



Genomics of the future: Identification of quantitative trait loci in the mouse

Lorraine Flaherty, Bruce Herron and Derek Symula

Genome Res. 2005 15: 1741-1745

Access the most recent version at doi:[10.1101/gr.3841405](https://doi.org/10.1101/gr.3841405)

References This article cites 38 articles, 11 of which can be accessed free at:
<http://genome.cshlp.org/content/15/12/1741.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

An advertisement banner with a teal background. On the left, the text "CRISPR and RNAi Genetic Screening. Your new superpower." is written in white. In the center, there is a white-bordered box containing the text "LEARN MORE". On the right, there is a photograph of a woman wearing a red and white superhero cape and mask, and a green molecular structure logo with the word "CELLECTA" below it.

To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Cold Spring Harbor Laboratory Press

Genomics of the future: Identification of quantitative trait loci in the mouse

Lorraine Flaherty,¹ Bruce Herron, and Derek Symula

Genomics Institute, Wadsworth Center, Troy, New York 12180, USA

Positional cloning of quantitative trait loci in rodents is a common approach to identify genes involved in complex phenotypes, including genes important to human disease. However, cloning the causative genes has proved to be more difficult than determining their positions. New tools such as genomic sequence, clone libraries, and new genomic-based methods offer new approaches to identify these genes. Here we review how these new tools and approaches will improve our ability to discover the genes important in complex traits.

Identifying genetic loci controlling complex traits (quantitative trait loci or QTLs) is one of the biggest challenges confronting genetics. These genes influence such traits as growth, morphology, and behavior and determine susceptibility and severity for nearly every disease. In particular, QTLs represent a gateway to the genetic factors controlling common, non-Mendelian diseases, such as heart disease and cancer, and affect many more people than the classic single gene diseases studied in the early days of positional cloning. These diseases have clear genetic components, yet the underlying genes have proven difficult to identify. Unfortunately, these genes are usually subtle in their expression and effect on the general phenotype of the organism; they interact with other genes and environmental effects making them difficult to isolate and they often have a low penetrance. Positional cloning of these QTLs in rodents has proved to be one of the most powerful tools for the functional identification of these genes (Georges 1997). The first step in this procedure is to map these loci to particular chromosomal regions, usually spanning anywhere from 10–40 cm. Over two thousand QTLs have been mapped (<http://www.pubmed.com>); however, the future narrowing of these regions and the subsequent identification of these QTLs are not easy (Flint et al. 2005). Sometimes, these QTLs contain polymorphisms resulting in obvious and deleterious effects on expression and function. More often, sequence differences are difficult to reconcile with the predicted phenotypes or cause expression and/or phenotypic variations that are subtler. Often there is no “smoking gun.” Given the importance of QTLs in common diseases and in pharmacogenomics, recent genetic studies have focused on the identification of these loci.

Genomic sequence and related tools promised to help us find these genes by improving existing approaches and making possible new approaches. Here, we discuss some of the present and future techniques for making identification of these QTLs an easier task. Now that we are several years into the post-genome, we can begin to evaluate some of these strategies. These fall into three categories: (1) making candidate intervals as small as possible, (2) efficiently evaluating large numbers of genes in candidate intervals, and (3) testing candidate genes in a more powerful and efficient manner. These strategies are schemed in Figure 1. It is important to note that, with the possible exception of a targeted knock-in mutation of a QTL, none of the methods yet

described can provide explicit proof of a gene’s involvement in a trait. Rather, each approach tests a gene or (ideally) multiple genes in a different way until the weight of evidence is considered sufficient to prove causation (Abiola et al. 2003).

