Published in final edited form as:
*Mamm Genome*. 2014 August ; 25(0): 271–292. doi:10.1007/s00335-014-9523-1.

# Fine mapping QTLs in advanced interbred lines and other outbred populations

**Natalia M. Gonzales**[1] and **Abraham A. Palmer**[1,2]

[1]Department of Human Genetics, University of Chicago, Chicago, IL 60637, USA

[2]Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL 60637, USA

## Abstract

Quantitative genetic studies in model organisms, particularly in mice, have been extremely successful in identifying chromosomal regions that are associated with a wide variety of behavioral and other traits. However, it is now widely understood that identification of the underlying genes will be far more challenging. In the last few years, a variety of populations have been utilized in an effort to more finely map these chromosomal regions with the goal of identifying specific genes. The common property of these newer populations is that linkage disequilibrium spans relatively short distances, which permits fine-scale mapping resolution. This review focuses on advanced intercross lines (AILs) which are the simplest such population. As originally proposed in 1995 by Darvasi and Soller, an AIL is the product of intercrossing two inbred strains beyond the $F_2$ generation. Unlike recombinant inbred strains, AILs are maintained as outbred populations; brother-sister matings are specifically avoided. Each generation of intercrossing beyond the $F_2$ further degrades linkage disequilibrium between adjacent makers, which allows for fine scale mapping of quantitative trait loci (QTLs). Advances in genotyping technology and techniques for the statistical analysis of AILs have permitted rapid advances in the application of AILs. We review some of the analytical issues and available software, including QTLRel, EMMA, EMMAX, GEMMA, TASSEL, GRAMMAR, WOMBAT, Mendel and others.

### Keywords

mice; mixed models; relatedness; behavior; inbred strains

## Introduction

Most heritable traits are polygenic, meaning that many different alleles contribute to the observed phenotype. The loci that influence quantitative traits are called quantitative trait loci (**QTLs**). While there is great interest in identifying the genes that give rise to QTLs for a wide range of heritable traits, this goal has been extremely challenging in both humans and model organisms. Mice have been extensively employed in so-called “forward genetic” studies, which are aimed at identifying QTLs. Mice have also been heavily used in “reverse

Corresponding Author: Abraham A Palmer, 920 E 58th St, CLSC-507D, Chicago, IL 60637, USA. Fax 773 834-0505; Phone 773 834-0525 aap@uchicago.edu.

genetic" studies in which mutant alleles are produced by genetically engineered insertions, deletions, and conditional alleles. Reverse genetic approaches are not the focus of the current review. Instead, we will focus on forward genetic studies with an emphasis on the use of advanced intercross lines (**AILs**). While we emphasize the use of mice, many of our points apply equally well to other model organisms.

QTL mapping studies are comprised of two primary types of data: phenotypes and genotypes. A simple example of a QTL study involves making an $F_2$ cross between two inbred strains, phenotyping the subjects for one or more quantitative traits and then genotyping the subjects at polymorphic markers across the genome. Such data can be analyzed by examining the evidence for association between the phenotype and the genotype at each marker. This basic approach has been used to identify hundreds of QTLs for a wide variety of phenotypes. However, these sorts of studies have seldom led to the identification of individual genes (Flint and Mackay 2009; Parker and Palmer 2011). This is because the QTL regions identified are very large and typically contain hundreds of genes. This review will focus on strategies for narrowing QTL regions, which is a critical prerequisite for gene identification.

In the past, it was desirable to use simple crosses like $F_2$s and backcrosses to map QTLs because they limited the amount of genotyping that was required. Only one or two recombinations per chromosome occur in these populations, so only a handful of markers per chromosome are needed. Unfortunately, this lack of recombination is exactly the factor that has prevented such studies from identifying smaller QTLs. As a result, large QTLs were more finely mapped using congenic strains. Congenic strains are created by introgressing a small interval from one of the two inbred strains used to map the QTL onto the other strain by repeated rounds of backcrossing in conjunction with maker-based selection for the putative QTL region. This approach is extremely time-, animal- and labor-intensive and has only occasionally been brought to fruition (Legare et al. 2000; Stylianou et al. 2004; Shirley et al. 2004; Park et al. 2005; Yazbek et al. 2011; Bryant et al. 2012; Buchner et al. 2012).

In addition to the difficulty of these studies, it is now clear that they are based on an unreliable premise; namely, that each locus identified using an $F_2$ cross is caused by a single polymorphism. In fact, it is often true that apparently large QTLs detected in an $F_2$ crosses are due to multiple smaller QTLs that happen to be clustered in a single region (e.g. Cheng et al. 2010; Parker et al. 2013b). In other cases, $F_2$ crosses may fail to identify true QTLs. For example, if two or more QTLs with opposite effects on the phenotype are closely linked, they may cancel out one another's effects on the phenotype. It is also possible that gene-gene interactions are sometimes required such that when a QTL region is broken into smaller pieces, none of them will affect the phenotype individually (Bryant et al. 2012). Due to all of these limitations, $F_2$ populations are being supplanted by other approaches.

While not the focus of this review, we wish to briefly mention the existence of recombinant inbred (**RI**) strains and the related collaborative cross (**CC**), which were recently reviewed by Flint and Eskin (2012). RI strains are generated by inbreeding the progeny of an $F_2$ cross (a closely related variant, termed recombinant inbred advanced intercross lines (**RIAIL**), are produced by inbreeding AILs rather than $F_2$ mice (Peirce et al. 2004; Rockman and

Kruglyak 2008)). When the cost of genotyping was high, the use of RI panels conferred a tremendous advantage because each strain only needed to be genotyped once. In addition, QTLs could be mapped to somewhat smaller regions than in an $F_2$ cross because new recombination events were produced in the generations required for inbreeding. Finally, huge amounts of data can be generated from RI strains across time and space; such data have been collected and are publicly available (www.genenetwork.org). The CC is similar to an RI but is created by crossing eight rather than two inbred strains. This distinction provides more genetic diversity and makes it possible to take advantage of recombination that occurred prior to the inbreeding of the eight founder strains. Thus, the CC has the potential to provide better mapping resolution than RI strains. However, both RI strains and the CC require large panels of strains to obtain sufficient power to detect QTLs. The largest mouse RI panel currently consists of about 100 strains derived from a cross between C57BL/6J and DBA/2J. While the initial plans for the CC were to make 1,000 strains, fewer than 70 are currently fully (or almost fully) inbred and available for distribution. Despite a continued enthusiasm for using RI strains and the CC to map QTLs, many labs have begun to use outbred strains in an effort to obtain superior resolution and power.

Outbred populations and specifically AILs will be the focus of this review. First, we will explain how an AIL population is generated and maintained. We then consider the genetic properties of AILs and other outbred populations as they relate to QTL mapping precision and power. We also describe genotyping strategies for outbred populations and discuss the challenges of mapping QTLs in the presence of complex genetic relationships. We then explain how mixed linear models can be used to address confounding due to complex relationships. Finally, we briefly review a selection of freely available software that has been used to address issues of relatedness using mixed models.

## Advanced intercross lines (AILs)

In the past decade, the cost of genotyping has rapidly decreased, eroding one advantage of $F_2$s, RIs and the CC. Accordingly, many researchers have shifted their attention towards outbred populations, which accumulate new recombinations in each generation, offering the potential to map QTLs with greater precision.

The simplest type of outbred population is an AIL. AILs were first proposed by Darvasi and Soller (1995) as a tool for fine mapping within established QTL regions; at the time genome-wide genotyping of AILs was too expensive to seriously contemplate. An AIL is produced by continuing to cross two inbred strains beyond the $F_2$ generation. Each successive generation of outbreeding leads to additional recombination, which is desirable because these recombinations improve QTL mapping precision. In mice, about four generations of breeding can be performed per year, so that a tenth generation AIL might be produced in less than three years. Although most AILs have been generated using rodents, AILs have been produced in other organisms (see Table 1).

In an $F_2$ population, all individuals are siblings, meaning that on average they should be equally related to one another. Beginning with the $F_3$ generation the relationships among individuals in an AIL are more complex. Some pairs of individuals will be full siblings,

whereas others will be more or less distant cousins. Complex pedigrees require a more sophisticated analysis because there are correlations among the phenotypes (more closely related individuals are more phenotypically similar) and among the genotypes (the probability of having the same genotype at a given marker is higher among more closely related individuals). As a result, markers that are not linked to any QTLs may still be associated with the phenotype because they provide information about relatedness among individuals (McPeek 2000; Peirce et al. 2008). This problem can be addressed by using more sophisticated analytical approaches that we will discuss later in this review.

Producing an AIL requires attention to both the size of the population and the breeding strategy. An AIL breeding population should be as large as possible to minimize genetic drift, inbreeding and relatedness among individuals (Darvasi and Soller 1995). In model organisms that are less expensive to maintain (e.g. insects), it is easy to support a large population; however, the expense of maintaining rodents typically limits populations to 50–100 breeder pairs. The simplest breeding scheme is to randomly mate individuals to produce the next generation. A somewhat better strategy is to avoid mating full siblings; better still is to avoid mating individuals that share a common grandparent. The optimal strategy is to use pedigree or genotype information to minimize the level of relatedness among all breeder pairs. This approach is similar to captive breeding programs used by zoos to maintain rare and endangered species (Fernandez 2005; Putnam and Ivy 2013). We have implemented these methods for using in an AIL (www.palmerlab.org). Familial contributions to the next generation are typically limited to one female and one male from each litter. This preserves genetic diversity by doubling the effective population size relative to its actual size, thus helping to prevent the fixation of parental alleles (Chia et al. 2005; Lawson et al. 2011a). For relatively inexpensive model organisms, random breeding is often the most practical strategy because any suboptimal pairings are easily overcome by increasing the population size.

## QTL mapping precision and power

The motivation for using AILs and other outbred populations is to improve QTL mapping precision, which is defined as the average width of the confidence interval for a QTL. Unfortunately, gains in precision come at the expense of power. Power is defined as the likelihood of detecting a QTL of a certain effect size at a given significance threshold. The trade-off between power and precision is due to the fact that a pair of markers at a given genetic distance will become increasingly uncorrelated (independent) with each generation of recombination. Therefore, the number of independent tests increases, which increases the threshold required for statistical significance. If one marker is closer to a QTL than another, this will become increasingly clear as recombinations accumulate over numerous generations. As a result, outbred mice offer better resolution for mapping QTLs, but have reduced power to detect true associations relative to an $F_2$ population of the same size. Of course, both precision and power increase with sample size.

