational prediction and in all cases, independent evidence of corresponding RNA transcripts and/or encoded proteins (Curwen et al. 2004). It is possible, however, that some of these "genes" may not encode proteins and are actually misannotated pseudogenes, as discussed below. The gene counts obtained by the other genome annotation databases at NCBI and UCSC (Table 8.1) are broadly similar, but there is an effort to improve and harmonize annotations of the protein-coding genes, known as the Consensus Coding Sequence (CCDS) project. The aim is to identify a core set of mouse (and human) protein coding regions that are annotated consistently and of high quality (Table 8.1).

The protein-encoding genes of the mouse vary wildly in size, but are typical for a mammalian genome—the largest, encoding dystrophin, spans about 2.6 Mb of the X chromosome. The average mouse exon is 280 bp; introns are much larger, with an average size of 4,981 bp, but the distribution of sizes is very wide—the largest intron being over 1 Mb.

8.3.3 Pseudogenes

Functionally inactive genes are labeled as pseudogenes and may either be transcribed, yet not apparently able to be translated to produce a protein, or nontranscribed. There are 5,228 pseudogenes annotated in Ensembl 65, yet this figure is open to a number of caveats. Firstly, some pseudogenes may be translated after mRNA editing replaces stop codons (Wagner et al. 2003). Secondly, there are may be computational artifacts that lead to misannotation of pseudogenes. Finally, in some cases, transcribed pseudogenes are still clearly functional—a recent example being the PTENP1 pseudogene, which regulates its "functional" homolog, PTEN, by interactions with microRNAs (Poliseno et al. 2010).

Pseudogenes have been described both as evolutionary garbage and as working material from which "new" genes may evolve (Lachmann 2010). The comparative analysis of genome sequences from wild mouse species, e.g., inbred

Table 8.3 Genome reference maps

Reference map	Type	Citations/URL	Description
Jackson Laboratory MGI	Integrated	http://www.informatics.jax.org/searches/linkmap_form.shtml	Incorporates information from various sources, including those mentioned below. As a consensus map, it has compromises between conflicting data. It represents the most comprehensive set of genetic maps of the mouse genome.
Jackson Laboratory/ University of Wisconsin	Linkage	Cox et al. (2009) http://cgd.jax.org/mousemapconverter/	Based on 10,195 SNP markers typed on a large heterogeneous stock (HS) population descended from eight inbred strains. This is now the primary mouse genetic map.
Whitehead	Linkage	Dietrich et al. (1996)	Based mainly on microsatellite markers typed on a small intersub-specific F1 intercross, allowing resolution of only ± 1 cM, but density of markers (about 7000), typed by one lab using a standardized protocol on a single cross gave high consistency. Was useful for positional cloning.
Jackson Laboratory BSS, BSB	Linkage	Rowe et al. (1994)	Based mainly on microsatellite markers typed on two, reciprocal, interspecific backcrosses, allowing resolution of the order of ± 0.5 cM. In contrast to the Whitehead map, these included data submitted by workers outside the Jackson Labs, but curated for quality control. Was used for positional cloning.
European Collaborative Interspecific Backcross (EUCIB)	Linkage	Rhodes et al. (1998)	Based mainly on microsatellite markers typed on two, reciprocal, interspecific backcrosses, allowing high resolution of ± 0.1 cM. Was very useful for positional cloning.
Whitehead/MRC	RH	Hudson et al. (2001)	A genome map in which STSs are positioned relative to one another on the basis of the frequency with which they are separated by radiation-induced breaks. Combines microsatellite, random STS and EST markers.
Genoscope	RH	Avner et al. (2001)	Further RH map based on microsatellite and EST markers.
Whitehead/MRC	YAC	Nusbaum et al. (1999)	YAC framework map, covering about 92 % of the genome.
WTSI/Baylor	BAC	Cai et al. (2001)	BAC map, based on a combination of screens of the library with microsatellite markers and individual clone fingerprints.
WTCHG/WTSI	Sequence-based	Yalcin et al. (2011)	Comprehensive map of structural variation in the genomes of 17 inbred strains.
Imprinting Maps, Harwell	Integrated	http://www.har.mrc.ac.uk/research/genomic_imprinting/	The data on these maps comprise information on mouse chromosome regions associated with imprinted phenotypes, imprinted genes within these regions, and imprinted genes in other regions of the genome.

People wishing to map a gene or other marker should first consult an annotated genome sequence assembly (either public or private) and cross-check with genetic consensus maps such as at MGI, or other independent maps such as one of the RH maps. There are still limitations in the genome sequence maps e.g., gaps and assembly problems in some regions and no coverage of heterochromatic parts of the genome, which may harbor some active genes

strains derived from *Mus spretus*, with the B6 reference sequence, may throw some light onto these hypotheses, because they shared a common ancestor about 1.5–2 million years ago (Guenet and Bonhomme 2003).

8.3.4 Major Protein Coding Gene Families

By comparison with the human genome sequence, it is clear that there has been both expansion and contraction of specific protein coding gene families in the mouse (Church et al. 2009; Waterston et al. 2002). The olfactory and vomeronasal receptor gene families are expanded in mice, which may be related to the dependence of these animals on olfaction both in exploring their environment and also in behavioral interactions. Other large gene families are the KRAB zinc-finger and high-mobility-group (1 and 2) DNA/RNA-binding proteins. There are over 50 IgG kappa light chain genes in the mouse genome, by comparison with 13 in humans. Some of these differences may be the result of lineage-specific selective pressure; perhaps we should not be surprised that the general functional classes of gene exhibiting the biggest differences are involved in olfaction, reproduction, and immunity, areas of biology that show large physiological and behavioral differences between mouse and man (Emes et al. 2003).

8.3.5 RNA Genes: Translational and Other Species

The genome includes many genes that do not encode protein, but instead encode functional RNA, involved in many different processes, including transcriptional regulation, mRNA processing, translation, and turnover. The challenge of annotating RNA genes is qualitatively distinct from the identification of protein-coding genes because of the lack of distinct signals or hallmarks. The best understood of the RNA genes are the 355 tRNA genes that match the expected anticodons and the 247 genes for ribosomal RNAs, spliceosomal RNAs, and telomerase

RNA (Waterston et al. 2002), which are scattered across the mouse genome.

The number of distinct classes of noncoding RNA molecules has grown dramatically over recent years, due, in part to the large-scale sequencing of transcripts (Carninci 2006; Okazaki et al. 2002; Strausberg et al. 2002) and to novel discoveries in other laboratory animals such as *C. elegans* (see Chap. 2). These large noncoding RNAs are mostly of unknown function, but some are apparently involved in various aspects of the regulation of imprinted gene expression, dosage compensation, development, and tumorigenesis (Esteller 2011; Guttman et al. 2009). Two examples of noncoding RNAs with specific functions are XIST and NRON. The XIST RNA regulates the transcriptional activity of the X chromosome as a major part of the mechanism of dosage compensation of X-linked genes (Avner and Heard 2001). A noncoding RNA called NRON, identified in the mouse and highly conserved in humans, apparently functions as a component of an RNA–protein complex that represses the NFAT transcription factor (Willingham et al. 2005). For a more extensive review of noncoding RNA genes, consult Esteller (2011).

8.3.6 Small RNA Genes: Regulatory MicroRNAs and Other Species

A highly heterogeneous mixture of small RNAs have been discovered in eukaryotes, including microRNA (miRNA), PIWI-interacting RNAs (piRNAs), and small nucleolar RNAs (snoRNAs) (Esteller 2011). The main function of miRNAs (19–24 nt) is posttranscriptional gene silencing by targeting of messenger RNAs and it is estimated that they regulate the translation of about 60 % of protein-coding genes. About 1,400 transcriptional units encoding these small RNAs are found throughout the mouse genome (Esteller 2011).

The main role of piRNAs (24–30 nt) is to bind the PIWI subfamily of Argonaute family proteins that are involved in maintaining genome stability in germline cells. They are transcribed from parts of the genome that contain transcriptionally active transposable elements. Recent evidence

suggests that piRNAs also regulate imprinting-related DNA methylation (Watanabe et al. 2011).