Detection and localization

In 1992, Dietrich and colleagues (Dietrich et al. 1992) forged a major tool for mapping QTLs—a dense, highly polymorphic linkage map based on molecular markers. Not only did this map become a wonderful tool to map QTLs in standard crosses, but it also made it practical to construct more specialized resources for use in linkage studies. These consisted of new mouse constructs, including large sets of recombinant inbred lines, heterogeneous stocks, consomics, advanced inbred lines, and congenic strains (Kruglyak and Lander 1995a,b; Lander and Kruglyak 1995; Darvasi 1998; Nadeau et al. 2000). For mapping studies, recombinant inbred lines and consomics have been the most useful and have led to the rough localization of a large number of QTLs (see http://www.informatics.jax.org/searches/allele_form.shtml). Perhaps, even more important, the complete C57BL/6 genome and low-coverage sequence for 129/S1, 129X1, A/J, and DBA/2J (Celera/Applera) became available on the Web, and 15 strains of mice are now being resequenced by the Center for Rodent Genetics (<http://www.niehs.nih.gov/crg/>). The single nucleotide polymorphisms (SNPs) revealed by these sequence databases provide an extremely dense map for linkage mapping.

What excited the QTL community more than improved mapping resolution, however, was the potential to use these SNPs for in silico localization using linkage disequilibrium (LD) analysis. Because most inbred mouse strains have a common ancestral heritage and ancestral genotypes account for most of the genetic variation among inbred strains (Wade et al. 2002; Frazer et al. 2004; Yalcin et al. 2004b), QTL genes are likely (though not necessarily) to be found in regions of different ancestry in strains with different phenotypes. Whole-genome association studies might be used to identify ancestral haplotype blocks, i.e., blocks of DNA inherited from a common ancestor, which correlate with different values of a trait among a large collection of inbred strains. This approach is attractive because it can use previously generated phenotypes, such as from the JAX Phenome Database, and requires no further crosses. Thus, in the hypothetical array of strains in Figure 2, sex-linked traits that follow strains A, C, F, and G would be due to the region between markers 5 and 6. For example, Pletcher et al (2004) used nearly 11,000 SNPs to derive inferred haplotypes for 48 strains and applied these haplotypes to

¹Corresponding author.

E-mail flaherty@wadsworth.org; fax (518) 880-1388.

Article and publication are at <http://www.genome.org/cgi/doi/10.1101/gr.3841405>.

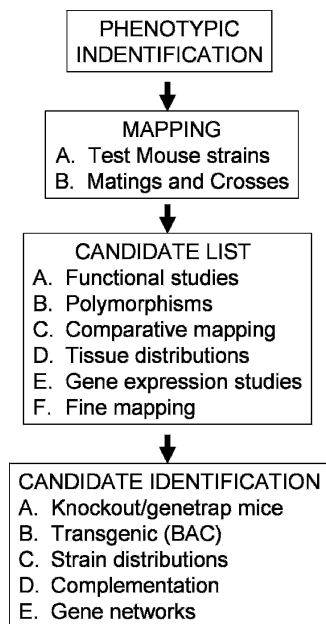


Figure 1. Scheme for the identification of QTLs.

existing phenotype data. Previously identified loci for several monogenic and several quantitative traits showed highly significant associations with the appropriate phenotypes (Pletcher et al. 2004). A similar approach would use haplotypes within a previously mapped QTL to potentially refine the QTL to a smaller haplotype block that associates with the trait (Manenti et al. 2004).

Despite the potential power of LD analysis, the emerging body of studies suggests major challenges to its application in mice (Frazer et al. 2004; Yalcin et al. 2004a). Complete resequencing has revealed greater detail than the whole genome shotgun reads used initially. Haplotype blocks in mice are not consistent among inbred strains; a given region may have different boundaries in different strains. There also are concerns that in silico analysis in mice may be impractical for precise QTL gene localization because of the large number of strains needed for good statistical power even under favorable conditions. It has been estimated that detecting a QTL would require between 40 and 150 strains, with 40 strains needed to determine a major gene controlling the majority of the phenotypic variance (Darvasi 2001). There is still hope for in silico strategies, though. While the structure of the mouse genome is more complex than originally thought, there are still blocks of sequence with clear phylogenetic relationships among the inbred strains (Frazer et al. 2004; Yalcin et al. 2004a). This information still allows for association tests that use a small set of SNPs to represent a given region. Alternatively, we can compare the strain distribution pattern (SDP) of each SNP with the SDP of the trait without regard for the structure of the genome. While this strategy requires complete sequence information for each strain, the limited number of SNP SDPs suggested by results of Yalcin and colleagues may have better statistical power than might be expected (Yalcin et al. 2004b). It seems likely that, given complete sequence coverage for many strains and analysis methods that balance power and resolution, some form of association will be important in QTL localization.