Wild-derived and wild-caught (**WC**) mice provide an extreme example of the relationship between precision and power. Wild mice are obtained by capturing individuals from natural populations. WC animals may be phenotyped directly or the progeny of the WC ancestors

may be studied. Studying the descendants of WC mice allows for better control over environmental variables; however, WC and even wild-derived inbred strains are more difficult to capture and handle than laboratory strains of mice (Wahlsten et al. 2003). This is a concern for phenotypes that require experimenter handling (Wahlsten et al. 2003). Rates of linkage disequilibrium (**LD**) in certain populations of WC mice are very low compared to laboratory populations (Laurie et al. 2007); thus, WC mice may offer considerable improvements in QTL mapping precision. Another potential advantage of using WC mice to map QTLs is that they possess a greater diversity of alleles (Laurie et al. 2007; Harper 2008). WC mice have also been subjected to different selection pressures as compared to laboratory mice (Harper 2008). WC mice may therefore be useful for exploring a wider spectrum of variation than mice derived from inbred or even outbred laboratory stocks (Laurie et al. 2007). On the other hand, high genetic diversity may result in reduced power to identify significant QTLs. For example, a genetically heterogeneous population is likely to segregate more variants that influence a given trait than a genetically homogenous population. Individual QTLs may therefore have smaller effect sizes, making them more difficult to detect (Ishikawa et al. 2000; Parker and Palmer 2011). Another limitation of WC mice is that numerous rare alleles will be present and consequently, a significant fraction of the variance may be due to alleles with a low minor allele frequency (**MAF**). Even with extremely large sample sizes it is difficult to detect alleles with low MAF, meaning that a significant amount of variation will be comprised of “dark matter”. These problems are similar to the well-known limitations of human genome-wide association studies (**GWAS**; Manolio et al. 2009).

A less extreme alternative to WC mice are commercial outbred (**CO**) populations. Available CO populations include MF-1 (Ghazalpour et al. 2008), CD-1 (Aldinger et al. 2009), NMR1 (Zhang et al. 2012) and CFW (Yalcin et al. 2010); a variety of CO populations were recently surveyed (Yalcin et al. 2010). Because they are derived from the same domesticated laboratory mice that were used to produce common inbred laboratory strains (Yalcin et al. 2010), they are behaviorally and physiologically well-adjusted to laboratory environments. Many CO stocks have been maintained for tens to hundreds of generations with relatively large population sizes. Although CO mice are expected to carry many alleles with low MAF, relatively recent population bottlenecks are likely to have reduced genetic diversity in CO mice relative to WC mice. Therefore, CO mice have the potential to offer excellent mapping precision but, like WC populations, CO are likely to require very large sample sizes because of low LD and low MAF.

Heterogeneous stock (**HS**) populations are produced by mixing more than two inbred strains (most commonly eight) and maintaining them as an outbred population for multiple generations. HS are preferable to WC and CO populations because the MAF of alleles in a HS is predicted to be no lower than 1/*n*, where *n* is the number of inbred strains used to create the population. In practice, allele frequencies are often higher because multiple inbred strains share the same alleles. However, some alleles will have frequencies below 1/*n* because of genetic drift or inadvertent selection against rare alleles that decreased fitness or fecundity. Examples of commonly used HS mouse populations include both the Boulder and Northport HS, which were derived from eight inbred mouse strains (Chia et al. 2005). More

recently, the HS-CC (Iancu et al. 2010) and the Diversity Outcross (**DO;** Logan et al. 2013) were derived from the same eight inbred strains that were used to create the CC. While most HS populations have been derived from inbred laboratory strains and thus trace their lineage back to *Mus musculus domesticus* (Yang et al. 2011), the HS-CC and the DO include several strains that are the inbred descendants of WC mice sampled from around the world (Chesler et al. 2008; Philip et al. 2011; Thaisz et al. 2012).

The size of the breeding population used to maintain an HS is an important parameter as it determines the rate of genetic drift. The Northport HS has been maintained by breeding 24 pairs per generation (Demarest et al. 2001), the Boulder HS uses 40 pairs (Mott et al. 2000), and the HS-CC uses 48 pairs (Iancu et al. 2010). The more recently created DO is maintained using 175 breeder pairs (Svenson et al. 2012; Logan et al. 2013). A final advantage of HS relative to WC and CO mice is the potential to impute founder haplotypes, which should allow imputation of millions of single nucleotide polymorphisms (**SNPs**) that segregate among founder strains in cases where those founders have been fully sequenced (Szatkiewicz et al. 2008; Wang et al. 2012; Baud et al. 2013). HS populations exist in other organisms including yeast (Cubillos et al. 2013), flies (Huang et al. 2012b), and rats (Baud et al. 2013; Parker et al. 2013a). The sample size needed to map QTLs in an HS population is likely to be smaller than that required for a CO or WC population because LD may be greater and because alleles should have a higher MAF.

AILs can be thought of as a special example of a HS; the distinguishing feature is that only two inbred strains are used to create an AIL. By using only two inbred strains, the genetic diversity of an AIL is held to a minimum, which is both a blessing and a curse. The advantages of limiting diversity are three fold: First, the MAF starts at 0.5, which maximizes power to detect the genetic effects of each allele. MAF may deviate from 0.5 due to inadvertent selection for fitness and fecundity and/or due to genetic drift, which is exacerbated by a small population size. In addition, rare alleles can arise due to *de novo* mutations or due to mutations that were segregating among the inbred progenitors; both situations are expected to be uncommon. Another advantage of AILs is that imputation is extremely simple; this is because identifying two alleles that are identical by state (**IBS**) necessarily means that they are also identical by descent (**IBD**). This reduces the number of markers that must be genotyped. Finally, AILs provide the best possible environment in which to analyze gene-by-gene interactions (epistasis), again, because allele frequencies are balanced (Parker and Palmer 2011; Pettersson et al. 2011). The disadvantage of limiting genetic diversity to two inbred strains is that at some genes/loci there will be no functionally significant differences between the two founding strains. Therefore, some genes that have the potential to alter a trait of interest will go undetected in any particular AIL. In contrast, HS, CO and WC populations have greater genetic diversity and are thus more likely to harbor functional variants at a given locus. Even so, the requirement for relatively small sample sizes as well as the relative simplicity of the analysis of an AIL are compelling advantages for forward genetic studies.

## Genotyping strategies

AILs and other outbred populations require greater genotyping density as compared to simpler $F_2$ crosses that have traditionally been used for QTL mapping. Whereas genotyping was once the most difficult and costly component of a QTL study, recent technological advances have revolutionized genotyping, permitting investigation of increasingly recombinant populations. Various array platforms have been developed for the analysis of mouse populations including the Mouse Diversity Array (**MDA**) which interrogates approximately 600,000 SNPs (Yang et al. 2009). More recently, alternative platforms have been developed which are much less expensive including the MUGA (Svenson et al. 2012; Logan et al. 2013; Iancu et al. 2013) and MEGAMUGA arrays; custom genotyping arrays are also routinely developed for specific projects (e.g. Cheng et al. 2010; Philip et al. 2011). Despite this multitude of options, the cost of genotyping mice on arrays remains substantially higher than comparable human arrays. This is presumably due to the much larger demand for human genotyping products and a lack of competition. As a result, there is a stronger incentive for mouse geneticists to identify alternative genotyping techniques.

As the cost of next-generation sequencing has continued to decline, techniques have been developed to obtain genotype information using next-generation sequencing platforms. For example, reduced representation methods such as genotyping-by-sequencing (**GBS**) limit sequencing to the DNA surrounding restriction enzyme digestion sites (Elshire et al. 2011). This approach is conceptually similar to sequence capture methods that target specific regions (e.g. exons; Clark et al. 2011; Rohland and Reich 2012). When applied to a large number of samples, both approaches rely on the ability to ligate uniquely indexed oligonucleotides to the ends of fragmented DNA. Another alternative is to sequence the whole genome at very low coverage in conjunction with genotype and haplotype imputation strategies (Pasaniuc et al. 2012). Because each sample can be uniquely identified during analysis, DNA from several individuals can be pooled and sequenced in a single reaction. AILs are particularly well-suited to these approaches because lower coverage is required to obtain reliable genotype calls for common alleles.

Determining the number of markers that should be genotyped per individual is critical to any QTL mapping study. While it is always best to genotype the largest number of markers possible, balancing cost with coverage and resolution is an important consideration. Unlike more complicated populations such as WC, CO and HS, the selection of markers to be genotyped in an AIL is very straightforward because any marker that is polymorphic between the two inbred strains is equally informative. Markers should also be spaced evenly throughout the genome.

Regardless of the genotyping strategies used, a certain amount of missing data (e.g. for a particular individual at a particular marker) is inevitable. Furthermore, it is sometimes of interest to infer the genotype at a marker that was not directly evaluated in any of the individuals. A variety of strategies for genotype imputation have been developed. In general, these methods rely on empirical data about the structure of LD among nearby markers (Szatkiewicz et al. 2008). The problem of imputation is relatively simple in an AIL because there are only two possible haplotypes and the full sequence of those haplotypes (the inbred

strains) is often available. Furthermore, pedigree information may be available; if genotypes from close relatives are known, that information can also be used for imputation (e.g. Cheng et al 2010).

## QTL mapping in structured populations

Identifying QTLs in the presence of population structure can be challenging. Population structure refers to the presence of systematic differences in allele frequencies among subsets of a population and can be due to recent pedigree-relationships, non-random mating and similar phenomena. Because genetic similarity will cause phenotypic similarity, individuals cannot be treated as independent observations in structured populations. This non-independence violates an assumption of certain statistical procedures used for QTL mapping. Failure to properly control for population structure can lead to inflation of the type I error rate (Newman et al. 2001; Peirce et al. 2008; Cheng et al. 2010; Cheng and Palmer 2012; Cheng et al. 2013). As a result, a number of methods have been developed to address the challenge of identifying QTLs in structured populations (Astle and Balding 2009).

Most methods account for population structure using structured association analysis, linear mixed models, or some combination of the two. Structured association analysis assumes that a population is comprised of a number of subpopulations or clusters, each tracing back to a distinct ancestral group (Pritchard et al. 2000). Genetic markers are used to determine an individual's membership in a cluster and this information is used to account for population structure in association testing, usually in the form of a linear regression model. Alternatively, principle components analysis (**PCA**) can be used to identify clusters (Price et al. 2006). PCA yields similar results to structured association analysis, but runs faster, making it ideal for large samples (Price et al. 2006). Both methods produce population membership vectors that can be used as fixed effects in a linear model or to adjust the raw phenotype and genotype data prior to association testing. Although structured association analysis and PCA are excellent at identifying population structure due to ancestry, they are not designed to account for recent familial relationships (Yu et al. 2005; Zhao et al. 2007). Therefore, genomic control is typically applied in addition to one of these methods to prevent the inflation of test statistics caused by cryptic relatedness (Devlin and Roeder 1999).

The disadvantage to this approach is that accounting for ancestral population structure and close familial relationships alone may not capture the more complex relationships present between pairs of individuals in a multigenerational cross (Abney et al. 2000, Cheng et al. 2010, Iancu et al. 2012). In contrast, linear mixed models can simultaneously account for multiple levels of relatedness. Several lines of evidence have shown mixed models to be highly effective in controlling the false positive (type I error) rate for association testing in structured populations (Abney et al. 2002; Kang et al. 2010; Cheng et al. 2010; Kenny et al. 2010; Listgarten et al. 2012; Cheng and Palmer 2012; Sul and Eskin 2013). This has renewed an interest in developing computationally efficient mixed model approaches for GWAS. A selection of freely available mixed model software for QTL mapping is discussed in this review. First, we provide a general overview of mixed model association analysis.

## Modeling association with mixed linear models

Mixed models have been applied to the quantitative genetics of plant and animal breeding for decades (Thompson 2008). In particular, the mixed model developed by Henderson (1975), called the “animal model”, has been widely used to predict phenotypic values and estimate heritability in animal populations. Unlike other methods available for estimating heritability at the time, Henderson’s model could incorporate information from complex, multigenerational pedigrees to estimate multiple causal components of the phenotypic variance (Kruuk 2004). Until recently, computational restrictions have limited the use of mixed models outside of plant and animal breeding. Henderson’s animal model has since been extended to genome-wide marker data (Goddard et al. 2009), which forms the basis of the mixed model association methods described here.