The snoRNAs (60–300 nt) are components of small nucleolar ribonucleoproteins (snoRNPs) and are involved in posttranscriptional methylation and pseudouridylation of ribosomal RNA. These modifications of the rRNA are essential for normal folding and stability.

8.3.7 Transposons

Transposons are mobile genetic elements—they are able to integrate into and also move around the mammalian genome (Kazazian 2004). The main classes of transposons in the mouse genome are DNA transposons and retrotransposons, with the latter class divided up by the presence or absence of flanking long terminal repeat (LTR) sequences. The LTR retrotransposons are responsible for most mobile-element insertions in mice; intracisternal A-particles (IAPs), early transposons (Etns), and mammalian LTR-retrotransposons (MaLRs). The largest group of non-LTR retrotransposons is the long inter-spersed repetitive element (LINE) L1 repeat family (see below) and about 3,000 copies behave as active transposons. It is likely that these mobile elements supply some of the working material from which "new" genes evolve (Kazazian 2004).

8.3.8 Repetitive Sequences

A large proportion of the mammalian genome consists of repetitive sequences, many of which are interspersed repeats, derived from transposable elements. The mouse genome consists of 42.1 % of repetitive elements (Church et al. 2009), belonging to 16 different major repeat families. As in humans, there is a strong correlation in the density of repeat element distribution with (G + C) content; LINE elements being concentrated more in (A + T) rich regions and SINE elements more abundant in (G + C) rich regions (Waterston et al. 2002). In general, these repetitive elements are viewed as mainly nonfunctional, "selfish" DNA, but there are some intriguing exceptions. For example, the repeat hypothesis for X inactivation proposes that the enrichment of LINE elements on the X chromosome, relative to the autosomes, acts as a binding signal for the Xist RNA (Lyon 2003).

8.4 Postgenomic Analysis

In the sections below, postgenomic approaches to analyze the expression, function, organization, and evolution of the mouse genome are considered briefly, under 13 headings.

8.4.1 Continued Annotation

There is an ongoing process of annotation of the B6 reference genomic sequence, e.g., improving gene predictions by incorporating manual annotations from the Vega database (Wilming et al. 2007). The genome annotation browsers such as Ensembl also include new data and computational tools to identify previously unpredicted genes (especially those producing noncoding RNAs), define new gene families, and reveal potential regulatory regions. Many projects are contributing to this process, but two worthy of specific mention are: firstly, the International Knockout Mouse Consortium, in which genes chosen for targeted mutagenesis are being annotated manually by the Havana group at the WTSI (http://vega.sanger.ac.uk/info/data/mouse_knock-outs.html) and integrated into Ensembl and secondly, the ENCODE project, which is identifying regulatory features for human and mouse inferred from experimental data and mapping them onto the mouse genome (http://www.genome.gov/10005107) (Chen et al. 2008).

8.4.2 Resequencing

The reference genome sequence is from the C57BL/6J inbred strain, which is one of the most widely used mouse strains. At the time of publication of the essentially "finished" genome

sequence, there remained ~1,200 sequence gaps (Church et al. 2009). There is an ongoing program to improve the public genome assembly, including filling gaps and resolving other problematic genome regions, being led by the Genome Reference Consortium (Table 8.1). This effort has now filled over 300 sequence gaps.

In addition to the B6 reference sequence, a project to produce high-quality draft sequences of 17 other inbred mouse strains was completed recently (http://www.sanger.ac.uk/resources/mouse/genomes/). This has generated an invaluable resource of sequence variation, ranging from over 56 million single nucleotide polymorphisms to a precise mapping of structural differences between strains (Keane et al. 2011; Yalcin et al. 2011). These resources will enable improved mapping of complex or quantitative trait loci (QTL), for example using in silico mapping (Pletcher et al. 2004) and also in the recombinant inbred strains being bred for the Collaborative Cross (Philip et al. 2011). Two of the inbred strains sequenced were derived from mouse subspecies (CAST/EiJ from *Mus musculus castaneus*) and species (SPRET/EiJ from *Mus spretus*) distinct from B6 and so will facilitate the analysis of natural variability in the *Mus* genus.

8.4.3 Transcriptome

Extensive cDNA sequencing has been undertaken in the mouse, systematically surveying gene expression in an enormous range of tissues. A combination of approaches, sampling different aspects of cDNAs, have generated a transcriptional landscape of the mouse, including the sequencing of random ESTs (expressed sequence tags) (Marra et al. 1999), serial analysis of gene expression (SAGE) (Yamamoto et al. 2001), cap analysis of gene expression (CAGE) (Carninci 2006), and full-length cDNA (Okazaki et al. 2002). With each technological advance, the mouse transcriptome has revealed numerous surprising findings, e.g., the abundance of long noncoding RNAs (Guttman et al. 2009; Ozsolak and Milos 2010), including and antisense pairs

(Carninci 2007) and the complexity of alternative splicing (Lee and Wang 2005).

8.4.4 Microarray Analysis

Expression profiling using microarrays has been applied extensively in the mouse, both with spotted cDNA (Hamatani et al. 2004; VanBuren et al. 2002) and oligonucleotide arrays (Cui et al. 2007; Granville and Dennis 2005). Microarray expression profiling remains a powerful technique and the legacy of data ensures comparisons with new experiments are productive, but as costs reduce it may soon be superseded by RNA-seq (Ozsolak and Milos 2010). This ultrahigh throughput sequencing method is highly sensitive, quantitative, and able to detect previously unknown transcripts.

8.4.5 Expression Analysis: Spatial and Temporal Patterns

Determining when and where genes are expressed is a critical step in predicting gene function. A fundamental requirement in such studies is an anatomical atlas; the reference work for the adult mouse (Cook 1965) is also available in electronic form (http://www.informatics.jax.org/cookbook/) and those for the developing mouse embryo and fetus (Kaufman 1992; Theiler 1989) underpin the Edinburgh Mouse Atlas Project (EMAP: http://www.emouseatlas.org/emap/home.html). These atlases typically combine anatomical drawings, photographs, and microphotographs using only nonspecific chemical dyes, but are essential in defining the location of gene expression. An extension of EMAP is the Atlas of Gene Expression (EMAGE: http://www.emouseatlas.org/emage/home.php), which combines in situ hybridization, protein immunohistochemistry, and transgenic reporter data for the mouse. A further valuable resource that integrates spatial and nonspatial expression data both from individual investigators and large-scale projects is the Gene Expression Database (GXD), maintained at the Jackson Laboratory by

the Mammalian Genome Informatics team (http://www.informatics.jax.org/expression.shtml). There have been two large-scale systematic efforts to characterize spatial expression patterns of mouse genes using in situ hybridization: the EUREXPRESS project, that characterized the developing mouse at embryonic day 14.5 (Diez-Roux et al. 2011) and the Allen Brain atlas, that focused on the adult mouse brain (Jones et al. 2009). An alternative approach to localizing gene expression was taken recently by Belgard and colleagues, in which RNA-seq was used to detect and quantify transcripts in the dissected layers of the mouse somatosensory cortex; this allowed them to identify candidate alternatively spliced transcripts that are differentially expressed across layers (Belgard et al. 2011).

8.4.6   Functional Analysis: Gene Replacement or Deletion

Targeted mutagenesis by homologous recombination is one of the great successes of mouse genetics in the twentieth century, allowing scientists to test the effect of gene inactivation or deletion on mammalian physiology and evaluate the likelihood that the orthologous human gene is associated with similar phenotypes or diseases (Mak 2007). Initially, the technology was exploited in individual laboratories, but more recently there have been systematic programs, e.g., KOMP, EUCOMM, and NorCOMM, now coordinated as the International Knockout Mouse Consortium (Collins et al. 2007). Their aim is to generate at least one knockout for every gene in the mouse genome and make available to the scientific community a catalog of the specific embryonic stem cell lines. A conventional knockout will produce a null allele; this is a first step in determining function, but in some cases may appear uninformative if the mutant mouse exhibits no gross phenotype, yet this apparent lack of abnormality may have several possible causes (Barbaric et al. 2007). One obvious possibility is that the chosen phenotype assay (s) are insufficiently sensitive, specific, or comprehensive; Lewis Wolpert, when told that a knockout mouse had no phenotype, famously

said "I say, have you taken your mice to the opera? Can they still tell Wagner from Mozart?" (Wolpert and Garcia-Bellido 1998). A further possibility is that the mutant mice die at such an early stage of development that, although the specific gene is clearly essential, the function of the gene product in later life (if any) is obscured. In order to study gene function, therefore, we require multiple alleles; e.g., conditional knockouts, in which one can deliberately choose a tissue or developmental stage at which a gene is inactivated. This has been achieved using a targeted recombination technology "borrowed" from various bacteriophages, e.g., the CRE/loxP recombinase system from P1 phage (Branda and Dymecki 2004; Gu et al. 1993).