Candidate gene identification

Genomic tools, techniques, and databases have transformed the hunt for candidate genes. Perhaps the greatest effect concerns the initial characterization of the candidate interval. Database searches focusing on the candidate interval have replaced arduous physical mapping and sequencing that, very recently, were state-of-the-art. In addition, multi-species sequence comparisons have identified conserved non-coding regions that are important in gene regulation (Boffelli et al. 2004; Loots et al. 2005). It is now possible, through database searches and multi-species sequence comparisons, to produce a nearly comprehensive list of candidate functional units from Web-based databases (Ahituv et al. 2005). Thus, genomic tools focus the search for causative genes and polymorphisms to those regions most likely to impact phenotype. Evaluating these genes and regulatory regions for appropriate function and changes in sequence and expression remains a challenge but the efficiency of the techniques used has been increased vastly.

Expression profiling would seem to be a perfect tool for evaluating candidates. As for Mendelian traits, an investigator can assay the entire interval to identify genes whose expression correlates with phenotype. The simplest application of expression profiling is to compare two groups of mice with different phenotypes and ask whether any genes in the candidate region are differentially expressed. The clearest reported success using this approach is susceptibility of mice to allergen-induced-airway hyper-responsiveness. Wills-Karp and colleagues (Ewart et al. 2000; Karp et al. 2000) mapped two QTLs for this trait with one differentially expressed gene, hemolytic component *Hc* (previously known as C5) as a prime candidate. Subsequently, differences in the *Hc* sequence were shown to be the cause of this hyper-responsiveness. As simple and attractive as this strategy is, it does not work very often. The investigator must assay the appropriate tissue, cell type, environmental conditions, and developmental stage to evaluate the correct and “bottleneck” phenotype, i.e., the place and time where the QTL makes a difference. Moreover, few cloned QTLs determine expression differences of a magnitude that would be detected by this approach (Flint et al. 2005). Therefore, recent efforts have focused on detecting subtle expression differences and building “gene networks” to take advantage of the tremendous throughput of expression profiling.

One such approach, genetical genomics, treats gene expression levels as quantitative traits and uses a segregating population to map loci that regulate expression of a given gene (Schadt et al. 2003; Bystrykh et al. 2005; Chesler et al. 2005; Hubner et al.

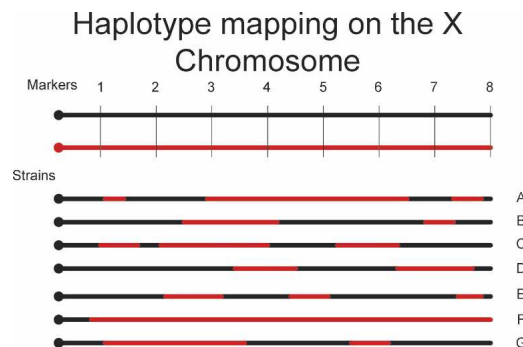


Figure 2. Haplotype block mapping of inbred strains to determine phenotype associated with red but not black chromosome.

2005; Li et al. 2005). Statistical analyses can then determine if a biological QTL co-segregates with an expression QTL (eQTL), which would suggest that the same gene is responsible for both the biological and expression phenotypes. It is possible to detect subtler expression differences by linkage mapping rather than through a simple two-group comparison (Schadt et al. 2003), in part because recombinant inbred lines, for example, will show trait segregation correlating with a particular genotype, thus increasing the statistical power through replication (Bystrykh et al. 2005; Chesler et al. 2005; Hubner et al. 2005).