By definition, mixed linear models include both fixed and random terms. Fixed effects are explanatory variables that affect the mean phenotype for all subjects. Random effects may be thought of as a random sample taken from an infinite population of levels; these are used to estimate the variance of the effects rather than a separate parameter for each level. In a simple genetic model, a fixed effect is the marker of interest and its levels are genotypes. Its parameters would describe the mean effect of each genotype group on the phenotype. In contrast, the random effects parameters would describe the magnitude of variability of genotypic effects.

A critical difference between fixed and random effects is that random effects have a covariance structure. Covariance describes the extent to which the difference of one individual from the mean is similar to the difference of a second individual. This is important for GWAS, which typically assume an additive model of inheritance, because phenotypic covariance between individuals is determined by how genetically similar they are. If phenotypic variation is caused by the sum of multiple genetic factors with individually minor effects, then the more alleles that individuals share, the more similar they will be. The key point here is that in outbred samples, pairs of individuals covary randomly with respect to genetic background. This effect is more pronounced for AILs because individuals are related to one another within *and* between families (Darvasi & Soller 1995, Peirce et al. 2008). Because fixed effects models do not account for covariance among relatives, they are prone to produce spurious associations. However, a mixed model can simultaneously account for the fixed effect of genotype at a locus and for latent effects such as polygenic background which influence the phenotype at random, making it a very powerful approach for mapping QTLs.

Mixed model association requires three steps. First, relatedness is modeled using a genetic relationship matrix. This can be constructed from pedigree data or genotyped markers. Next, a mixed model is used to estimate the influence of relatedness on the phenotypic variance. Finally, evidence for association is evaluated using statistics that account for the component of phenotypic variance explained by genetic relationships. We describe each of these steps below.

### Step One: Constructing a genetic relationship matrix

Phenotypic variation is caused by multiple components of variance, each of which represents a different proportion of genetic or environmental effects on the phenotype. Genetic variance can be partitioned into additive, dominance and epistatic components. Dominance and epistasis are rarely modeled in association studies because the computational requirements of variance component estimation are high and there is typically low power to detect these effects (Abney et al. 2000). Therefore, we focus our discussion on the additive genetic variance, which is commonly used to model the random effect of polygenic background in GWAS.

A matrix that summarizes all pairwise relationships among individuals in a sample is required to estimate the polygenic effect. This is called a genetic relationship matrix, which is derived from kinship coefficients calculated using pedigree or marker data. If a pedigree is used, kinship is calculated as the probability that any two alleles sampled at random from two individuals are IBD. Alternatively, genotyped markers can be used to calculate the probability that two alleles sampled from two individuals at a locus are IBS given genotype information at other loci. Multiple adaptations have been applied to these basic methods (Thompson 2013), but in general, the kinship coefficients they produce are similar (Kang et al. 2010).

In most cases, the use of empirical kinship coefficients derived from densely genotyped SNPs offers improved power over pedigree-based methods (Cheng et al. 2013). Full siblings share one half of their genome on average, but in mice, rats and humans (all of who have about 20 chromosomes that are 100–200 cM in length), the actual sharing commonly ranges between 40–60% (standard deviation is about 5%). Kinships computed from the pedigree give the expected rather than the realized degree of sharing, so using genotype data is advantageous. Finally, in studies where only incomplete pedigrees and sparse genotyping data are available, it may be useful to combine both types of information (Cheng et al. 2013).

Computationally efficient methods have been developed that estimate relatedness using only a subset of genotyped markers (e.g. Lippert et al. 2011; Sul and Eskin 2013). In one approach, markers that show strong evidence for association with the phenotype under a linear regression model are used to construct the relationship matrix (Lippert et al. 2013). Conditioning on markers that are likely to influence the phenotype can improve power (Lippert et al. 2013), but using tentatively associated markers to estimate variance components may compromise the ability to control for the effects of population structure (Yang et al. 2014). An alternative strategy is to use a subset of markers sampled uniformly across the genome (uniform sampling helps to avoid bias caused by LD) (Lippert et al. 2011). When feasible, including all available genotype data provides the most accurate correction for subtle population structure (Yang et al. 2014).

Importantly, markers within the region being tested should be excluded when estimating relatedness (Listgarten et al. 2012; Cheng et al. 2013; Yang et al. 2014). Nearby markers may be correlated with the QTL (due to linkage) *and* genetic background (due to population structure) (Atwell et al. 2010). If the markers used to describe genetic background are

correlated with the SNP being tested, the test statistic will be overly conservative. Thus, failure to exclude nearby markers can lead to a significant loss of power (Listgarten et al. 2012a; Cheng et al. 2013; Yang et al. 2014). This problem has been referred to as “proximal contamination” (Listgarten et al. 2012). A convenient way to avoid it is to exclude all markers on the chromosome currently being scanned (Lippert et al. 2011; Cheng et al. 2013; Yang et al. 2014). A more detailed analysis of the various methods for modeling relatedness using genotyped markers is outside the scope of this review and has been described elsewhere (Thompson 2013; Yang et al. 2014).

### Step Two: Estimating variance components

Variance components methods can be used to model how environmental and genetic factors differentially affect phenotypic variation by estimating parameters for each component of variance. In mixed model GWAS, the additive genetic relationship matrix is treated as a constant (Thompson 2008). Maximum likelihood estimation (**MLE**) is then used to derive the additive and residual genetic variance components. This is accomplished by identifying the probability distribution that is most likely to have produced the pattern of IBD or IBS sharing across the genome. The distribution is then searched to obtain the set of parameters that maximizes the likelihood of the data under the null and alternative hypotheses.

Restricted maximum likelihood estimation (**REML**) may also be used to estimate variance components. REML structures the likelihood such that its maximization does not require estimation of the fixed effects. In general, the two methods perform similarly (Searle et al. 2008).

### Step Three: Association testing

The typical mixed model association procedure fits a single marker as a fixed effect and the polygenic component as a random effect. As described above, MLE is used to derive parameters under the null and alternative hypotheses, which are evaluated by taking the ratio of their likelihoods. Although the likelihood ratio test is considered to be the gold standard for GWAS, estimating parameters for every marker in a large study is computationally expensive (Abney et al. 2000). Various procedures have been developed to address this limitation.

A common strategy is to assume that for most traits, the phenotypic variance can be explained by a large number of markers with individually weak effects (Kang et al. 2010; Zhang et al. 2010). If this assumption holds, then variance components are not expected to differ significantly from one marker to the next, making repetitive iterations unnecessary. Instead, the random effect estimated under the null hypothesis is used as a fixed effect in separate association tests for each marker (Kang et al. 2010; Zhang et al. 2010). Other techniques are available which do not require variance components to be uniform across all markers (Lippert et al. 2011; Zhou and Stephens 2012); instead, improvements in speed are achieved by modifying the algorithms used to estimate variance components. Variations of these approaches and other strategies for mixed model association mapping are available in the software discussed below.

In addition to performing an appropriate analysis, it is also necessary to set a threshold for significance. Permutation is widely used to determine significance thresholds in QTL mapping (Churchill and Doerge 1994). However, the standard method of randomly shuffling phenotypes and genotypes is only suitable when all observations are exchangeable, which is not generally the case when individuals are related (Abney et al. 2002; Churchill and Doerge 2008; Peirce et al. 2008; Cheng and Palmer 2012). We have explored several methods for obtaining significance thresholds in an AIL. Simulations indicate that when relatedness is properly accounted for in the model, permutation is an efficient and robust option in many situations (Cheng and Palmer 2012). Other methods include bootstrapping, gene dropping (MacCluer et al. 1986; Cheng et al. 2010; Cheng and Palmer 2012) and a version of gene dropping that is implemented in GRAIP (Peirce et al. 2008).

## Mixed model association software

A variety of mixed model association packages are freely available for the analysis of AILs and other outbred populations (see Table 2). We provide an overview of different methods that have been proposed for mixed model association mapping and explain how they apply to various types of data. First we describe some early programs including TASSEL, GRAMMAR and EMMA and discuss how they evolved to the meet the demands of ever larger and more densely genotyped samples. We also highlight alternatives to these methods, including QTLRel, GEMMA, and FaST-LMM. We conclude by mentioning software designed for special applications, such as multi-locus mapping.

### First-generation mixed model association software

TASSEL, GRAMMAR and EMMA were among the first freely available programs designed for mapping QTLs using a mixed model. TASSEL accounts for relatedness by modeling broad population structure as a fixed effect and polygenic background as a random effect (Yu et al. 2005). It summarizes broad population structure using structured association and can estimate polygenic background from either marker or pedigree data. One or both terms may be included in the model along with covariates. TASSEL's ability to model both ancestry and familial relatedness works well for traits that are caused by few QTLs with large effects (e.g., flowering time in maize). Its flexibility also makes it appropriate for a variety of populations and architectures (Yu et al. 2005). The main disadvantage of TASSEL in its early years was that it re-estimated variance components for each marker, making it inefficient for large data sets.

EMMA improves on the efficiency of TASSEL by adjusting the covariance matrix so that the computational requirement at each iteration of the likelihood function is reduced (Kang et al. 2008). EMMA's approach is powerful and has been used to analyze a variety of phenotypes in mice (Bennett et al. 2010; Kirby et al. 2010; Johnson et al. 2012; Hersch et al. 2012; Ghazalpour et al. 2012; Courtney and Massett 2012; Himes et al. 2013; Mott et al. 2014) and other organisms (Uchiyama et al. 2013; Rosas et al. 2013). However, because EMMA also re-estimates variance components for each individual marker (Kang et al. 2008), it is very slow when analyzing large, densely genotyped samples.

The first release of GRAMMAR was specifically designed for pedigreed populations. GRAMMAR alleviates the computational burden of variance component estimation by first optimizing a reduced model that includes all effects except the effect of the marker (in other words, parameters are estimated under the null hypothesis) (Aulchenko et al. 2007). GRAMMAR then uses residuals from the reduced model (as opposed to the raw phenotypic data) to test for association. The motivation for this is that the marker signal is still captured by the residual once all of the other effects are removed; therefore the mixed model equation only needs to be solved once for each phenotype, and association tests can be performed rapidly with linear regression (Aulchenko et al. 2007). Although this approach improves efficiency, it tends to underestimate significance when population structure or QTL effect sizes are high (Zhou and Stephens 2012). Accordingly, GRAMMAR-GC, which features a modified version of genomic control to account for the inflation in type I error, was released shortly after GRAMMAR was first described (Amin et al. 2007).

Additional mixed model association software was available in the early 2000s, including WOMBAT and Mendel. WOMBAT is a flexible mixed model program capable of analyzing large pedigreed populations (Meyer 2007). Special features include the ability to incorporate multiple covariates, analyze multivariate traits, and model multiple random effects. WOMBAT also allows the user to model the covariance matrix using only its leading principal components (similar to the PCA approach used in human GWAS) in addition to modeling the full structure with a relationship matrix (Meyer 2007). An updated version of WOMBAT is now available that uses genome-wide marker data to estimate relatedness for large GWAS (Meyer and Tier 2011). Mendel is a multi-purpose genetics program that can be used for mixed model QTL mapping with marker or pedigree-based estimates of relatedness (Lange et al. 2013). Mendel can also summarize relatedness using strain coefficients rather than kinship coefficients. Strain coefficients represent the probability that a pair of alleles are derived from the same ancestral strain and are analogous to other measures of IBD (Bauman et al. 2008). One advantage of using strain coefficients is that they are useful for haplotype inference in samples derived from inbred lines (Bauman et al. 2008). Other notable features include the ability to handle multivariate traits and accommodate complex crosses (Zhou et al. 2011).