8.4.7   Functional Analysis: RNAi Knockdowns

RNA interference (RNAi) is a method of reducing gene function using short double-stranded RNA (dsRNA) molecules, discovered in the nematode worm, *C. elegans* (Chap. 2), that is now used in many other organisms, including mice (Hitz et al. 2009). This method has not been taken up to a large extent in mice, due in part to difficulties in delivering bioactive dsRNA to cells in whole animals. However, some exciting biology has resulted, e.g., in the fields of type 1 diabetes (Kissler et al. 2006), transcriptional network interactions during infection (Amit et al. 2009), tumor-suppressor genomics (Zender et al. 2008), and neurobiology (Bai et al. 2003; Thakker et al. 2005). New ways to deliver plasmid DNA encoding dsRNAs in mice would be needed to allow experiments similar to those done in *C. elegans*, where simply feeding them bacteria carrying plasmids expressing shRNA (Timmons et al. 2001) is highly effective.

8.4.8   Functional Analysis: Random Mutagenesis

Spontaneous mutation in mice has generated many interesting and valuable abnormal traits, it was soon realized however, that it was possible to

accelerate this random process using various irradiation, chemical, or biological treatments (reviewed in more detail by Flaherty and colleagues 1998). The mutagenic potential of high-energy radiation had been described first in Drosophila (Muller 1927), but was soon used to produce mutations in mice. Various chemicals have been used successfully as mutagens in mice, e.g., ethylmethane sulfonate (EMS) and chlorambucil, but that most widely used is ethyl-nitroso-urea (ENU). The main reasons for the predominance of ENU are that it is probably the most potent known mutagen, introducing about 1 functional sequence change per locus, per 750 progeny, and because it produces, almost exclusively, point mutations (Hitotsumachi et al. 1985). ENU mutagenesis produces a wide variety of functional, which can include null (loss), hypomorphic (decreased), (increased), and even neomorphic (functional gain). This variety of alleles can tell one something new about function, not previously suspected about a gene based on simple null alleles, or help in dissecting out different aspects of function of a single protein, e.g., the Ikaros gene (Papathanasiou et al. 2003), where a null allele has multiple, pleiotropic effects or the cytoplasmic dynein heavy chain 1 (*Dnchc1*) gene, where animals homozygous for the null allele die at an early stage of embryogenesis and heterozygotes show no obvious abnormalities (Hafezparast et al. 2003). Point mutations produced by ENU in mice mirror closely the spectrum of common human mutations, but it is still relatively time consuming to determine the gene affected. Developments in methods for sequence capture (Olson 2007) and next generation DNA sequencing (Bentley et al. 2008; Wheeler et al. 2008) make this obstacle less important.

There are also various biological mutagenesis methods that depend upon random insertion of novel DNA sequences that disrupt gene function and produce abnormal traits, e.g., based on retroviruses (Soriano et al. 1987) or transposable elements (Carlson and Largaespada 2005; Ivics et al. 2009). The Sleeping Beauty transposon, which was derived from a fish, in combination with the Tc1/MARINER transposase gene, has been adapted for insertional mutagenesis in mice. These methods have lower mutagenesis efficiency than ENU, but have the advantage of easier gene identification.

8.4.9 Functional Analysis: Genetic Reference Populations

The concept of mouse genetic reference populations (GRP) (Argmann et al. 2005) covers the range from sets of the conventional inbred strains described earlier to the more sophisticated set of recombinant inbred (RI) strains known as the Collaborative Cross (CC) strains (Churchill et al. 2004). The CC strains are being produced by combining the genomes of eight genetically diverse founder strains (A/J, C57BL/6J, 129S1/SvImJ, NOD/LtJ, NZO/HiLtJ, CAST/EiJ, PWK/PhJ, and WSB/EiJ) and then inbreeding for at least 23 generations. The ultimate goal of the CC is to breed about 1,000 RI lines, but even with 66 lines, it is possible to map QTL to a resolution of <1.2 Mb (Durrant et al. 2011). These strains will be powerful tools in dissecting complex traits, particularly because the eight founder strains were included in those chosen for resequencing (Yalcin et al. 2011) and so by inference, there will be complete genomic sequence information for all the CC strains.

8.4.10 Functional Analysis: Systematic Phenotyping

The mouse clinic concept (Gailus-Durner et al. 2005; Brown et al. 2006) grew from the idea of the broad health check carried out on a human being at a hospital or general health practice. Humans may present with vague symptoms and so a broad range of different tests is often conducted to aid the clinician in reaching a differential diagnosis. This is a systematic, explorative approach, rather than the hypothesis-driven approach that has been typical of many labs investigating knockout mice. The guiding principle is that our understanding of mammalian gene function is imperfect and that the most unbiased method of detecting novel phenotypes in mutant

mice is to subject animals to a broad range of assays, covering all physiological systems. The success of this approach is shown by the detection of many novel phenotypes, even in conventional inbred mouse strains that have been studied intensively for many years (http://www.europhenome.org/).

8.4.11 Interactome and Gene Networks

Within the past decade, high-throughput genomic technologies have revolutionized biomedical science. The genetic, expression or protein data generated by these technologies, when combined with functional annotations, allows inference of predicted network interactions and the definition of subsets of molecules forming interactome(s). These approaches were developed first in simple species such as yeast (Goll and Uetz 2006), but have since been expanded to humans, mice, and other more complex organisms.

The interactomes of the hair cells of the inner ear and photoreceptor cells of the retina include dynamic Usher syndrome protein complexes in mouse and man (Kremer et al. 2006). Spatiotemporal expression, immuno-histochemistry, and immuno-electronmicroscopy in wild-type and mutant mice have contributed significantly to dissecting the interactions of the Usher complexes in stereocilia.

A transcriptional network mediating mouse dendritic cell responses to pathogen-associated molecules has been constructed using a combination of expression profiling and targeted knockdown using shRNA (Amit et al. 2009).

8.4.12 Proteomics and Structural Genomics

Mouse proteomics has lagged behind studies in humans, yet there have been numerous valuable investigations, e.g., of proteomic alterations in disease or disease models (Madan and Amar 2008; Nath et al. 2009; Zhai et al. 2009). Furthermore, Prokai and colleagues have demonstrated that "shotgun" proteomics for protein abundance profiling from mouse tissue is practical using liquid chromatography and mass spectrometry (Prokai et al. 2009).

The Protein Structure Initiative aims at the determination of the 3D structure of all proteins, predominantly from humans, but also some from mouse (http://www.nigms.nih.gov/Research/FeaturedPrograms/PSI/) (Markley et al. 2009).

8.4.13 Comparative Genomics

One of the motivating factors behind the study of mouse genetics and genomics has been the expectation that the fundamental biological similarity between humans and mice could inform a better understanding of human health and disease. It has become clear that this similarity extends to the molecular level: if we align the mouse and human genomes, 15,187 protein-coding genes have simple 1:1 orthologous relationships, corresponding to about 80 % of human genes (Church et al. 2009); furthermore, these genes tend to be clustered into blocks of conserved sequence. Indeed, in the process of finishing the mouse genome, one approach that was used to help detect BAC contig overlaps was to use the syntenic human sequence as a virtual probe (Church et al. 2009; Denny et al. 2001).