Cis-acting eQTLs are defined as those that are closely linked to the target gene, while *trans*-acting eQTLs map elsewhere in the genome. A *cis*-acting eQTL is of particular interest for candidate gene identification since it is likely to regulate the closely linked structural gene and implicate it as a strong candidate. *Trans*-acting eQTLs are more difficult to understand since they suggest candidate genes more indirectly. The genes regulated by a given *trans*-acting eQTL probably represent downstream effects of the regulator and, as might be expected, can have similar functions or act in the same pathway (Chesler et al. 2005). Thus, gene networks defined by *trans*-acting eQTLs will broaden our interpretations of quantitative traits and provide a wealth of information for understanding gene regulation and coordinately regulated genes. Several reports recently demonstrated the power of this approach in linking expression of specific genes to behavioral (Chesler et al. 2005) and diabetes-related phenotypes in mice (Doss et al. 2005), hypertension in rats (Hubner et al. 2005), and turnover of mouse hematopoietic stem cells (Bystrykh et al. 2005). Importantly, these candidate genes frequently were polymorphic or mapped to regions of differing haplotypes between the parental strains, and thus were strong candidates for regulatory QTLs.

Like the LD studies described above, genetical genomics is a highly *in silico* method that has increased power to compare phenotypes from existing databases with emerging expression data and sequencing variations so that probable candidates can be identified. Some of this analysis software is available on the WebQTL Web site (<http://www.genenetwork.org/home>). While the authors of these initial studies generated their own expression data, data for other experiments are becoming increasingly available in expression databases such as NCBI GEO (<http://www.ncbi.nlm.nih.gov/geo/>). This approach is a powerful one and is likely to become a common one to use for QTL studies.

Causative gene identification

Once strong candidates are identified, it is crucial to test them. The main problems include the difficulty in predicting the effects of existing polymorphisms on gene expression and function of the candidate gene, in extrapolating these effects to biological impact, and in ascribing phenotypic differences between two strains to one of many genes polymorphic between those strains. In some cases, the primary candidate gene, e.g., *Rgs2* (Yalcin et al. 2004b) in the case of behavior, has not been shown to have a structural defect. In others, the only observed polymorphism may have no clear effect on candidate gene function. For example, the *Apoa2* gene was a strong functional and positional candidate for a QTL controlling plasma HDL level but only one *Apoa2* polymorphism was significantly associated with plasma HDL level when the coding region was examined, a conserved alanine that is changed to valine in high HDL strains (Wang et al. 2004). Furthermore, the common origin of the inbred strains has

contributed to several polymorphisms in a candidate interval and may have nothing causative to do with the expected complex trait phenotype. For most QTLs, investigators must explicitly test new alleles of the candidate gene. They may generate null alleles, complement an existing allele, or, in the most recent twist, find new alleles of the candidate's ortholog in a different species (Lyons et al. 2005; Wang et al. 2005).

Targeted knock-in mutations of a candidate gene remain the gold standard for QTL gene identification (Abiola et al. 2003). A central reason for the utilization of mice in complex trait analysis has been the ability to isolate and genetically manipulate mouse embryonic stem (ES) cells that can be used to create mice with specific gene defects. Historically, generating a targeted mutation in a specific candidate gene was a major undertaking and the resources required to do so on a large scale were unlikely to be found in academic laboratories. Publicly available targeted mutations covered only a fraction of genome until recently. The availability of the mouse genome sequence and the clones used in the public sequencing effort make generating a targeting construct easier (Cotta-de-Almeida et al. 2003). One can identify the target site in a gene using a sequence database and generate homology arms by PCR of a purified BAC template. More sophisticated targeting constructs for conditional knockouts can be made by recombineering (Copeland et al. 2001). In addition, it soon may become unnecessary to build a single targeting construct. Instead, one can use MICER, a library of pre-made targeting vectors for insertional mutations (Adams et al. 2004). In addition, there is an initiative to make and screen a massive library of ES cells that contain a reporter gene in place of each putative mouse open reading frame (Austin et al. 2004). Thus, heterozygous mice created from each of these lines would provide spatial and temporal information about where the gene is expressed, while homozygous null mice could provide insight to gene function *in vivo*. A tiered strategy can also be used for mutant line characterization. Here multiple levels of phenotype ascertainment will be used to funnel specific loci into more focused studies. For example, if heterozygous mice show strong reporter expression in brain, a battery behavioral testing might be warranted for that mutant line (Austin et al. 2004).