These programs made an impact in the field of quantitative genetics by bringing attention to the problem of complex relationships in human pedigrees and model organism populations. Although population structure was widely recognized as a confound in human GWAS, family-based association tests, structured association and PCA could not fully account for the complex structures present in every case. TASSEL, GRAMMAR and EMMA provided a temporary solution for researchers performing linkage studies in model organisms and human pedigrees. However, as the cost of genotyping continued to decrease, the computationally expensive animal model which formed the basis of TASSEL, GRAMMAR and EMMA became of limited use to the increasing number of investigators using GWAS to study complex traits. In response, the quantitative genetics community began to devise ways to apply a powerful mixed model framework to larger, more densely genotyped samples.

### Faster mixed models for large data sets

The need to analyze larger sets of data resulted in upgrades for existing methods. A recent version of TASSEL improves computational speed by combining mixed model compression with the algorithm 'population parameters previously determined' (**P3D**; Zhang et al. 2010). The first method "compresses" the mixed model by clustering individuals into groups according to kinship. Relatedness between pairs of *groups* rather than pairs of *individuals* is used as the random effect in the mixed model. Reducing the size of the random effect in this way can greatly improve efficiency when the number of groups is small. However, different combinations of compression levels and clustering algorithms have been shown to produce variable results; therefore, these parameters should be determined empirically prior to association testing (Zhang et al. 2010). Using the P3D approach, a mixed model is optimized without including the effect of the marker. This is solved to determine the population parameters (e.g. genetic and residual variances), which are fixed as Bayesian priors in the second step to estimate the non-population parameters (e.g. marker effect and polygenic variance) for each marker (Zhang et al. 2010). By holding the population parameters from step one constant, P3D avoids multiple iterations of the mixed model equation and improves speed. P3D may be applied individually or in combination with mixed model compression.

A faster version of EMMA, called EMMAX, has also been developed. The key difference between the two programs is that instead of re-estimating variance components for each alternative hypothesis, EMMAX uses a single estimate based on the null hypothesis (Kang et al. 2010). This approach is similar to the P3D algorithm used in TASSEL and running time is comparable for the two programs (Lippert et al. 2011). Although EMMAX may result in lower power compared to EMMA in cases where population structure is high or when QTL effect sizes are strong (Wu et al. 2011; Zhou and Stephens 2012), the magnitude of this difference is small and the gain in speed can be appreciable (Kang et al. 2010).

A number of modifications have been made since GRAMMAR was first introduced. At its inception, GRAMMAR was not equipped to calculate relatedness from marker genotypes, nor was it practical for large pedigrees (Aulchenko et al. 2007); however, current versions of GRAMMAR support marker-based kinship estimates, large sample sizes and offer improved power to identify rare variants (Amin et al. 2007; Svishcheva et al. 2012; Belonogova et al. 2013). Additional features include the ability to model gene-by-environment interactions, dominance, epistasis and parent-of-origin effects (Amin et al. 2007).

The recognition that mixed models could effectively account for population structure also inspired the development of new software. QTLRel was designed specifically for analysis in multigenerational crosses among model organisms and has been used to map QTLs for a variety of behavioral (Cheng et al. 2010; Samocha et al. 2010; Yoshizawa et al. 2012; Parker et al. 2012; Weber et al. 2013; Logan et al. 2013) and physiological (Lionikas et al. 2010; Parker et al. 2011; Heydemann et al. 2012; Bartnikas et al. 2012; Svenson et al. 2012; Leamy et al. 2012; Leamy et al. 2013a) traits in mouse AIL and HS populations. QTLRel estimates kinship using a rapid approach that can accommodate deep, complex pedigrees (Cheng et al. 2011). QTLRel's ability to incorporate multiple random terms allows it to estimate relatedness using a combination of pedigree and/or marker data, which is especially

useful when marker or pedigree data are missing or incomplete (Cheng et al. 2013). Including additional variance terms may also improve power by accounting for more of the "missing" heritability for a given trait (Cheng et al. 2010; Cheng et al. 2011). DOQTL is an R package that calls QTLRel; DOQTL provides additional tools for the analysis of HS populations such as the DO.

GEMMA avoids the repetitive iterations used in EMMA by devising a computationally inexpensive matrix multiplication algorithm to calculate genetic variance components for each marker (Zhou and Stephens 2012). Unlike most other software, GEMMA uses an exact test for association. An exact significance test requires that the sample distribution agree with the assumptions used to generate the distribution of the test statistic, whereas an approximate test may or may not be valid if the sample is sufficiently large. A comparison of GEMMA to EMMAX and GRAMMAR revealed that the two approximate methods had reduced power in highly structured populations, particularly when the strongest associations had relatively large effect sizes (Zhou and Stephens 2012). This effect was more pronounced for GRAMMAR, which also underperformed in a less structured population in which the strongest associations explained only a small proportion of the phenotypic variation (Zhou and Stephens 2012). EMMAX and GEMMA were comparable in the latter case, suggesting that the benefit of using an exact method will vary according to the method as well as the genetic architecture of the trait and the population under study.

FaST-LMM is another exact method that reduces running time by factoring the genetic relationship matrix into a simpler form (Lippert et al. 2011). This makes it possible to transform the phenotypes, markers and covariates so that the data become independent; linear regression may then be used for association testing (Lippert et al. 2011). FaST-LMM is very efficient provided that the number of markers used to estimate the relationship matrix is less than the sample size. When the number of markers is greater than the size of the cohort, a subset of them may be used to estimate relatedness. In this case, markers should be sampled uniformly across the genome (excluding the region containing the marker of interest) to avoid bias caused by LD (Lippert et al. 2011).

### Mixed model software for other QTL mapping applications

Most of software discussed in this review is designed for testing one marker or haplotype at a time. Although the majority of methods can incorporate covariates into the analysis, modeling the effect of many markers is problematic because the maximum number of markers fitted at once must be smaller than the number of individuals in the sample. Software designed for multi-locus analysis may therefore be preferable for certain genetic architectures (e.g. when multiple variants have opposing effects on the trait or when many weak effects are masked by a few large effects) (Yang et al. 2010). Examples include LMM Lasso (Rakitsch et al. 2013), MLMM (Segura et al. 2012), Qxpak5 (Pérez-Enciso and Misztal 2011), GCTA (Yang et al. 2012) and a forthcoming version of Bagpipe (Valdar et al 2009).

Another feature shared by most software is the ability to condition on genotypes and treat phenotypes as random. This is a reasonable approach for randomly sampled individuals. An alternative strategy, featured in the programs ROADTRIPS and MASTOR, is to condition

on phenotypes and treat *genotypes* as random. This method is robust to misspecification of the phenotypic covariance because it makes fewer assumptions about the phenotypic distribution; therefore, it may be more appropriate for samples of individuals ascertained by phenotype (Thornton and McPeek 2010; Jakobsdottir and McPeek 2013). A second advantage of ROADTRIPS and MASTOR is that they can easily incorporate information from individuals with missing genotype, phenotype or covariate information into the analysis by leveraging information from relatives (Thornton and McPeek 2010; Jakobsdottir and McPeek 2013). MASTOR is designed for mapping QTLs and ROADTRIPS is designed for case-control studies. Both programs are suitable for analyzing outbred pedigrees, complex inbred pedigrees and large GWAS containing any number of related and unrelated samples.

We have highlighted software suitable for the analysis of outbred populations such as AILs across a wide range of scenarios. However, many other mixed model programs are available that account for relatedness in other types of populations. For example, the R packages wgaim (Taylor and Verbyla 2011) and dlmap (Huang et al. 2012a) apply the framework of ASReml-R (a licensed program) (Gilmour et al. 2009) to interval mapping in inbred crosses. ASReml-R was also used to develop the MTMM package (Korte et al. 2012) which is designed for analyzing correlated phenotypes. Other licensed software capable of performing mixed model association analysis includes SAS, JMP Genomics and others. Important considerations when choosing software include the number of genotypes used, the size of the sample and the genetic architecture of the trait. For example, very large and/or densely genotyped samples may require extremely fast programs like GEMMA or FaST-LMM, and software such as QTLRel, EMMAX, TASSEL or GEMMA are appropriate for most model organism GWAS.

## Summary

In summary, AILs and related populations are useful because they allow for finer mapping of QTLs than studies using $F_2$ populations or other traditional crosses. Genotyping approaches have been developed to meet the need for dense genotyping, and a number of analytic strategies are available for studying these populations. The primary advantages of AILs relative to other outbred populations are balanced allele frequency and simplicity of analysis. AILs have been used in a variety of species to examine a wide array of phenotypes, as shown in Table 1.

## References

Abney M, McPeek MS, Ober C. Estimation of variance components of quantitative traits in inbred populations. Am J Hum Genet. 2000; 66:629–650. [PubMed: 10677322]

Abney M, Ober C, McPeek MS. Quantitative-trait homozygosity and association mapping and empirical genomewide significance in large, complex pedigrees: fasting serum-insulin level in the Hutterites. Am J Hum Genet. 2002; 70:920–934. [PubMed: 11880950]

Aldinger KA, Sokoloff G, Rosenberg DM, et al. Genetic Variation and Population Substructure in Outbred CD-1 Mice: Implications for Genome-Wide Association Studies. PLoS ONE. 2009; 4:e4729.10.1371/journal.pone.0004729 [PubMed: 19266100]

Amin N, van Duijn CM, Aulchenko YS. A Genomic Background Based Method for Association Analysis in Related Individuals. PLoS ONE. 2007; 2:e1274.10.1371/journal.pone.0001274 [PubMed: 18060068]

Astle W, Balding DJ. Population Structure and Cryptic Relatedness in Genetic Association Studies. Stat Sci. 2009; 24:451–471.10.1214/09-STS307

Atwell S, Huang YS, Vilhjálmsson BJ, et al. Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature. 2010; 465:627–631.10.1038/nature08800 [PubMed: 20336072]

Aulchenko YS, de Koning D-J, Haley C. Genomewide Rapid Association Using Mixed Model and Regression: A Fast and Simple Method For Genomewide Pedigree-Based Quantitative Trait Loci Association Analysis. Genetics. 2007; 177:577–585.10.1534/genetics.107.075614 [PubMed: 17660554]

Backdahl L, Guo JP, Jagodic M, et al. Definition of arthritis candidate risk genes by combining rat linkage-mapping results with human case-control association data. Ann Rheum Dis. 2008; 68:1925–1932.10.1136/ard.2008.090803 [PubMed: 19066175]

Bartnikas TB, Parker CC, Cheng R, et al. QTLs for murine red blood cell parameters in LG/J and SM/J F2 and advanced intercross lines. Mamm Genome. 2012; 23:356–366.10.1007/s00335-012-9393-3 [PubMed: 22322356]

Baud A, Hermsen R, Guryev V, et al. Combined sequence-based and genetic mapping analysis of complex traits in outbred rats. Nat Genet. 2013; 45:767–775.10.1038/ng.2644 [PubMed: 23708188]