We can also think of comparative genomics as a way of analyzing the mouse C57BL/6J reference sequence, using the recently acquired data from four wild-derived inbred strains CAST/EiJ, WSB/EiJ, PWK/PhJ, and SPRET/EiJ, which include the ancestors of the common lab strains and are representative of the *Mus musculus castaneus, Mus musculus domesticus, Mus musculus musculus,* and *Mus spretus* taxa, respectively (Keane et al. 2011). The evolutionary divergence between the laboratory mouse and *Mus spretus* is about 1.5 million years, so it will be fascinating to perform de novo assemblies on these sequences and then compare them with the B6 genome. The three *Mus musculus* subspecies, *castaneus, musculus,* and *domesticus* will interbreed and their F_1 hybrid offspring are fertile, whereas only the female F_1 hybrid offspring of

crosses with *Mus spretus* are fertile. Some of the molecular bases of this difference are understood (Dejager et al. 2009), yet it is likely that comparative genome analysis will give further clues to the causes of hybrid infertility and insights into the process of speciation.

An excellent comparison of rat, mouse, and human genome maps is given in Chap. 9, but Szpirer and Goran do not extend their comparisons to include the newest mouse genome sequences. Keane et al. (2011) sequenced 17 inbred mouse strains and noted that some of the sequences in the non-B6 strains failed to align with the B6 reference genome, yet could be aligned with rat or rabbit genomes. The authors do not describe the content of these "missing" sequences.

Conclusion

A major challenge for the biologists of the twenty-first century is determining functions for the 22,000–37,000 genes in the human genome. One of the tools we can use to help us rise to this challenge is the study of "model" organisms such as mice. One might believe, now that the sequences for human and mouse genomes are in hand, that determining gene function would be a simple, albeit time consuming, matter of studying, one at a time, the phenotype of mice in which a single gene has been inactivated. It is clear, however, that this will give only a partial picture of gene function, because of the structural and functional complexity of mammalian genes. A large proportion of genes are expressed as multiple transcripts, produced from distinct promoters, or by alternative splicing. Genes are subject to spatial and temporal regulation, e.g., activated during embryonic development, then switched off and later reactivated only in specific adult tissue(s). Additional complexity is imposed by post-translational modification of proteins modulating function. Furthermore, some genes are functionally redundant, i.e., multiple genes encode similar proteins, so that deletion of only one of the genes may not produce a phenotype. And what about RNA-coding genes? We need therefore, other, subtler, alleles such as point mutations, conditional knockouts, and ways to partially reduce gene function, such as RNAi, as described earlier. The genetic and functional genomic "toolbox" in the mouse has grown extensively in the last two decades and as systematic mutagenesis and phenotyping are combined, we grow ever closer to achieving that goal of determining function for every mouse gene and so, by inference, for their human equivalent.

Glossary

Congenic strains In which selected chromosomal regions from one inbred strain have been introgressed onto the genetic background of a distinct strain.

Conplastic strains Strains in which the mitochondrial genome from one strain has been transferred onto the nuclear genome of another.

Consomic strains Special case of congenic strain, in which a whole chromosome from one inbred strain has been bred onto the genetic background of a distinct strain.

Heterogeneous stocks Genetically heterogeneous stock of mice descended from a few, e.g., eight, inbred progenitor strains and derived from a pseudorandom breeding scheme over a large (~50) number of generations.

Inbred strains Produced under a strict regime of at least 20 generations of sister–brother or equivalent matings—these mice will have minimal or no heterozygous loci.

Mutant stocks/strains A mouse stock or inbred strain in which any permanent alteration in the DNA sequence, including chromosomal rearrangement or point mutation, has occurred.

Recombinant congenic strains Strains produced by intercrossing two inbred strains, then back-crossing to one of the parental strains for a few generations (typically ~2–3 generations), then inbred, without any selection.

Recombinant inbred strains A set of strains derived by the mating of individuals from the F2 generation of a cross of two inbred strains and from each subsequent generation in accordance with a strict routine of inbreeding.

Wild-derived strains Inbred strains derived from wild-caught mice, e.g., MSM/Ms, which is derived from a Japanese wild mouse, *Mus musculus molossinus*.

References

Amit I, Garber M, Chevrier N, Leite AP, Donner Y, Eisenhaure T, Guttman M, Grenier JK, Li W, Zuk O, Schubert LA, Birditt B, Shay T, Goren A, Zhang X, Smith Z, Deering R, McDonald RC, Cabili M, Bernstein BE, Rinn JL, Meissner A, Root DE, Hacohen N, Regev A (2009) Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. Science 326:257–263

Argmann CA, Chambon P, Auwerx J (2005) Mouse phenogenomics: the fast track to "systems metabolism". Cell Metab 2:349–360

Avner P, Heard E (2001) X-chromosome inactivation: counting, choice and initiation. Nat Rev Genet 2:59–67

Avner P, Bruls T, Poras I, Eley L, Gas S, Ruiz P, Wiles MV, Sousa-Nunes R, Kettleborough R, Rana A, Morissette J, Bentley L, Goldsworthy M, Haynes A, Herbert E, Southam L, Lehrach H, Weissenbach J, Manenti G, Rodriguez-Tome P, Beddington R, Dunwoodie S, Cox RD (2001) A radiation hybrid transcript map of the mouse genome. Nat Genet 29:194–200

Bai J, Ramos RL, Ackman JB, Thomas AM, Lee RV, LoTurco JJ (2003) RNAi reveals doublecortin is required for radial migration in rat neocortex. Nat Neurosci 6:1277–1283

Barbaric I, Miller G, Dear TN (2007) Appearances can be deceiving: phenotypes of knockout mice. Brief Funct Genomic Proteomic 6:91–103

Belgard TG, Marques AC, Oliver PL, Abaan HO, Sirey TM, Hoerder-Suabedissen A, Garcia-Moreno F, Molnar Z, Margulies EH, Ponting CP (2011) A transcriptomic atlas of mouse neocortical layers. Neuron 71:605–616

Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DM, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Boodhun A, Brennan JS, Bridgham JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC, Carter NP, Castillo N, Chiara ECM, Chang S, Neil Cooley R, Crake NR, Dada OO, Diakoumakos KD, Dominguez-Fernandez B, Earnshaw DJ, Egbujor UC, Elmore DW, Etchin SS, Ewan MR, Fedurco M, Fraser LJ, Fuentes Fajardo KV, Scott Furey W, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM, Ho JT, Horgan AM, Hoschler K, Hurwitz S, Ivanov DV, Johnson MQ, James T, Huw Jones TA, Kang GD, Kerelska TH, Kersey AD, Khrebtukova I, Kindwall AP, Kingsbury Z, Kokko-Gonzales PI, Kumar A, Laurent MA, Lawley CT, Lee SE, Lee X, Liao AK, Loch JA, Lok M, Luo S, Mammen RM, Martin JW, McCauley PG, McNitt P, Mehta P, Moon KW, Mullens JW, Newington T, Ning Z, Ling Ng B, Novo SM, O'Neill MJ, Osborne MA, Osnowski A, Ostadan O, Paraschos LL, Pickering L, Pike AC, Chris Pinkard D, Pliskin DP, Podhasky J, Quijano VJ, Raczy C, Rae VH, Rawlings SR, Chiva Rodriguez A, Roe PM, Rogers J, Rogert Bacigalupo MC, Romanov N, Romieu A, Roth RK, Rourke NJ, Ruediger ST, Rusman E, Sanches-Kuiper RM, Schenker MR, Seoane JM, Shaw RJ, Shiver MK, Short SW, Sizto NL, Sluis JP, Smith MA, Ernest Sohna Sohna J, Spence EJ, Stevens K, Sutton N, Szajkowski L, Tregidgo CL, Turcatti G, Vandevondele S, Verhovsky Y, Virk SM, Wakelin S, Walcott GC, Wang J, Worsley GJ, Yan J, Yau L, Zuerlein M, Mullikin JC, Hurles ME, McCooke NJ, West JS, Oaks FL, Lundberg PL, Klenerman D, Durbin R, Smith AJ (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456:53–59

Branda CS, Dymecki SM (2004) Talking about a revolution: the impact of site-specific recombinases on genetic analyses in mice. Dev Cell 6:7–28

Brown SDM, Hancock JM, Gates H (2006) Understanding mammalian genetic systems: the challenge of phenotyping in the mouse. PLoS Genet 2:e118

Burgio G, Szatanik M, Guenet J-L, Arnau M-R, Panthier J-J, Montagutelli X (2007) Interspecific recombinant congenic strains between C57BL/6 and mice of the mus spretus species: a powerful tool to dissect genetic control of complex traits. Genetics 177:2321–2333