Moreover, there is a wealth of genetraps available to the rodent research community. These are mutations that result from the insertion of an expression vector into the target gene. This expression vector contains a reporter gene, a selectable marker, poly(A) site, and translation stop site, but no promoter sequences. The vector is electroporated into mouse ES cells and, if it inserts in a transcribed sequence and produces a fusion transcript composed of part of the endogenous transcribed sequence (i.e., the "trapped" gene), the reporter, and the selectable marker, it will allow these ES cell clones to be selected. The trapped gene then is identified by 5' gene sequencing. The insertion often serves as a hypomorphic or null mutation, since it prematurely terminates transcription of the endogenous gene, and contains a reporter for the expression patterning of the endogenous gene. More recent genetraps feature site-specific recombinase sequences to create conditional mutations (Schnütgen et al. 2005). Several genetraps ES cell libraries have been compiled and over 45,000 are available publicly (see <http://www.sanger.ac.uk/PostGenomics/genetraps/>) covering more than one-third of the genes in the mouse genome (Skarnes et al. 2004). These genetraps libraries are particularly useful to investigators who are not mouse geneticists but wish to test a QTL candidate of interest.

Classical complementation is another powerful method for

gene testing that has been made more practical by genomics. Correlation of the QTL with the presence of a single BAC would greatly refine the search for the causal mutation. BACs have become broadly utilized in functional mouse genomics and are the vector of choice for large scale sequencing because of its superior stability and DNA yield compared with other large insert vectors (Haldi et al. 1994). BACs also broaden the spectrum of candidate gene alleles available for testing and through the availability of libraries from different strains and through recent advances in BAC engineering that may facilitate the use of BAC transgenic mice to test for strain-specific dominant effects. The limitation of this approach is the availability of the proper BAC library and considerably large interval that most QTLs currently cover. BAC libraries from many strains are now publicly available with other strain libraries nearing completion (<http://bacpac.chori.org/>). Mapping of these BACs on the mouse genome will greatly facilitate their utilization.

SNP data from other species make possible a new, powerful candidate gene test, the cross-species comparison. These SNPs make it possible to ask the following question: Do different alleles of the human ortholog (or rat, platypus, etc.) of a candidate mouse gene correlate with phenotypic differences caused by the QTL? This procedure has been successfully used by Wang and colleagues (Wang et al. 2005), who identified *Tnfrsf4*, as a candidate for diet-induced atherosclerosis based on position, expression, and sequence polymorphism. In addition, the authors found a SNP in the human *TNFSF4* gene that was associated with elevated risk of coronary artery disease and myocardial infarction. This approach provides new and independent populations in which to test the original hypothesis. This can be important given the limited polymorphism present in the inbred mouse strains. It further points to the conservation of sequence (though probably not individual polymorphisms) across millions of years of evolutionary distance as a very persuasive demonstration of normal and variant function. However, the failure to recapitulate the effects of a rodent gene in humans (or any other species) is more difficult to interpret, as the results could be attributed to any number of complicating genetic or environmental factors. There is also concern about the low reproducibility of human association studies (Ioannidis 2005). Still, such cross-species comparisons provide powerful evidence of causation when successful and may yet prove to be important.

Mapping the location of the QTL is an extremely important first step in QTL identification. While it is clear that location is not *sufficient* to identify a QTL gene, it is an integral part of the investigation and substantially limits the candidate gene pool. Given the rapid increases in high-throughput evaluation of genes and functional databases and the difficulties in studying genes of small effect in heterogeneous segregating populations, is it possible to robustly identify candidate genes without positional information? Our hope is that approaches such as genetic transcription networks will soon be refined in a manner that is sufficient to “guess” what genes may be involved in a physiological function. The insertion and/or the deletion of these same candidates could then lead to their identification as QTLs that influence function. For example, Soriano and colleagues demonstrated a position-independent approach to identify genetraps clones mutant in genes important in PDGF signaling (Chen et al. 2004). Such position-independent strategies will certainly be used in the future especially when expression profiling becomes reproducible and available for a large number of tissues and strains.