Bauman LE, Sinsheimer JS, Sobel EM, Lange K. Mixed Effects Models for Quantitative Trait Loci Mapping With Inbred Strains. Genetics. 2008; 180:1743–1761.10.1534/genetics.108.091058 [PubMed: 18791243]

Becanovic K, Jagodic M, Sheng JR, et al. Advanced intercross line mapping of Eae5 reveals Ncf-1 and CLDN4 as candidate genes for experimental autoimmune encephalomyelitis. J Immunol. 2006; 176:6055–6064. [PubMed: 16670314]

Behnke JM, Iraqi FA, Mugambi JM, et al. High resolution mapping of chromosomal regions controlling resistance to gastrointestinal nematode infections in an advanced intercross line of mice. Mamm Genome. 2006; 17:584–597.10.1007/s00335-005-0174-0 [PubMed: 16783640]

Belonogova NM, Svishcheva GR, van Duijn CM, et al. Region-Based Association Analysis of Human Quantitative Traits in Related Individuals. PLoS ONE. 2013; 8:e65395.10.1371/journal.pone.0065395 [PubMed: 23799013]

Bennett BJ, Farber CR, Orozco L, et al. A high-resolution association mapping panel for the dissection of complex traits in mice. Genome Res. 2010; 20:281–290.10.1101/gr.099234.109 [PubMed: 20054062]

Bennett KE, Flick D, Fleming KH, et al. Quantitative Trait Loci That Control Dengue-2 Virus Dissemination in the Mosquito Aedes aegypti. Genetics. 2005; 170:185–194.10.1534/genetics.104.035634 [PubMed: 15781707]

Benson AK, Kelly SA, Legge R, et al. Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. Proc Natl Acad Sci U S A. 2010; 107:18933–18938.10.1073/pnas.1007028107 [PubMed: 20937875]

Besnier F, Wahlberg P, Rönneg\aard L, et al. Fine mapping and replication of QTL in outbred chicken advanced intercross lines. Genet Sel Evol. 2011; 43

Bradbury PJ, Zhang Z, Kroon DE, et al. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics. 2007; 23:2633–2635.10.1093/bioinformatics/btm308 [PubMed: 17586829]

Bryant CD, Kole LA, Guido MA, et al. Congenic dissection of a major QTL for methamphetamine sensitivity implicates epistasis. Genes Brain Behav. 2012; 11:623–632.10.1111/j.1601-183X.2012.00795.x [PubMed: 22487465]

Buchner DA, Geisinger JM, Glazebrook PA, et al. The juxtaparanodal proteins CNTNAP2 and TAG1 regulate diet-induced obesity. Mamm Genome. 2012; 23:431–442.10.1007/s00335-012-9400-8 [PubMed: 22752552]

Chen Y-P, Prashar A, Erichsen JT, et al. Heritability of ocular component dimensions in chickens: genetic variants controlling susceptibility to experimentally induced myopia and pretreatment eye size are distinct. Invest Ophthalmol Vis Sci. 2011; 52:4012–4020.10.1167/iovs.10-7045 [PubMed: 21436281]

Cheng R, Abney M, Palmer AA, Skol AD. QTLRel: an R package for genome-wide association studies in which relatedness is a concern. BMC Genet. 2011; 12:66. [PubMed: 21794153]

Cheng R, Lim JE, Samocha KE, et al. Genome-Wide Association Studies and the Problem of Relatedness Among Advanced Intercross Lines and Other Highly Recombinant Populations. Genetics. 2010; 185:1033–1044.10.1534/genetics.110.116863 [PubMed: 20439773]

Cheng R, Palmer AA. A Simulation Study of Permutation, Bootstrap, and Gene Dropping for Assessing Statistical Significance in the Case of Unequal Relatedness. Genetics. 2012; 193:1015–1018.10.1534/genetics.112.146332 [PubMed: 23267053]

Cheng R, Parker CC, Abney M, Palmer AA. Practical Considerations Regarding the Use of Genotype and Pedigree Data to Model Relatedness in the Context of Genome-Wide Association Studies. G358 Genes Genomes Genetics. 2013; 3:1861–1867.10.1534/g3.113.007948

Chesler EJ, Miller DR, Branstetter LR, et al. The Collaborative Cross at Oak Ridge National Laboratory: developing a powerful resource for systems genetics. Mamm Genome. 2008; 19:382–389.10.1007/s00335-008-9135-8 [PubMed: 18716833]

Cheverud JM, Lawson HA, Fawcett GL, et al. Diet-Dependent Genetic and Genomic Imprinting Effects on Obesity in Mice. Obesity. 2010; 19:160–170.10.1038/oby.2010.141 [PubMed: 20539295]

Chia R, Achilli F, Festing MFW, Fisher EMC. The origins and uses of mouse outbred stocks. Nat Genet. 2005; 37:1181–1186.10.1038/ng1665 [PubMed: 16254564]

Churchill GA, Doerge RW. Naive Application of Permutation Testing Leads to Inflated Type I Error Rates. Genetics. 2008; 178:609–610.10.1534/genetics.107.074609 [PubMed: 18202402]

Churchill GA, Doerge RW. Empirical threshold values for quantitative trait mapping. Genetics. 1994; 138:963–971. [PubMed: 7851788]

Clark MJ, Chen R, Lam HYK, et al. Performance comparison of exome DNA sequencing technologies. Nat Biotechnol. 2011; 29:908–914.10.1038/nbt.1975 [PubMed: 21947028]

Courtney SM, Massett MP. Identification of exercise capacity QTL using association mapping in inbred mice. Physiol Genomics. 2012; 44:948–955.10.1152/physiolgenomics.00051.2012 [PubMed: 22911454]

Cubillos FA, Parts L, Salinas F, et al. High Resolution Mapping of Complex Traits with a Four-Parent Advanced Intercross Yeast Population. Genetics. 201310.1534/genetics.113.155515

Darvasi A, Soller M. Advanced intercross lines, an experimental population for fine genetic mapping. Genetics. 1995; 141:1199. [PubMed: 8582624]

Demarest K, Koyner J, McCaughran J Jr, et al. Further characterization and high-resolution mapping of quantitative trait loci for ethanol-induced locomotor activity. Behav Genet. 2001; 31:79–91. [PubMed: 11529277]

Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999; 55:997–1004. [PubMed: 11315092]

Ehrich TH, Hrbek T, Kenney-Hunt JP, et al. Fine-mapping gene-by-diet interactions on chromosome 13 in a LG/J$\times$ SM/J murine model of obesity. Diabetes. 2005; 54:1863–1872. [PubMed: 15919810]

Elshire RJ, Glaubitz JC, Sun Q, et al. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLoS ONE. 2011; 6:e19379.10.1371/journal.pone.0019379 [PubMed: 21573248]

Fawcett GL, Jarvis JP, Roseman CC, et al. Fine-mapping of Obesity-related Quantitative Trait Loci in an F9/10 Advanced Intercross Line. Obesity. 2009; 18:1383–1392.10.1038/oby.2009.411 [PubMed: 19910941]

Fawcett GL, Roseman CC, Jarvis JP, et al. Genetic Architecture of Adiposity and Organ Weight Using Combined Generation QTL Analysis. Obesity. 2008; 16:1861–1868.10.1038/oby.2008.300 [PubMed: 18551125]

Fernandez J. Efficiency of the Use of Pedigree and Molecular Marker Information in Conservation Programs. Genetics. 2005; 170:1313–1321.10.1534/genetics.104.037325 [PubMed: 15879510]

Flint J, Eskin E. Genome-wide association studies in mice. Nat Rev Genet. 2012; 13:807–817.10.1038/nrg3335 [PubMed: 23044826]

Flint J, Mackay TFC. Genetic architecture of quantitative traits in mice, flies, and humans. Genome Res. 2009; 19:723–733.10.1101/gr.086660.108 [PubMed: 19411597]

Frésard L, Leroux S, Dehais P, et al. Fine mapping of complex traits in non-model species: using next generation sequencing and advanced intercross lines in Japanese quail. BMC Genomics. 2012; 13:551. [PubMed: 23066875]

Ghazalpour A, Doss S, Kang H, et al. High-Resolution Mapping of Gene Expression Using Association in an Outbred Mouse Stock. PLoS Genet. 2008; 4:e1000149.10.1371/journal.pgen.1000149 [PubMed: 18688273]

Ghazalpour A, Rau CD, Farber CR, et al. Hybrid mouse diversity panel: a panel of inbred mouse strains suitable for analysis of complex genetic traits. Mamm Genome. 2012; 23:680–692.10.1007/s00335-012-9411-5 [PubMed: 22892838]

Gillett A, Marta M, Jin T, et al. TNF production in macrophages is genetically determined and regulates inflammatory disease in rats. J Immunol Baltim Md 1950. 2010; 185:442–450.10.4049/jimmunol.0904101

Gilmour, AR.; Gogel, BJ.; Cullis, BR.; Thompson, R. ASReml user guide release 3.0. VSN Int. Ltd; Hemel Hempstead UK: 2009.

Goddard ME, Wray NR, Verbyla K, Visscher PM. Estimating Effects and Making Predictions from Genome-Wide Marker Data. Stat Sci. 2009; 24:517–529.10.1214/09-STS306

Gomez-Machorro C, Bennett KE, del Lourdes Munoz M, Black Wc. Quantitative trait loci affecting dengue midgut infection barriers in an advanced intercross line of Aedes aegypti. Insect Mol Biol. 2004; 13:637–648. [PubMed: 15606812]

Harper JM. Wild-derived mouse stocks: an underappreciated tool for aging research. AGE. 2008; 30:135–145.10.1007/s11357-008-9057-0 [PubMed: 19424863]

Hasenstein, J.; Lamont, SJ. Chicken Gallinacin Gene Cluster Associated with Salmonella Colonization in Two Advanced Intercross Lines. Iowa State University; 2007.

Heifetz EM, Fulton JE, O'Sullivan NP, et al. Mapping QTL affecting resistance to Marek's disease in an F6 advanced intercross population of commercial layer chickens. BMC Genomics. 2009; 10:20.10.1186/1471-2164-10-20 [PubMed: 19144166]

Henderson CR. Best linear unbiased estimation and prediction under a selection model. Biometrics. 1975; 31:423–447. [PubMed: 1174616]

Hernandez-Valladares M, Naessens J, Gibson JP, et al. Confirmation and dissection of QTL controlling resistance to malaria in mice. Mamm Genome Off J Int Mamm Genome Soc. 2004a; 15:390–398.10.1007/s00335-004-3042-4

Hernandez-Valladares M, Rihet P, ole-MoiYoi OK, Iraqi FA. Mapping of a new quantitative trait locus for resistance to malaria in mice by a comparative mapping approach with human Chromosome 5q31-q33. Immunogenetics. 2004b; 56:115–117.10.1007/s00251-004-0667-0 [PubMed: 15118851]

Hersch M, Peter B, Kang HM, et al. Mapping Genetic Variants Associated with Beta-Adrenergic Responses in Inbred Mice. PLoS ONE. 2012; 7:e41032.10.1371/journal.pone.0041032 [PubMed: 22859963]

Heydemann A, Swaggart KA, Kim GH, et al. The superhealing MRL background improves muscular dystrophy. Skelet Muscle. 2012; 2:26.10.1186/2044-5040-2-26 [PubMed: 23216833]

Himes BE, Sheppard K, Berndt A, et al. Integration of Mouse and Human Genome-Wide Association Data Identifies KCNIP4 as an Asthma Gene. PLoS ONE. 2013; 8:e56179.10.1371/journal.pone.0056179 [PubMed: 23457522]

Huang BE, Shah R, George AW. dlmap: An R Package for Mixed Model QTL and Association Analysis. J Stat Softw. 2012a; 50:1–22.