Cai WW, Chow CW, Damani S, Gregory SG, Marra M, Bradley A (2001) An SSLP marker-anchored BAC framework map of the mouse genome. Nat Genet 29:133–134

Carlson CM, Largaespada DA (2005) Insertional mutagenesis in mice: new perspectives and tools. Nat Rev Genet 6:568–580

Carninci P (2006) Tagging mammalian transcription complexity. Trends Genet 22:501–510

Carninci P (2007) Constructing the landscape of the mammalian transcriptome. J Exp Biol 210:1497–1506

Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G, Sung WK, Clarke ND, Wei CL, Ng HH (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 133:1106–1117

Church DM, Goodstadt L, Hillier LW, Zody MC, Goldstein S, She X, Bult CJ, Agarwala R, Cherry JL, DiCuccio M, Hlavina W, Kapustin Y, Meric P, Maglott D, Birtle Z, Marques AC, Graves T, Zhou S, Teague B, Potamousis K, Churas C, Place M, Herschleb J, Runnheim R, Forrest D, Amos-Landgraf J, Schwartz DC, Cheng Z, Lindblad-Toh K, Eichler EE, Ponting CP, The Mouse Genome Sequencing C (2009) Lineage-specific biology revealed by a finished genome assembly of the mouse. PLoS Biol 7:e1000112

Churchill GA, Airey DC, Allayee H, Angel JM, Attie AD, Beatty J, Beavis WD, Belknap JK, Bennett B, Berretini W, Bleich A, Bogue M, Broman KW, Buck KJ, Buckler E, Burmeister M, Chesler EJ, Cheverud JM, Clapcote S, Cook MN, Cox RD, Crabbe JC, Crusio WE, Darvasi A, Deschepper CF, Doerge RW, Farber CR, Forejt J, Gaile D, Garlow SJ, Geiger H, Gershenfeld H, Gordon T, Gu J, Gu W, de Haan G, Hayes NL, Heller C, Himmelbauer H, Hitzemann R, Hunter K, Hsu HC, Iraqi FA, Ivandic B, Jacob HJ, Jansen RC, Jepsen KJ, Johnson DK, Johnson TE, Kempermann G, Kendziorski C, Kotb M, Kooy RF, Llamas B, Lammert F, Lassalle JM, Lowenstein PR, Lu L, Lusis A, Manly KF, Marcucio R, Matthews D, Medrano JF, Miller DR, Mittleman G, Mock BA, Mogil JS, Montagutelli X, Morahan G, Morris DG, Mott R, Nadeau JH, Nagase H, Nowakowski RS, O'Hara BF, Osadchuk AV, Page GP, Paigen B, Paigen K, Palmer AA, Pan HJ, Peltonen-Palotie L, Peirce J, Pomp D, Pravenec M, Prows DR, Qi Z, Reeves RH, Roder J, Rosen GD, Schadt EE, Schalkwyk LC, Seltzer Z, Shimomura K, Shou S, Sillanpaa MJ, Siracusa LD, Snoeck HW, Spearow JL, Svenson K, Tarantino LM, Threadgill D, Toth LA, Valdar W, de Villena FP, Warden C, Whatley S, Williams RW, Wiltshire T, Yi N, Zhang D, Zhang M, Zou F (2004) The collaborative cross, a community resource for the genetic analysis of complex traits. Nat Genet 36:1133–1137

Collins FS, Rossant J, Wurst W (2007) A mouse for all reasons. Cell 128:9–13

Cook MJ (1965) The anatomy of the laboratory mouse. Elsevier, Amsterdam

Cormier RT, Hong KH, Halberg RB, Hawkins TL, Richardson P, Mulherkar R, Dove WF, Lander ES (1997) Secretory phospholipase Pla2g2a confers resistance to intestinal tumorigenesis. Nat Genet 17:88–91

Cox A, Ackert-Bicknell CL, Dumont BL, Ding Y, Bell JT, Brockmann GA, Wergedal JE, Bult C, Paigen B, Flint J, Tsaih SW, Churchill GA, Broman KW (2009) A new standard genetic map for the laboratory mouse. Genetics 182:1335–1344

Cuenot L (1902) Notes et revues. Arch Zool Exp Gen: xxvii

Cui XS, Li XY, Shen XH, Bae YJ, Kang JJ, Kim NH (2007) Transcription profile in mouse four-cell, morula, and blastocyst: genes implicated in compaction and blastocoel formation. Mol Reprod Dev 74:133–143

Curwen V, Eyras E, Andrews TD, Clarke L, Mongin E, Searle SMJ, Clamp M (2004) The Ensembl automatic gene annotation system. Genome Res 14:942–950

Dejager L, Libert C, Montagutelli X (2009) Thirty years of Mus spretus: a promising future. Trends Genet 25:234–241

Denny P, Bate R, Mallon AM (2001) What use is the human genome for understanding the mouse? Genome Biol 2(11):COMMENT2009

Dietrich W, Katz H, Lincoln SE, Shin HS, Friedman J, Dracopoli NC, Lander ES (1992) A genetic map of the mouse suitable for typing intraspecific crosses. Genetics 131:423–447

Dietrich WF, Miller J, Steen R, Merchant MA, Damron-Boles D, Husain Z, Dredge R, Daly MJ, Ingalls KA, O'Connor ETJA, DeAngelis MM, Levinson KDML, Goodman N, Copeland NG, Jenkins NA, Hawkins TL, Stein L, Page DC, Lander ES (1996) A comprehensive genetic map of the mouse genome. Nature 380:149–152

Diez-Roux G, Banfi S, Sultan M, Geffers L, Anand S, Rozado D, Magen A, Canidio E, Pagani M, Peluso I, Lin-Marq N, Koch M, Bilio M, Cantiello I, Verde R, De Masi C, Bianchi SA, Cicchini J, Perroud E, Mehmeti S, Dagand E, Schrinner S, Nürnberger A, Schmidt K, Metz K, Zwingmann C, Brieske N, Springer C, Hernandez AM, Herzog S, Grabbe F, Sieverding C, Fischer B, Schrader K, Brockmeyer M, Dettmer S, Helbig C, Alunni V, Battaini M-A, Mura C, Henrichsen CN, Garcia-Lopez R, Echevarria D, Puelles E, Garcia-Calero E, Kruse S, Uhr M, Kauck C, Feng G, Milyaev N, Ong CK, Kumar L, Lam M, Semple CA, Gyenesei A, Mundlos S, Radelof U, Lehrach H, Sarmientos P, Reymond A, Davidson DR, Dollé P, Antonarakis SE, Yaspo M-L, Martinez S, Baldock RA, Eichele G, Ballabio A (2011) A high-resolution anatomical atlas of the transcriptome in the mouse embryo. PLoS Biol 9:e1000582

Durrant C, Tayem H, Yalcin B, Cleak J, Goodstadt L, Pardo-Manuel de Villena F, Mott R, Iraqi FA (2011) Collaborative cross mice and their power to map host susceptibility to Aspergillus fumigatus infection. Genome Res 21:1239–1248

Emes RD, Goodstadt L, Winter EE, Ponting CP (2003) Comparison of the genomes of human and mouse lays the foundation of genome zoology. Hum Mol Genet 12:701–709

Esteller M (2011) Non-coding RNAs in human disease. Nat Rev Genet 12:861–874

Evans EP (1996) Standard idiogram. In: Lyon M, Rastan S, Brown S (eds) Genetic variants and strains of the laboratory mouse, 3rd edn. Oxford University Press, Oxford, pp 1446–1448

Flaherty L: Generation, identification, and recovery of mouse mutations. Methods 1998, 14:107–118. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9571070

Gailus-Durner V, Fuchs H, Becker L, Bolle I, Brielmeier M, Calzada-Wack J, Elvert R, Ehrhardt N, Dalke C, Franz TJ, Grundner-Culemann E, Hammelbacher S, Holter SM, Holzlwimmer G, Horsch M, Javaheri A, Kalaydjiev SV, Klempt M, Kling E, Kunder S, Lengger C, Lisse T, Mijalski T, Naton B, Pedersen V, Prehn C, Przemeck G, Racz I, Reinhard C, Reitmeir P, Schneider I, Schrewe A, Steinkamp R, Zybill C, Adamski J, Beckers J, Behrendt H, Favor J, Graw J, Heldmaier G, Hofler H, Ivandic B, Katus H, Kirchhof P, Klingenspor M, Klopstock T, Lengeling A, Muller W, Ohl F, Ollert M, Quintanilla-Martinez L, Schmidt J, Schulz H, Wolf E, Wurst W, Zimmer A, Busch DH, de Angelis MH (2005) Introducing the German mouse clinic: open access platform for standardized phenotyping. Nat Methods 2:403–404