Conclusions

SNP mappings of 15 strains of mice (sponsored by the Center for Rodent Genetics) will only be the beginning. The mouse investigator will be able to compare and contrast different chromosomal regions to identify new genes influencing the fate of QTLs affecting minor functions. It is likely that the contributions from each strain will focus on narrow regions for identification of these QTLs, leading to ~100–500 genes with a known tissue-determining effect. Microarray analyses of gene activity should then narrow these genes down to a short list, functionally distinguishing them from their neighbors. Thus, QTL identifications should become easier tasks and ones that are easily accomplished by the amplification of our current mouse and molecular genetic techniques.

References

- Abiola, O., Angel, J.M., Avner, P., Bachmanov, A.A., Belknap, J.K., Bennett, B., Blankenhorn, E.P., Blizard, D.A., Bolivar, V., Brockmann, G.A., et al. 2003. The nature and identification of quantitative trait loci: A community's view. *Nat. Rev. Genet.* **4**: 911–916.
- Adams, D.J., Biggs, P.J., Cox, T., Davies, R., van der Weyden, L., Jonkers, J., Smith, J., Plumb, B., Taylor, R., Nishijima, I., et al. 2004. Mutagenic insertion and chromosome engineering resource (MICER). *Nat. Genet.* **36**: 867–871.
- Ahituv, N., Prabhakar, S., Poulin, F., Rubin, E.M., and Couronne, O. 2005. Mapping *cis*-regulatory domains in the human genome using multi-species conservation of synteny. *Hum. Mol. Genet.* **14**: 3057–3063.
- Austin, C.P., Battey, J.F., Bradley, A., Bucan, M., Capocchi, M., Collins, F.S., Dove, W.F., Duyk, G., Dymecki, S., Eppig, J.T., et al. 2004. The knockout mouse project. *Nat. Genet.* **36**: 921–924.
- Boffelli, D., Weer, C.V., Weng, L., Lewis, K.D., Shoukry, M.I., Pachter, L., Keys, D.N., and Rubin, E.M. 2004. Intraspecies sequence comparisons for annotating genomes. *Genome Res.* **14**: 2406–2411.
- Bystrykh, L., Weersing, E., Dontje, B., Sutton, S., Pletcher, M.T., Wiltshire, T., Su, A.I., Vellenga, E., Wang, J., Manly, K.F., et al. 2005. Uncovering regulatory pathways that affect hematopoietic stem cell function using ‘genetical genomics’. *Nat. Genet.* **37**: 225–232.
- Chen, W.V., Delrow, J., Corrin, P.D., Frazier, J.P., and Soriano, P. 2004. Identification and validation of PDGF transcriptional targets by microarray-coupled gene-trap mutagenesis. *Nat. Genet.* **36**: 304–312.
- Chesler, E.J., Lu, L., Shou, S., Qu, Y., Gu, J., Wang, J., Hsu, H.C., Mountz, J.D., Baldwin, N.E., Langston, M.A., et al. 2005. Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. *Nat. Genet.* **37**: 233–242.
- Copeland, N.G., Jenkins, N.A., and Court, D.L. 2001. Recombineering: A powerful new tool for mouse functional genomics. *Nat. Rev. Genet.* **2**: 769–779.
- Cotta-de-Almeida, V., Schonhoff, S., Shibata, T., Leiter, A., and Snapper, S.B. 2003. A new method for rapidly generating gene-targeting vectors by engineering BACs through homologous recombination in bacteria. *Genome Res.* **13**: 2190–2194.
- Darvasi, A. 1998. Experimental strategies for the genetic dissection of complex traits in animal models. *Nat. Genet.* **18**: 19–24.
- . 2001. In silico mapping of mouse quantitative trait loci. *Science* **294**: 2423.
- Dietrich, W., Katz, H., Lincoln, S.E., Shin, H.S., Friedman, J., Dracopoli, N.C., and Lander, E.S. 1992. A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics* **131**: 423–447.
- Doss, S., Schadt, E.E., Drake, T.A., and Lusis, A.J. 2005. *Cis*-acting expression quantitative trait loci in mice. *Genome Res.* **15**: 681–691.
- Ewart, S.L., Kuperman, D., Schadt, E., Tankersley, C., Grupe, A., Shubitowski, D.M., Peltz, G., and Wills-Karp, M. 2000. Quantitative trait loci controlling allergen-induced airway hyperresponsiveness in inbred mice. *Am. J. Respir. Cell Mol. Biol.* **23**: 537–545.
- Flint, J., Valdar, W., Shifman, S., and Mott, R. 2005. Strategies for mapping and cloning quantitative trait genes in rodents. *Nat. Rev. Genet.* **6**: 271–286.
- Frazer, K.A., Tao, H., Osoegawa, K., de Jong, P.J., Chen, X., Doherty, M.F., and Cox, D.R. 2004. Noncoding sequences conserved in a limited number of mammals in the SIM2 interval are frequently functional. *Genome Res.* **14**: 367–372.