Huang W, Richards S, Carbone MA, et al. Epistasis dominates the genetic architecture of Drosophila quantitative traits. Proc Natl Acad Sci. 2012b; 109:15553–15559. [PubMed: 22949659]

Huberle A, Beyeen AD, Ockinger J, et al. Advanced Intercross Line Mapping Suggests That Ncf1 (Ean6) Regulates Severity in an Animal Model of Guillain-Barre Syndrome. J Immunol. 2009; 182:4432–4438.10.4049/jimmunol.0803847 [PubMed: 19299744]

Iancu O, Darakjian P, Walter N, et al. Genetic diversity and striatal gene networks: focus on the heterogeneous stock-collaborative cross (HS-CC) mouse. BMC Genomics. 2010; 11:585. [PubMed: 20959017]

Iancu OD, Oberbeck D, Darakjian P, et al. Selection for Drinking in the Dark Alters Brain Gene Coexpression Networks. Alcohol Clin Exp Res. 2013; 37:1295–1303.10.1111/acer.12100 [PubMed: 23550792]

Illingworth CJR, Parts L, Bergström A, et al. Inferring Genome-Wide Recombination Landscapes from Advanced Intercross Lines: Application to Yeast Crosses. PLoS ONE. 2013; 8:e62266.10.1371/journal.pone.0062266 [PubMed: 23658715]

Iraqi F, Clapcott SJ, Kumari P, et al. Fine mapping of trypanosomiasis resistance loci in murine advanced intercross lines. Mamm Genome. 2000; 11:645–648.10.1007/s003350010133 [PubMed: 10920233]

Ishikawa A, Matsuda Y, Namikawa T. Detection of quantitative trait loci for body weight at 10 weeks from Philippine wild mice. Mamm Genome. 2000; 11:824–830.10.1007/s003350010145 [PubMed: 11003694]

Jagodic M, Becanovic K, Sheng JR, et al. An advanced intercross line resolves Eae18 into two narrow quantitative trait loci syntenic to multiple sclerosis candidate loci. J Immunol. 2004; 173:1366–1373. [PubMed: 15240732]

Jakobsdottir J, McPeek MS. MASTOR: mixed-model association mapping of quantitative traits in samples with related individuals. Am J Hum Genet. 2013; 92:652–666.10.1016/j.ajhg.2013.03.014 [PubMed: 23643379]

Jarvis JP, Cheverud JM. Mapping the Epistatic Network Underlying Murine Reproductive Fatpad Variation. Genetics. 2010; 187:597–610.10.1534/genetics.110.123505 [PubMed: 21115969]

Jennen DG, Vereijken AL, Bovenhuis H, et al. Confirmation of quantitative trait loci affecting fatness in chickens. Genet Sel Evol. 2005; 37:215.10.1186/1297-9686-37-3-215 [PubMed: 16194525]

Johannesson M, Karlsson J, Wernhoff P, et al. Identification of epistasis through a partial advanced intercross reveals three arthritis loci within the Cia5 QTL in mice. Genes Immun. 2005; 6:175–185.10.1038/sj.gene.6364155 [PubMed: 15716976]

Johnson NV, Ahn SH, Deshmukh H, et al. Haplotype Association Mapping Identifies a Candidate Gene Region in Mice Infected With Staphylococcus aureus. G358 Genes Genomes Genetics. 2012; 2:693–700.10.1534/g3.112.002501

Ka S, Markljung E, Ring H, et al. Expression of carnitine palmitoyl-CoA transferase-1B is influenced by a cis-acting eQTL in two chicken lines selected for high and low body weight. Physiol Genomics. 2013; 45:367–376.10.1152/physiolgenomics.00078.2012 [PubMed: 23512741]

Kang HM, Sul JH, Service SK, et al. Variance component model to account for sample structure in genome-wide association studies. Nat Genet. 2010; 42:348–354.10.1038/ng.548 [PubMed: 20208533]

Kang HM, Zaitlen NA, Wade CM, et al. Efficient Control of Population Structure in Model Organism Association Mapping. Genetics. 2008; 178:1709–1723.10.1534/genetics.107.080101 [PubMed: 18385116]

Kärst S, Strucken EM, Schmitt AO, et al. Effect of the myostatin locus on muscle mass and intramuscular fat content in a cross between mouse lines selected for hypermuscularity. BMC Genomics. 2013; 14:16. [PubMed: 23324137]

Kelly SA, Nehrenberg DL, Hua K, et al. Parent-of-origin effects on voluntary exercise levels and body composition in mice. Physiol Genomics. 2009; 40:111–120.10.1152/physiolgenomics.00139.2009 [PubMed: 19903762]

Kelly SA, Nehrenberg DL, Hua K, et al. Functional Genomic Architecture of Predisposition to Voluntary Exercise in Mice: Expression QTL in the Brain. Genetics. 2012; 191:643–654.10.1534/genetics.112.140509 [PubMed: 22466041]

Kelly SA, Nehrenberg DL, Peirce JL, et al. Genetic architecture of voluntary exercise in an advanced intercross line of mice. Physiol Genomics. 2010; 42:190–200.10.1152/physiolgenomics.00028.2010 [PubMed: 20388837]

Kenny EE, Kim M, Gusev A, et al. Increased power of mixed models facilitates association mapping of 10 loci for metabolic traits in an isolated population. Hum Mol Genet. 2010; 20:827–839.10.1093/hmg/ddq510 [PubMed: 21118897]

Kirby A, Kang HM, Wade CM, et al. Fine Mapping in 94 Inbred Mouse Strains Using a High-Density Haplotype Resource. Genetics. 2010; 185:1081–1095.10.1534/genetics.110.115014 [PubMed: 20439770]

Korte A, Vilhjálmsson BJ, Segura V, et al. A mixed-model approach for genome-wide association studies of correlated traits in structured populations. Nat Genet. 2012; 44:1066–1071.10.1038/ng.2376 [PubMed: 22902788]

Kraja AT, Lawson HA, Arnett DK, et al. Obesity–insulin targeted genes in the 3p26-25 region in human studies and LG/J and SM/J mice. Metabolism. 2012; 61:1129–1141.10.1016/j.metabol.2012.01.008 [PubMed: 22386932]

Kruuk LEB. Estimating genetic parameters in natural populations using the “animal model. Philos Trans R Soc B Biol Sci. 2004; 359:873–890.10.1098/rstb.2003.1437

Lange K, Papp JC, Sinsheimer JS, et al. Mendel: the Swiss army knife of genetic analysis programs. Bioinformatics. 2013; 29:1568–1570.10.1093/bioinformatics/btt187 [PubMed: 23610370]

Laurie CC, Nickerson DA, Anderson AD, et al. Linkage disequilibrium in wild mice. PLoS Genet. 2007; 3:e144. [PubMed: 17722986]

Lawson HA, Cady JE, Partridge C, et al. Genetic Effects at Pleiotropic Loci Are Context-Dependent with Consequences for the Maintenance of Genetic Variation in Populations. PLoS Genet. 2011a; 7:e1002256.10.1371/journal.pgen.1002256 [PubMed: 21931559]

Lawson HA, Lee A, Fawcett GL, et al. The importance of context to the genetic architecture of diabetes-related traits is revealed in a genome-wide scan of a LG/J × SM/J murine model. Mamm Genome. 2011b; 22:197–208.10.1007/s00335-010-9313-3 [PubMed: 21210123]

Lawson HA, Zelle KM, Fawcett GL, et al. Genetic, epigenetic, and gene-by-diet interaction effects underlie variation in serum lipids in a LG/JxSM/J murine model. J Lipid Res. 2010; 51:2976–2984.10.1194/jlr.M006957 [PubMed: 20601649]

Leamy LJ, Kelly SA, Hua K, et al. Quantitative trait loci for bone mineral density and femoral morphology in an advanced intercross population of mice. Bone. 2013a; 55:222–229.10.1016/j.bone.2013.02.014 [PubMed: 23486184]

Leamy LJ, Kelly SA, Hua K, et al. Quantitative trait loci for bone mineral density and femoral morphology in an advanced intercross population of mice. Bone. 2013b; 55:222–229.10.1016/j.bone.2013.02.014 [PubMed: 23486184]

Leamy LJ, Kelly SA, Hua K, Pomp D. Exercise and diet affect quantitative trait loci for body weight and composition traits in an advanced intercross population of mice. Physiol Genomics. 2012; 44:1141–1153.10.1152/physiolgenomics.00115.2012 [PubMed: 23048196]

Legare ME, Bartlett FS, Frankel WN. A major effect QTL determined by multiple genes in epileptic EL mice. Genome Res. 2000; 10:42–48. [PubMed: 10645948]

Lionikas A, Cheng R, Lim JE, et al. Fine-mapping of muscle weight QTL in LG/J and SM/J intercrosses. Physiol Genomics. 2010; 42A:33–38.10.1152/physiolgenomics.00100.2010 [PubMed: 20627939]

Lippert C, Listgarten J, Liu Y, et al. FaST linear mixed models for genome-wide association studies. Nat Methods. 2011; 8:833–835.10.1038/nmeth.1681 [PubMed: 21892150]

Lippert C, Quon G, Kang EY, et al. The benefits of selecting phenotype-specific variants for applications of mixed models in genomics. Sci Rep. 2013; 3

Listgarten J, Lippert C, Kadie CM, et al. Improved linear mixed models for genome-wide association studies. Nat Meth. 2012a; 9:525–526.10.1038/nmeth.2037

Listgarten J, Lippert C, Kadie CM, et al. Improved linear mixed models for genome-wide association studies - nmeth.2037.pdf. Nat Methods. 2012b; 9:525–526. [PubMed: 22669648]

Logan RW, Robledo RF, Recla JM, et al. High-precision genetic mapping of behavioral traits in the diversity outbred mouse population: Genetic mapping of behavioral traits in the outbred mouse. Genes Brain Behav. 2013; 12:424–437.10.1111/gbb.12029 [PubMed: 23433259]

Loschiavo M, Nguyen QK, Duselis AR, Vrana PB. Mapping and identification of candidate loci responsible for Peromyscus hybrid overgrowth. Mamm Genome. 2007; 18:75–85.10.1007/s00335-006-0083-x [PubMed: 17242862]

MacCluer JW, VandeBerg JL, Read B, Ryder OA. Pedigree analysis by computer simulation. Zoo Biol. 1986; 5:147–160.

Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature. 2009; 461:747–753.10.1038/nature08494 [PubMed: 19812666]

Marta M, Stridh P, Becanovic K, et al. Multiple loci comprising immune-related genes regulate experimental neuroinflammation. Genes Immun. 2010; 11:21–36.10.1038/gene.2009.62 [PubMed: 19675581]

McGuire JL, Bergstrom HC, Parker CC, et al. Traits of fear resistance and susceptibility in an advanced intercross line. Eur J Neurosci. 2013; 38:3314–3324.10.1111/ejn.12337 [PubMed: 23968228]

McNeil CL, Bain CL, Macdonald SJ, Fay JC. Multiple Quantitative Trait Loci Influence the Shape of a Male-Specific Genital Structure in Drosophila melanogaster. G358 Genes Genomes Genetics. 2011; 1:343–351.10.1534/g3.111.000661

McPeek MS. From mouse to human: fine mapping of quantitative trait loci in a model organism. Proc Natl Acad Sci U S A. 2000; 97:12389–12390.10.1073/pnas.240463597 [PubMed: 11050190]

Meyer K. WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). J Zhejiang Univ Sci B. 2007; 8:815–821.10.1631/jzus.2007.B0815 [PubMed: 17973343]

Meyer K, Tier B. “SNP Snappy”: A Strategy for Fast Genome-Wide Association Studies Fitting a Full Mixed Model. Genetics. 2011; 190:275–277.10.1534/genetics.111.134841 [PubMed: 22021386]

Moradi Marjaneh M, Martin ICA, Kirk EP, et al. QTL mapping of complex binary traits in an advanced intercross line: QTL mapping of complex binary traits in an advanced intercross line. Anim Genet. 2012; 43:97–101.10.1111/j.1365-2052.2012.02383.x [PubMed: 22742507]

Mott R, Yuan W, Kaisaki P, et al. The Architecture of Parent-of-Origin Effects in Mice. Cell. 2014; 156:332–342.10.1016/j.cell.2013.11.043 [PubMed: 24439386]

Newman DL, Abney M, McPeek MS, et al. The importance of genealogy in determining genetic associations with complex traits. Am J Hum Genet. 2001; 69:1146. [PubMed: 11590549]

Norgard EA, Jarvis JP, Roseman CC, et al. Replication of long-bone length QTL in the F9-F10 LG, SM advanced intercross. Mamm Genome. 2009; 20:224–235.10.1007/s00335-009-9174-9 [PubMed: 19306044]

Norgard EA, Lawson HA, Pletscher LS, et al. Genetic factors and diet affect long-bone length in the F34 LG, SM advanced intercross. Mamm Genome. 2010; 22:178–196.10.1007/s00335-010-9311-5 [PubMed: 21170743]

Ockinger J, Serrano-Fernández P, Möller S, et al. Definition of a 1.06-Mb region linked to neuroinflammation in humans, rats and mice. Genetics. 2006; 173:1539–1545.10.1534/genetics.106.057406 [PubMed: 16624898]

Ockinger J, Stridh P, Beyeen AD, et al. Genetic variants of CC chemokine genes in experimental autoimmune encephalomyelitis, multiple sclerosis and rheumatoid arthritis. Genes Immun. 2010; 11:142–154.10.1038/gene.2009.82 [PubMed: 19865101]

Park Y-G, Zhao X, Lesueur F, et al. Sipa1 is a candidate for underlying the metastasis efficiency modifier locus Mtes1. Nat Genet. 2005; 37:1055–1062.10.1038/ng1635 [PubMed: 16142231]

Parker CC, Chen H, Flagel SB, et al. Rats are the smart choice: Rationale for a renewed focus on rats in behavioral genetics. Neuropharmacology. 2013a10.1016/j.neuropharm.2013.05.047

Parker CC, Cheng R, Sokoloff G, et al. Fine-mapping alleles for body weight in LG/J × SM/J F2 and F34 advanced intercross lines. Mamm Genome. 2011; 22:563–571.10.1007/s00335-011-9349-z [PubMed: 21761260]

Parker CC, Cheng R, Sokoloff G, Palmer AA. Genome-wide association for methamphetamine sensitivity in an advanced intercross mouse line. Genes Brain Behav. 2012; 11:52–61.10.1111/j.1601-183X.2011.00747.x [PubMed: 22032291]

Parker CC, Palmer AA. Dark Matter: Are Mice the Solution to Missing Heritability? Front Genet. 201110.3389/fgene.2011.00032

Parker CC, Sokoloff G, Leung E, et al. A large QTL for fear and anxiety mapped using an $F_2$ cross can be dissected into multiple smaller QTLs: Dissection of a large QTL. Genes Brain Behav. 2013b:n/a–n/a.10.1111/gbb.12064

Pasaniuc B, Rohland N, McLaren PJ, et al. Extremely low-coverage sequencing and imputation increases power for genome-wide association studies. Nat Genet. 2012; 44:631–635.10.1038/ng.2283 [PubMed: 22610117]

Pavlicev M, Wagner GP, Noonan JP, et al. Genomic Correlates of Relationship QTL Involved in Fore- versus Hind Limb Divergence in Mice. Genome Biol Evol. 2013; 5:1926–1936.10.1093/gbe/evt144 [PubMed: 24065733]

Peirce JL, Broman KW, Lu L, et al. Genome Reshuffling for Advanced Intercross Permutation (GRAIP): Simulation and Permutation for Advanced Intercross Population Analysis. PLoS ONE. 2008; 3:e1977.10.1371/journal.pone.0001977 [PubMed: 18431467]

Peirce JL, Lu L, Gu J, et al. A new set of BXD recombinant inbred lines from advanced intercross populations in mice. BMC Genet. 2004; 5:7.10.1186/1471-2156-5-7 [PubMed: 15117419]

Pérez-Enciso M, Misztal I. Qxpak. 5: old mixed model solutions for new genomics problems. BMC Bioinformatics. 2011; 12:202. [PubMed: 21612630]

Pettersson M, Besnier F, Siegel PB, Carlborg Ö. Replication and Explorations of High-Order Epistasis Using a Large Advanced Intercross Line Pedigree. PLoS Genet. 2011; 7:e1002180.10.1371/journal.pgen.1002180 [PubMed: 21814519]

Philip VM, Sokoloff G, Ackert-Bicknell CL, et al. Genetic analysis in the Collaborative Cross breeding population. Genome Res. 2011; 21:1223–1238.10.1101/gr.113886.110 [PubMed: 21734011]

Prashar A, Hocking PM, Erichsen JT, et al. Common determinants of body size and eye size in chickens from an advanced intercross line. Exp Eye Res. 2009; 89:42–48.10.1016/j.exer.2009.02.008 [PubMed: 19249299]

Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006; 38:904–909.10.1038/ng1847 [PubMed: 16862161]

Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155:945–959. [PubMed: 10835412]

Putnam AS, Ivy JA. Kinship-Based Management Strategies for Captive Breeding Programs When Pedigrees Are Unknown or Uncertain. J Hered. 201310.1093/jhered/est068

Rakitsch B, Lippert C, Stegle O, Borgwardt K. A Lasso multi-marker mixed model for association mapping with population structure correction. Bioinforma Oxf Engl. 2013; 29:206–214.10.1093/bioinformatics/bts669

Redmond SB, Chuammitri P, Andreasen CB, et al. Genetic control of chicken heterophil function in advanced intercross lines: associations with novel and with known Salmonella resistance loci and a likely mechanism for cell death in extracellular trap production. Immunogenetics. 2011; 63:449–458.10.1007/s00251-011-0523-y [PubMed: 21455609]

Rockman MV, Kruglyak L. Breeding designs for recombinant inbred advanced intercross lines. Genetics. 2008; 179:1069–1078.10.1534/genetics.107.083873 [PubMed: 18505881]

Rohland N, Reich D. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. 2012; 22:939–946.10.1101/gr.128124.111 [PubMed: 22267522]

Rosas U, Cibrian-Jaramillo A, Ristova D, et al. Integration of responses within and across Arabidopsis natural accessions uncovers loci controlling root systems architecture. Proc Natl Acad Sci U S A. 2013; 110:15133–15138.10.1073/pnas.1305883110 [PubMed: 23980140]

Saavedra-Rodriguez K, Strode C, Flores Suarez A, et al. Quantitative Trait Loci Mapping of Genome Regions Controlling Permethrin Resistance in the Mosquito Aedes aegypti. Genetics. 2008; 180:1137–1152.10.1534/genetics.108.087924 [PubMed: 18723882]

Samocha KE, Lim JE, Cheng R, et al. Fine mapping of QTL for prepulse inhibition in LG/J and SM/J mice using F2 and advanced intercross lines. Genes Brain Behav. 2010; 9:759–767.10.1111/j.1601-183X.2010.00613.x [PubMed: 20597988]

Searle, SR.; Casella, G.; McCulloch, CE. Var Compon. John Wiley & Sons, Inc; 2008. Maximum Likelihood (ML) and Restricted Maximum Likelihood (REML); p. 232-257.

Segura V, Vilhjálmsson BJ, Platt A, et al. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat Genet. 2012; 44:825–830.10.1038/ng.2314 [PubMed: 22706313]

Sheng JR, Jagodic M, Dahlman I, et al. Eae19, a New Locus on Rat Chromosome 15 Regulating Experimental Autoimmune Encephalomyelitis. Genetics. 2005; 170:283–289.10.1534/genetics.104.035261 [PubMed: 15716504]

Shirley RL, Walter NAR, Reilly MT, et al. Mpdz is a quantitative trait gene for drug withdrawal seizures. Nat Neurosci. 2004; 7:699–700.10.1038/nn1271 [PubMed: 15208631]

Stridh P, Thessen Hedreul M, Beyeen AD, et al. Fine-mapping resolves Eae23 into two QTLs and implicates ZEB1 as a candidate gene regulating experimental neuroinflammation in rat. PloS One. 2010; 5:e12716.10.1371/journal.pone.0012716 [PubMed: 20856809]

Stylianou IM, Christians JK, Keightley PD, et al. Genetic complexity of an obesity QTL (Fob3) revealedby detailed genetic mapping. Mamm Genome. 2004; 15:472–481.10.1007/s00335-004-3039-z [PubMed: 15181539]

Sul JH, Eskin E. Mixed models can correct for population structure for genomic regions under selection. Nat Rev Genet. 2013; 14:300–300.10.1038/nrg2813-c1 [PubMed: 23438871]

Svenson KL, Gatti DM, Valdar W, et al. High-Resolution Genetic Mapping Using the Mouse Diversity Outbred Population. Genetics. 2012; 190:437–447.10.1534/genetics.111.132597 [PubMed: 22345611]

Svishcheva GR, Axenovich TI, Belonogova NM, et al. Rapid variance components–based method for whole-genome association analysis. Nat Genet. 2012; 44:1166–1170.10.1038/ng.2410 [PubMed: 22983301]

Szatkiewicz JP, Beane GL, Ding Y, et al. An imputed genotype resource for the laboratory mouse. Mamm Genome. 2008; 19:199–208.10.1007/s00335-008-9098-9 [PubMed: 18301946]

Taylor J, Verbyla A. R package wgaim: QTL analysis in bi-parental populations using linear mixed models. J Stat Softw. 2011; 40:1–18.