Goll J, Uetz P (2006) The elusive yeast interactome. Genome Biol 7:223

Granville CA, Dennis PA (2005) An overview of lung cancer genomics and proteomics. Am J Respir Cell Mol Biol 32:169–176

Gregorova S, Divina P, Storchova R, Trachtulec Z, Fotopulosova V, Svenson KL, Donahue LR, Paigen B, Forejt J (2008) Mouse consomic strains: exploiting genetic divergence between Mus m. musculus and Mus m. domesticus subspecies. Genome Res 18:509–515

Gu H, Zou Y-R, Rajewsky K (1993) Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. Cell 73:1155–1164

Guenet JL, Bonhomme F (2003) Wild mice: an ever-increasing contribution to a popular mammalian model. Trends Genet 19:24–31

Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N, Bernstein BE, Kellis M, Regev A, Rinn JL, Lander ES (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458:223–227

Hafezparast M, Klocke R, Ruhrberg C, Marquardt A, Ahmad-Annuar A, Bowen S, Lalli G, Witherden AS, Hummerich H, Nicholson S, Morgan PJ, Oozageer R, Priestley JV, Averill S, King VR, Ball S, Peters J, Toda T, Yamamoto A, Hiraoka Y, Augustin M, Korthaus D, Wattler S, Wabnitz P, Dickneite C, Lampel S, Boehme F, Peraus G, Popp A, Rudelius M, Schlegel J, Fuchs H, Hrabe de Angelis M, Schiavo G, Shima DT, Russ AP, Stumm G, Martin JE, Fisher EM (2003) Mutations in dynein link motor neuron degeneration to defects in retrograde transport. Science 300:808–812

Haldane J, Sprunt A, Haldane N (1915) Reduplication in mice. J Genet 5:133–135

Hamatani T, Carter MG, Sharov AA, Ko MS (2004) Dynamics of global gene expression changes during mouse preimplantation development. Dev Cell 6:117–131

Hitotsumachi S, Carpenter DA, Russell WL (1985) Dose-repetition increases the mutagenic effectiveness of N-ethyl-N-nitrosourea in mouse spermatogonia. Proc Natl Acad Sci USA 82:6619–6621

Hitz C, Steuber-Buchberger P, Delic S, Wurst W, Kuhn R (2009) Generation of shRNA transgenic mice. Methods Mol Biol 530:101–129

Hudson TJ, Church DM, Greenaway S, Nguyen H, Cook A, Steen RG, Van Etten WJ, Castle AB, Strivens MA, Trickett P, Heuston C, Davison C, Southwell A, Hardisty R, Varela-Carver A, Haynes AR, Rodriguez-Tome P, Doi H, Ko MS, Pontius J, Schriml L, Wagner L, Maglott D, Brown SD, Lander ES, Schuler G, Denny P (2001) A radiation hybrid map of mouse genes. Nat Genet 29:201–205

Ivics Z, Li MA, Mates L, Boeke JD, Nagy A, Bradley A, Izsvak Z (2009) Transposon-mediated genome manipulation in vertebrates. Nat Methods 6:415–422

Jones AR, Overly CC, Sunkin SM (2009) The Allen brain atlas: 5 years and beyond. Nat Rev Neurosci 10:821–828

Kaufman MH (1992) The atlas of mouse development. Academic, San Diego

Kazazian HH Jr (2004) Mobile elements: drivers of genome evolution. Science 303:1626–1632

Keane TM, Goodstadt L, Danecek P, White MA, Wong K, Yalcin B, Heger A, Agam A, Slater G, Goodson M, Furlotte NA, Eskin E, Nellaker C, Whitley H, Cleak J, Janowitz D, Hernandez-Pliego P, Edwards A, Belgard TG, Oliver PL, McIntyre RE, Bhomra A, Nicod J, Gan X, Yuan W, van der Weyden L, Steward CA, Bala S, Stalker J, Mott R, Durbin R, Jackson IJ, Czechanski A, Guerra-Assuncao JA, Donahue LR, Reinholdt LG, Payseur BA, Ponting CP, Birney E, Flint J, Adams DJ (2011) Mouse genomic variation and its effect on phenotypes and gene regulation. Nature 477:289–294

Keeler CE, Fuji S (1937) The antiquity of mouse variations in the orient. J Hered 28:93–96

Kim UJ, Birren BW, Slepak T, Mancino V, Boysen C, Kang HL, Simon MI, Shizuya H (1996) Construction and characterization of a human bacterial artificial chromosome library. Genomics 34:213–218

Kissler S, Stern P, Takahashi K, Hunter K, Peterson LB, Wicker LS (2006) In vivo RNA interference demonstrates a role for Nramp1 in modifying susceptibility to type 1 diabetes. Nat Genet 38:479–483

Kouri RE, Miller DA, Miller OJ, Dev VG, Grewal MS, Hutton JJ (1971) Identification by quinacrine fluorescence of the chromosome carrying mouse linkage group I in the Cattanach translocation. Genetics 69:129–132

Kozak C, Nichols E, Ruddle FH (1975) Gene linkage analysis in the mouse by somatic cell hybridization: assignment of adenine phosphoribosyltransferase to chromosome 8 and alpha-galactosidase to the X chromosome. Somatic Cell Genet 1:371–382

Kremer H, van Wijk E, Marker T, Wolfrum U, Roepman R (2006) Usher syndrome: molecular links of pathogenesis, proteins and pathways. Hum Mol Genet 15(2):R262–270

Lachmann P (2010) Genetic and cultural evolution: from fossils to proteins, and from behaviour to ethics. Eur Rev 18:297–309

Lee C, Wang Q (2005) Bioinformatics analysis of alternative splicing. Brief Bioinform 6:23–33

Lyon MF (2003) The Lyon and the LINE hypothesis. Semin Cell Dev Biol 14:313–318

Lyon MF, Rastan S, Brown SDM (1996) Genetic variants and strains of the laboratory mouse, 3rd edn. Oxford University Press, Oxford

Madan M, Amar S (2008) Toll-like receptor-2 mediates diet and/or pathogen associated atherosclerosis: proteomic findings. PLoS One 3:e3204

Mak TW (2007) Gene targeting in embryonic stem cells scores a knockout in Stockholm. Cell 131:1027–1031

Markley JL, Aceti DJ, Bingman CA, Fox BG, Frederick RO, Makino S, Nichols KW, Phillips GN Jr, Primm JG, Sahu SC, Vojtik FC, Volkman BF, Wrobel RL, Zolnai Z (2009) The center for eukaryotic structural genomics. J Struct Funct Genomics 10:165–179

Marra M, Hillier L, Kucaba T, Allen M, Barstead R, Beck C, Blistain A, Bonaldo M, Bowers Y, Bowles L, Cardenas M, Chamberlain A, Chappell J, Clifton S, Favello A, Geisel S, Gibbons M, Harvey N, Hill F, Jackson Y, Kohn S, Lennon G, Mardis E, Martin J, Mila L, McCann R, Morales R, Pape D, Person B, Prange C, Ritter E, Soares M, Schurk R, Shin T, Steptoe M, Swaller T, Theising B, Underwood K, Wylie T, Yount T, Wilson R, Waterston R (1999) An encyclopedia of mouse genes. Nat Genet 21:191–194

McCarthy LC, Terrett J, Davis ME, Knights CJ, Smith AL, Critcher R, Schmitt K, Hudson J, Spurr NK, Goodfellow PN (1997) A first-generation whole genome-radiation hybrid map spanning the mouse genome. Genome Res 7:1153–1161