- Georges, M. 1997. QTL mapping to QTL cloning: Mice to the rescue. *Genome Res.* **7**: 663–665.
- Haldi, M., Perrot, V., Saumier, M., Desai, T., Cohen, D., Cherif, D., Ward, D., and Lander, E.S. 1994. Large human YACs constructed in a rad52 strain show a reduced rate of chimerism. *Genomics* **24**: 478–484.
- Hubner, N., Wallace, C.A., Zimdahl, H., Petretto, E., Schulz, H., Maciver, F., Mueller, M., Hummel, O., Monti, J., Zidek, V., et al. 2005. Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. *Nat. Genet.* **37**: 243–253.
- Ioannidis, J.P. 2005. Why most published research findings are false. *PLoS Med.* **2**: e124.
- Karp, C.L., Grupe, A., Schadt, E., Ewart, S.L., Keane-Moore, M., Cuomo, P.J., Kohl, J., Wahl, L., Kuperman, D., Germer, S., et al. 2000. Identification of complement factor 5 as a susceptibility locus for experimental allergic asthma. *Nat. Immunol.* **1**: 221–226.
- Kruglyak, L. and Lander, E.S. 1995a. A nonparametric approach for mapping quantitative trait loci. *Genetics* **139**: 1421–1428.
- . 1995b. High-resolution genetic mapping of complex traits. *Am. J. Hum. Genet.* **56**: 1212–1223.
- Lander, E. and Kruglyak, L. 1995. Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. *Nat. Genet.* **11**: 241–247.
- Li, H., Lu, L., Manly, K.F., Chesler, E.J., Bao, L., Wang, J., Zhou, M., Williams, R.W., and Cui, Y. 2005. Inferring gene transcriptional modulatory relations: A genetical genomics approach. *Hum. Mol. Genet.* **14**: 1119–1125.
- Loots, G.G., Kneissel, M., Keller, H., Baptist, M., Chang, J., Collette, N.M., Ovcharenko, D., Plajzer-Frick, I., and Rubin, E.M. 2005. Genomic deletion of a long-range bone enhancer misregulates *sclerostin* in Van Buchem disease. *Genome Res.* **15**: 928–935.
- Lyons, M.A., Korstanje, R., Li, R., Sheehan, S.M., Walsh, K.A., Rollins, J.A., Carey, M.C., Paigen, B., and Churchill, G.A. 2005. Single and interacting QTLs for cholesterol gallstones revealed in an intercross between mouse strains NZB and SM. *Mamm. Genome* **16**: 152–163.
- Manenti, G., Galbiati, F., Gianni-Barrera, R., Pettinicchio, A., Acevedo, A., and Dragani, T.A. 2004. Haplotype sharing suggests that a genomic segment containing six genes accounts for the pulmonary adenoma susceptibility 1 (Pas1) locus activity in mice. *Oncogene* **23**: 4495–4504.
- Nadeau, J.H., Singer, J.B., Matin, A., and Lander, E.S. 2000. Analysing complex genetic traits with chromosome substitution strains. *Nat. Genet.* **24**: 221–225.
- Pletcher, M.T., McClurg, P., Batalov, S., Su, A.I., Barnes, S.W., Lagler, E., Korstanje, R., Wang, X., Nusskern, D., Bogue, M.A., et al. 2004. Use of a dense single nucleotide polymorphism map for in silico mapping in the mouse. *PLoS Biol.* **2**: e393.