Muller H (1927) Artificial transmutation of the gene. Science 66:84–87

Mural RJ, Adams MD, Myers EW, Smith HO, Miklos GL, Wides R, Halpern A, Li PW, Sutton GG, Nadeau J, Salzberg SL, Holt RA, Kodira CD, Lu F, Chen L, Deng Z, Evangelista CC, Gan W, Heiman TJ, Li J, Li Z, Merkulov GV, Milshina NV, Naik AK, Qi R, Shue BC, Wang A, Wang J, Wang X, Yan X, Ye J, Yooseph S, Zhao Q, Zheng L, Zhu SC, Biddick K, Bolanos R, Delcher AL, Dew IM, Fasulo D, Flanigan MJ, Huson DH, Kravitz SA, Miller JR, Mobarry CM, Reinert K, Remington KA, Zhang Q, Zheng XH, Nusskern DR, Lai Z, Lei Y, Zhong W, Yao A, Guan P, Ji RR, Gu Z, Wang ZY, Zhong F, Xiao C, Chiang CC, Yandell M, Wortman JR, Amanatides PG, Hladun SL, Pratts EC, Johnson JE, Dodson KL, Woodford KJ, Evans CA, Gropman B, Rusch DB, Venter E, Wang M, Smith TJ, Houck JT, Tompkins DE, Haynes C, Jacob D, Chin SH, Allen DR, Dahlke CE, Sanders R, Li K, Liu X, Levitsky AA, Majoros WH, Chen Q, Xia AC, Lopez JR, Donnelly MT, Newman MH, Glodek A, Kraft CL, Nodell M, Ali F, An HJ, Baldwin-Pitts D, Beeson KY, Cai S, Carnes M, Carver A, Caulk PM, Center A, Chen YH, Cheng ML, Coyne MD, Crowder M, Danaher S, Davenport LB, Desilets R, Dietz SM, Doup L, Dullaghan P, Ferriera S, Fosler CR, Gire HC, Glucksmann A, Gocayne JD, Gray J, Hart B, Haynes J, Hoover J, Howland T, Ibegwam C, Jalali M, Johns D, Kline L, Ma DS, MacCawley S, Magoon A, Mann F, May D, McIntosh TC, Mehta S, Moy L, Moy MC, Murphy BJ, Murphy SD, Nelson KA, Nuri Z, Parker KA, Prud-homme AC, Puri VN, Qureshi H, Raley JC, Reardon MS, Regier MA, Rogers YH, Romblad DL, Schutz J, Scott JL, Scott R, Sitter CD, Smallwood M, Sprague AC, Stewart E, Strong RV, Suh E, Sylvester K, Thomas R, Tint NN, Tsonis C, Wang G, Williams MS, Williams SM, Windsor SM, Wolfe K, Wu MM, Zaveri J, Chaturvedi K, Gabrielian AE, Ke Z, Sun J, Subramanian G, Venter JC, Pfannkoch CM, Barnstead M, Stephenson LD (2002) A comparison of whole-genome shotgun-derived mouse chromosome 16 and the human genome. Science 296:1661–1671

Nakagata N (2000) Cryopreservation of mouse spermatozoa mammalian genome. Mamm Genome 11:572–576

Nath AK, Krauthammer M, Li P, Davidov E, Butler LC, Copel J, Katajamaa M, Oresic M, Buhimschi I, Buhimschi C, Snyder M, Madri JA (2009) Proteomic-based detection of a protein cluster dysregulated during cardiovascular development identifies biomarkers of congenital heart defects. PLoS One 4:e4221

Nusbaum C, Slonim DK, Harris KL, Birren BW, Steen RG, Stein LD, Miller J, Dietrich WF, Nahf R, Wang V, Merport O, Castle AB, Husain Z, Farino G, Gray D, Anderson MO, Devine R, Horton LT Jr, Ye W, Wu X, Kouyoumjian V, Zemsteva IS, Wu Y, Collymore AJ, Courtney DF, Tam J, Cadman M, Haynes AR, Heuston C, Marsland T, Southwell A, Trickett P, Strivens MA, Ross MT, Makalowski W, Xu Y, Boguski MS, Carter NP, Denny P, Brown SDM, Hudson TJ, Lander ES (1999) A YAC-based physical map of the mouse genome. Nat Genet 22:388–393

Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, Nikaido I, Osato N, Saito R, Suzuki H, Yamanaka I, Kiyosawa H, Yagi K, Tomaru Y, Hasegawa Y, Nogami A, Schonbach C, Gojobori T, Baldarelli R, Hill DP, Bult C, Hume DA, Quackenbush J, Schriml LM, Kanapin A, Matsuda H, Batalov S, Beisel KW, Blake JA, Bradt D, Brusic V, Chothia C, Corbani LE, Cousins S, Dalla E, Dragani TA, Fletcher CF, Forrest A, Frazer KS, Gaasterland T, Gariboldi M, Gissi C, Godzik A, Gough J, Grimmond S, Gustincich S, Hirokawa N, Jackson IJ, Jarvis ED, Kanai A, Kawaji H, Kawasawa Y, Kedzierski RM, King BL, Konagaya A, Kurochkin IV, Lee Y, Lenhard B, Lyons PA, Maglott DR, Maltais L, Marchionni L, McKenzie L, Miki H, Nagashima T, Numata K, Okido T, Pavan WJ, Pertea G, Pesole G, Petrovsky N, Pillai R, Pontius JU, Qi D, Ramachandran S, Ravasi T, Reed JC, Reed DJ, Reid J, Ring BZ, Ringwald M, Sandelin A, Schneider C, Semple CA, Setou M, Shimada K, Sultana R, Takenaka Y, Taylor MS, Teasdale RD, Tomita M, Verardo R, Wagner L, Wahlestedt C, Wang Y,

Watanabe Y, Wells C, Wilming LG, Wynshaw-Boris A, Yanagisawa M, Yang I, Yang L, Yuan Z, Zavolan M, Zhu Y, Zimmer A, Carninci P, Hayatsu N, Hirozane-Kishikawa T, Konno H, Nakamura M, Sakazume N, Sato K, Shiraki T, Waki K, Kawai J, Aizawa K, Arakawa T, Fukuda S, Hara A, Hashizume W, Imotani K, Ishii Y, Itoh M, Kagawa I, Miyazaki A, Sakai K, Sasaki D, Shibata K, Shinagawa A, Yasunishi A, Yoshino M, Waterston R, Lander ES, Rogers J, Birney E, Hayashizaki Y (2002) Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature 420:563–573

Olson M (2007) Enrichment of super-sized resequencing targets from the human genome. Nat Methods 4:891–892

Ozsolak F, Milos PM (2010) RNA sequencing: advances, challenges and opportunities. Nat Rev Genet 12:87–98

Papathanasiou P, Perkins AC, Cobb BS, Ferrini R, Sridharan R, Hoyne GF, Nelms KA, Smale ST, Goodnow CC (2003) Widespread failure of hematolymphoid differentiation caused by a recessive niche-filling allele of the Ikaros transcription factor. Immunity 19:131–144

Philip VM, Sokoloff G, Ackert-Bicknell CL, Striz M, Branstetter L, Beckmann MA, Spence JS, Jackson BL, Galloway LD, Barker P, Wymore AM, Hunsicker PR, Durtschi DC, Shaw GS, Shinpock S, Manly KF, Miller DR, Donohue KD, Culiat CT, Churchill GA, Lariviere WR, Palmer AA, O'Hara BF, Voy BH, Chesler EJ (2011) Genetic analysis in the collaborative cross breeding population. Genome Res 21:1223–1238

Pletcher MT, McClurg P, Batalov S, Su AI, Barnes SW, Lagler E, Korstanje R, Wang X, Nusskern D, Bogue MA, Mural RJ, Paigen B, Wiltshire T (2004) Use of a dense single nucleotide polymorphism map for in silico mapping in the mouse. PLoS Biol 2:e393

Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP (2010) A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature 465:1033–1038

Prokai L, Stevens SM, Rauniyar N, Nguyen V (2009) Rapid label-free identification of estrogen-induced differential protein expression in vivo from mouse brain and uterine tissue. J Proteome Res 8:3862–3871

Rhodes M, Straw R, Fernando S, Evans A, Lacey T, Dearlove A, Greystrong J, Walker J, Watson P, Weston P, Kelly M, Taylor D, Gibson K, Mundy C, Bourgade F, Poirier C, Simon D, Brunialti ALB, Montagutelli X, Guenet JL, Haynes A, Brown SDM (1998) A high-resolution microsatellite map of the mouse genome. Genome Res 8:531–542

Rowe LB, Nadeau JH, Turner R, Frankel WN, Letts VA, Eppig JT, Ko MS, Thurston SJ, Birkenmeier EH (1994) Maps from two interspecific backcross DNA panels available as a community genetic mapping resource. Mamm Genome 5:253–274 [published erratum appears in Mamm Genome 1994 Jul;5(7):463]

Sharpe J, Lettice L, Hecksher-Sorensen J, Fox M, Hill R, Krumlauf R (1999) Identification of sonic hedgehog as a candidate gene responsible for the polydactylous mouse mutant Sasquatch. Curr Biol 9:97–S91

Silver LM (1995) Mouse genetics: concepts and applications. Oxford University Press, New York

Soriano P, Gridley T, Jaenisch R (1987) Retroviruses and insertional mutagenesis in mice: proviral integration at the Mov 34 locus leads to early embryonic death. Genes Dev 1:366–375

Stassen AP, Groot PC, Eppig JT, Demant P (1996) Genetic composition of the recombinant congenic strains. Mamm Genome 7:55–58

Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Villalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Ketteman M, Madan A, Rodrigues S, Sanchez A, Whiting M, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW, Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U, Smailus DE, Schnerch A, Schein JE, Jones SJ, Marra MA (2002) Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. Proc Natl Acad Sci USA 99:16899–16903

Takada T, Mita A, Maeno A, Sakai T, Shitara H, Kikkawa Y, Moriwaki K, Yonekawa H, Shiroishi T (2008) Mouse inter-subspecific consomic strains for genetic dissection of quantitative complex traits. Genome Res 18:500–508

Takahashi H, Liu C, Paul MW, Philippe MS (2010) Chapter 4—archiving and distributing mouse lines by sperm cryopreservation, IVF, and embryo transfer. Methods Enzymol 476:53–69, Academic

Thakker DR, Natt F, Husken D, van der Putten H, Maier R, Hoyer D, Cryan JF (2005) siRNA-mediated knockdown of the serotonin transporter in the adult mouse brain. Mol Psychiatry 10:782–789

Theiler K (1989) The house mouse: atlas of embryonic development. Springer, New York

Thuan NV, Kishigami S, Wakayama T (2010) How to improve the success rate of mouse cloning technology. J Reprod Dev 56:20–30

Timmons L, Court DL, Fire A (2001) Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans. Gene 263:103–112

VanBuren V, Piao Y, Dudekula DB, Qian Y, Carter MG, Martin PR, Stagg CA, Bassey UC, Aiba K, Hamatani T,

Kargul GJ, Luo AG, Kelso J, Hide W, Ko MS (2002) Assembly, verification, and initial annotation of the NIA mouse 7.4 K cDNA clone set. Genome Res 12:1999–2003

Wagner K-D, Wagner N, Schedl A (2003) The complex life of WT1. J Cell Sci 116:1653–1658

Watanabe T, Tomizawa S, Mitsuya K, Totoki Y, Yamamoto Y, Kuramochi-Miyagawa S, Iida N, Hoki Y, Murphy PJ, Toyoda A, Gotoh K, Hiura H, Arima T, Fujiyama A, Sado T, Shibata T, Nakano T, Lin H, Ichiyanagi K, Soloway PD, Sasaki H (2011) Role for piRNAs and noncoding RNA in de novo DNA methylation of the imprinted mouse Rasgrf1 locus. Science 332:848–852

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Eswara P, Eyras E, Felsenfeld A, Fewell GA, Flicek P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S, Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigo R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, Karolchik D, Kasprzyk A, Kawai J, Keibler E, Kells C, Kent WJ, Kirby A, Kolbe DL, Korf I, Kucherlapati RS, Kulbokas EJ, Kulp D, Landers T, Leger JP, Leonard S, Letunic I, Levine R, Li J, Li M, Lloyd C, Lucas S, Ma B, Maglott DR, Mardis ER, Matthews L, Mauceli E, Mayer JH, McCarthy M, McCombie WR, McLaren S, McLay K, McPherson JD, Meldrim J, Meredith B, Mesirov JP, Miller W, Miner TL, Mongin E, Montgomery KT, Morgan M, Mott R, Mullikin JC, Muzny DM, Nash WE, Nelson JO, Nhan MN, Nicol R, Ning Z, Nusbaum C, O'Connor MJ, Okazaki Y, Oliver K, Overton-Larty E, Pachter L, Parra G, Pepin KH, Peterson J, Pevzner P, Plumb R, Pohl CS, Poliakov A, Ponce TC, Ponting CP, Potter S, Quail M, Reymond A, Roe BA, Roskin KM, Rubin EM, Rust AG, Santos R, Sapojnikov V, Schultz B, Schultz J, Schwartz MS, Schwartz S, Scott C, Seaman S, Searle S, Sharpe T, Sheridan A, Shownkeen R, Sims S, Singer JB, Slater G, Smit A, Smith DR, Spencer B, Stabenau A, Stange-Thomann N, Sugnet C, Suyama M, Tesler G, Thompson J, Torrents D, Trevaskis E, Tromp J, Ucla C, Ureta-Vidal A, Vinson JP, Von Niederhausern AC, Wade CM, Wall M, Weber RJ, Weiss RB, Wendl MC, West AP, Wetterstrand K, Wheeler R, Whelan S, Wierzbowski J, Willey D, Williams S, Wilson RK, Winter E, Worley KC, Wyman D, Yang S, Yang SP, Zdobnov EM, Zody MC, Lander ES (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420:520–562

Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, Gomes X, Tartaro K, Niazi F, Turcotte CL, Irzyk GP, Lupski JR, Chinault C, Song XZ, Liu Y, Yuan Y, Nazareth L, Qin X, Muzny DM, Margulies M, Weinstock GM, Gibbs RA, Rothberg JM (2008) The complete genome of an individual by massively parallel DNA sequencing. Nature 452:872–876

Willingham AT, Orth AP, Batalov S, Peters EC, Wen BG, Aza-Blanc P, Hogenesch JB, Schultz PG (2005) A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. Science 309:1570–1573

Wilming LG, Gilbert JGR, Howe K, Trevanion S, Hubbard T, Harrow JL (2007) The vertebrate genome annotation (Vega) database. Nucl Acids Res 36(Database issue):D753–D760

Wolpert L, Garcia-Bellido A (1998) Debatable issues. Interview by Alain Ghysen. Int J Dev Biol 42:511–518

Yalcin B, Wong K, Agam A, Goodson M, Keane TM, Gan X, Nellaker C, Goodstadt L, Nicod J, Bhomra A, Hernandez-Pliego P, Whitley H, Cleak J, Dutton R, Janowitz D, Mott R, Adams DJ, Flint J (2011) Sequence-based characterization of structural variation in the mouse genome. Nature 477:326–329

Yamamoto M, Wakatsuki T, Hada A, Ryo A (2001) Use of serial analysis of gene expression (SAGE) technology. J Immunol Methods 250:45–66

Zender L, Xue W, Zuber J, Semighini CP, Krasnitz A, Ma B, Zender P, Kubicka S, Luk JM, Schirmacher P, Richard McCombie W, Wigler M, Hicks J, Hannon GJ, Powers S, Lowe SW (2008) An oncogenomics-based in vivo RNAi screen identifies tumor suppressors in liver cancer. Cell 135:852–864

Zhai J, Strom AL, Kilty R, Venkatakrishnan P, White J, Everson WV, Smart EJ, Zhu H (2009) Proteomic characterization of lipid raft proteins in amyotrophic lateral sclerosis mouse spinal cord. FEBS J 276:3308–3323

Zhang Y, Proenca R, Maffei M, Barone M, Leopold L, Friedman JM (1994) Positional cloning of the mouse obese gene and its human homologue. Nature 372:425–432

Zhou S, Herschleb J, Schwartz DC (2007) A single molecule system for whole genome analysis, chapter 9. In: Mitchelson KR (ed) New high throughput technologies for DNA sequencing and genomics, Perspectives in bioanalysis. Elsevier, Amsterdam, pp 265–300