- Schadt, E.E., Monks, S.A., Drake, T.A., Lusk, A.J., Che, N., Colinayo, V., Ruff, T.G., Milligan, S.B., Lamb, J.R., Cavet, G., et al. 2003. Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422**: 297–302.
- Schnütgen, F., De-Zolt, S., Van Sloun, P., Hollatz, M., Floss, T., Hansen, J., Altschmied, J., Seisenberger, C., Ghyselinck, N.B., Ruiz, P., et al. 2005. Genomewide production of multipurpose alleles for the functional analysis of the mouse genome. *Proc. Natl. Acad. Sci.* **102**: 7221–7226.
- Skarnes, W.C., von Melchner, H., Wurst, W., Hicks, G., Nord, A.S., Cox, T., Young, S.G., Ruiz, P., Soriano, P., Tessier-Lavigne, M., et al. 2004. A public gene trap resource for mouse functional genomics. *Nat. Genet.* **36**: 543–544.
- Wade, C.M., Kulbokas III, E.J., Kirby, A.W., Zody, M.C., Mullikin, J.C., Lander, E.S., Lindblad-Toh, K., and Daly, M.J. 2002. The mosaic structure of variation in the laboratory mouse genome. *Nature* **420**: 574–578.
- Wang, X., Korstanje, R., Higgins, D., and Paigen, B. 2004. Haplotype analysis in multiple crosses to identify a QTL gene. *Genome Res.* **14**: 1767–1772.
- Wang, X., Ishimori, N., Korstanje, R., Rollins, J., and Paigen, B. 2005. Identifying novel genes for atherosclerosis through mouse–human comparative genetics. *Am. J. Hum. Genet.* **77**: 1–15.
- Yalcin, B., Fullerton, J., Miller, S., Keays, D.A., Brady, S., Bhomra, A., Jefferson, A., Volpi, E., Copley, R.R., Flint, J., et al. 2004a. Unexpected complexity in the haplotypes of commonly used inbred strains of laboratory mice. *Proc. Natl. Acad. Sci.* **101**: 9734–9739.
- Yalcin, B., Willis-Owen, S.A., Fullerton, J., Meesaq, A., Deacon, R.M., Rawlins, J.N., Copley, R.R., Morris, A.P., Flint, J., and Mott, R. 2004b. Genetic dissection of a behavioral quantitative trait locus shows that *Rgs2* modulates anxiety in mice. *Nat. Genet.* **36**: 1197–1202.

Web site references

- <http://www.pubmed.com>; PubMed.
- <http://www.genenetwork.org/home>; A set of linked resources for systems genetics.
- http://www.informatics.jax.org/searches/allele_form.shtml; Database of QTLs searchable phenotype, disease, nomenclature, chromosomal location, or allele categories.
- <http://www.niehs.nih.gov/crg/>; Includes description of mouse resequencing project and links to appropriate databases.
- <http://www.ncbi.nlm.nih.gov/geo/>; A gene expression/molecular abundance repository MIAME compliant data submissions, and a curated, online resource for gene expression data browsing, query and retrieval.
- <http://www.sanger.ac.uk/PostGenomics/genetrp/>; Description of the Sanger Institute Gene Trap Resource and searchable database of genetrp mutations produced by the International Gene Trap Consortium.
- <http://bacpac.chori.org/>; Web site run by the BACPAC Resource Center (BPRC) at Children's Hospital Oakland Research Institute in Oakland, California. Includes descriptions of clone libraries and services available.