Terenina E, Babigumira BM, Le Mignon G, et al. Association study of molecular polymorphisms in candidate genes related to stress responses with production and meat quality traits in pigs. Domest Anim Endocrinol. 2013; 44:81–97.10.1016/j.domaniend.2012.09.004 [PubMed: 23063408]

Thaisz J, Tsaih S-W, Feng M, et al. Genetic analysis of albuminuria in collaborative cross and multiple mouse intercross populations. AJP Ren Physiol. 2012; 303:F972–F981.10.1152/ajprenal.00690.2011

Thompson EA. Identity by Descent: Variation in Meiosis, Across Genomes, and in Populations. Genetics. 2013; 194:301–326.10.1534/genetics.112.148825 [PubMed: 23733848]

Thompson R. Estimation of quantitative genetic parameters. Proc Biol Sci. 2008; 275:679–686.10.1098/rspb.2007.1417 [PubMed: 18211869]

Thornton T, McPeek MS. ROADTRIPS: Case-Control Association Testing with Partially or Completely Unknown Population and Pedigree Structure. Am J Hum Genet. 2010; 86:172–184.10.1016/j.ajhg.2010.01.001 [PubMed: 20137780]

Uchiyama K, Iwata H, Moriguchi Y, et al. Demonstration of genome-wide association studies for identifying markers for wood property and male strobili traits in Cryptomeria japonica. PloS One. 2013; 8:e79866.10.1371/journal.pone.0079866 [PubMed: 24260312]

Wahlsten D, Metten P, Crabbe JC. A rating scale for wildness and ease of handling laboratory mice: results for 21 inbred strains tested in two laboratories. Genes Brain Behav. 2003; 2:71–79. [PubMed: 12884964]

Wang JR, de Villena FP-M, Lawson HA, et al. Imputation of Single-Nucleotide Polymorphisms in Inbred Mice Using Local Phylogeny. Genetics. 2012; 190:449–458.10.1534/genetics.111.132381 [PubMed: 22345612]

Wang M, Lemon WJ, Liu G, et al. Fine mapping and identification of candidate pulmonary adenoma susceptibility 1 genes using advanced intercross lines. Cancer Res. 2003a; 63:3317–3324. [PubMed: 12810665]

Wang X, Le Roy I, Nicodeme E, et al. Using advanced intercross lines for high-resolution mapping of HDL cholesterol quantitative trait loci. Genome Res. 2003b; 13:1654–1664. [PubMed: 12805272]

Weber JN, Peterson BK, Hoekstra HE. Discrete genetic modules are responsible for complex burrow evolution in Peromyscus mice. Nature. 2013; 493:402–405.10.1038/nature11816 [PubMed: 23325221]

Wiren A, Gunnarsson U, Andersson L, Jensen P. Domestication-related genetic effects on social behavior in chickens - Effects of genotype at a major growth quantitative trait locus. Poult Sci. 2009; 88:1162–1166.10.3382/ps.2008-00492 [PubMed: 19439625]

Wu C, DeWan A, Hoh J, Wang Z. A Comparison of Association Methods Correcting for Population Stratification in Case-Control Studies: Method Comparison in Population Structure. Ann Hum Genet. 2011; 75:418–427.10.1111/j.1469-1809.2010.00639.x [PubMed: 21281271]

Yalcin B, Nicod J, Bhomra A, et al. Commercially Available Outbred Mice for Genome-Wide Association Studies. PLoS Genet. 2010; 6:e1001085.10.1371/journal.pgen.1001085 [PubMed: 20838427]

Yang H, Ding Y, Hutchins LN, et al. A customized and versatile high-density genotyping array for the mouse. Nat Methods. 2009; 6:663–666.10.1038/nmeth.1359 [PubMed: 19668205]

Yang H, Wang JR, Didion JP, et al. Subspecific origin and haplotype diversity in the laboratory mouse. Nat Genet. 2011; 43:648–655.10.1038/ng.847 [PubMed: 21623374]

Yang J, Benyamin B, McEvoy BP, et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet. 2010; 42:565–569.10.1038/ng.608 [PubMed: 20562875]

Yang J, Ferreira T, Morris AP, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet. 2012; 44:369–375. S1–3.10.1038/ng.2213 [PubMed: 22426310]

Yang J, Zaitlen NA, Goddard ME, et al. Advantages and pitfalls in the application of mixed-model association methods. Nat Genet. 2014; 46:100–106.10.1038/ng.2876 [PubMed: 24473328]

Yazbek SN, Buchner DA, Geisinger JM, et al. Deep congenic analysis identifies many strong, context-dependent QTLs, one of which, Slc35b4, regulates obesity and glucose homeostasis. Genome Res. 2011; 21:1065–1073.10.1101/gr.120741.111 [PubMed: 21507882]

Yoshizawa M, Yamamoto Y, O'Quin KE, Jeffery WR. Evolution of an adaptive behavior and its sensory receptors promotes eye regression in blind cavefish. BMC Biol. 2012; 10:108.10.1186/1741-7007-10-108 [PubMed: 23270452]

Yu J, Pressoir G, Briggs WH, et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet. 2005; 38:203–208.10.1038/ng1702 [PubMed: 16380716]

Yu X, Bauer K, Wernhoff P, et al. Fine mapping of collagen-induced arthritis quantitative trait loci in an advanced intercross line. J Immunol. 2006; 177:7042–7049. [PubMed: 17082620]

Yu X, Bauer K, Wernhoff P, Ibrahim SM. Using an advanced intercross line to identify quantitative trait loci controlling immune response during collagen-induced arthritis. Genes Immun. 2007; 8:296–301. [PubMed: 17361202]

Yu X, Teng H, Marques A, et al. High resolution mapping of Cia3: a common arthritis quantitative trait loci in different species. J Immunol Baltim Md 1950. 2009; 182:3016–3023.10.4049/jimmunol.0803005

Zhang S, Lou Y, Amstein TM, et al. Fine mapping of a major locus on Chromosome 10 for exploratory and fear-like behavior in mice. Mamm Genome. 2005; 16:306–318.10.1007/s00335-004-2427-8 [PubMed: 16104379]

Zhang W, Korstanje R, Thaisz J, et al. Genome-Wide Association Mapping of Quantitative Traits in Outbred Mice. G358 Genes Genomes Genetics. 2012; 2:167–174.10.1534/g3.111.001792

Zhang Z, Ersoz E, Lai C-Q, et al. Mixed linear model approach adapted for genome-wide association studies. Nat Genet. 2010; 42:355–360.10.1038/ng.546 [PubMed: 20208535]

Zhao K, Aranzana MJ, Kim S, et al. An Arabidopsis example of association mapping in structured samples. PLoS Genet. 2007; 3:e4.10.1371/journal.pgen.0030004 [PubMed: 17238287]

Zhou JJ, Ghazalpour A, Sobel EM, et al. Quantitative Trait Loci Association Mapping by Imputation of Strain Origins in Multifounder Crosses. Genetics. 2011; 190:459–473.10.1534/genetics.111.135095 [PubMed: 22143921]

Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 2012; 44:821–824.10.1038/ng.2310 [PubMed: 22706312]

**Table 1**

QTL mapping studies using advanced intercross lines (AILs).

| Organism | Study | Phenotype | Cross | Generation | Sample size | Marker number | Marker type | Genome-wide? | Analysis software | Controlled for relatedness? |
|---|---|---|---|---|---|---|---|---|---|---|
| *Aedes aegypti* | Bennett *et al*. 2005 | dengue-2 virus susceptibility | *D2S3*[a] x *D2MEB*[b] | $F_5$ | 436 | NA | SNP | Y | QTL Cartographer | NA |
| *Aedes aegypti* | Gomez-Machorro *et al*. 2004 | dengue-2 virus susceptibility | *Ibo11*[c] x *DS3*[c] | $F_5$ | 147 | 44 | SNP | Y | FORTRAN (BINARYQTL), QTL Cartographer (MIMQTL) | NA |
| *Aedes aegypti* | Saavedra-Rodriguez *et al.* 2008 | permethrin resistance | IMU[d] x NO | $F_3$ | 439 | 34 | SNP | Y | QTL Cartographer | NA |
| *Coturnix japonica* | Frésard *et al.* 2012 | fear (duration of tonic immobility) | LTI[e] x STI[e] | $F_6$ | 679 | 332 | SNP | Y | QTLmap, R/nlme | Random effects of dam and sire |
| *Coturnix japonica* | Frésard *et al.* 2012 | fear (duration of tonic immobility) | LTI[e] x DD | $F_7$ | 603 | 305 | SNP | Y | QTLmap, R/nlme | Random effects of dam and sire |
| *Drosophila melanogaster* | McNeil *et al*. 2011 | male-specific genital structure | b3852 x *Sam* | $F_{17}$ | 344 | 87 | SNP | Y | R/qtl | NA |
| *Gallus gallus* | Besnier *et al*. 2011 | body weight | HWS[f] x LWS[f] | $P_0$-$F_8$ | 1529 | 304 | SNP | N | | Random polygenic effect |
| *Gallus gallus* | Hasenstein and Lamont 2007 | *Salmonella enteritidis* colonization | Bro[g] x Leg | $F_8$ | 10 | 13 | SNP | N | | Fixed effects of dam and sire |
| *Gallus gallus* | Hasenstein and Lamont 2007 | *Salmonella enteritidis* colonization | Bro[g] x Fay | $F_8$ | 13 | 13 | SNP | N | | Fixed effects of dam and sire |
| *Gallus gallus* | Heifetz *et al.* 2009 | Marek's disease resistance | Leg1[h] x Leg2[h] | $F_6$ | 1615 | 232 | SNP, VNTR | Y | FitModel JMP (SAS) | NA |
| *Gallus gallus* | Jennen *et al*. 2005 | fatness (body weight, abdominal fat weight, percent abdominal fat) | Bro1[i] x Bro2[i] | $F_{8\text{-}9}$ | 1030 | 22 | VNTR | N | | Included family means to account for polygenic differences between families |
| *Gallus gallus* | Redmond *et al.* 2011 | immune response to *Salmonella enteritidis* (phagocytosis, oxidative burst, extracellular trap production) | Bro[g] x Leg | $F_{13}$ | 152 | 12,456 | SNP | Y | PROC MIXED (SAS) | NA |
| *Gallus gallus* | Redmond *et al.* 2011 | immune response to *Salmonella enteritidis* (phagocytosis, oxidative burst, extracellular trap production) | Bro[g] x Fay | $F_{13}$ | 189 | 13,052 | SNP | Y | PROC MIXED (SAS) | NA |
| *Mus musculus* | Bartnikas *et al.* 2012 | red blood cell parameters | LG/J x SM/J | $F_{34}$ | 472 | 3,144 | SNP | Y | QTLrel | Random polygenic effect |
| *Mus musculus* | Behnke *et al*. 2006 | resistance to GI nematode infections | SWR/J x CBA/J | $F_{6\text{-}7}$ | 1,100 | NA | VNTR | N | MapMaker EXP/QTL, QTL Express | NA |
| *Mus musculus* | Benson et al. 2010 | abundance of gut microbiota | C57BL/6J x HR[j] | $F_4$ | 645 | 530 | SNP | Y | R/qtl | Random effect of cohort, family and litter; GRAIP |
| *Mus musculus* | Cheng *et al*. 2010 | saline- and methamphetamine-induced locomotor activity | LG/J x SM/J | $F_{34}$ | 695 | 3,144 | SNP | Y | QTLrel | Random polygenic effect |
| *Mus musculus* | Cheverud *et al.* 2010 | organ weight, fat pad weight, body weight | LG/J x SM/J | $F_{16}$ | 1002 | 1,402 | SNP | Y | PROC MIXED (SAS) | Random effect of family |

| Organism | Study | Phenotype | Cross | Generation | Sample size | Marker number | Marker type | Genome-wide? | Analysis software | Controlled for relatedness? |
|---|---|---|---|---|---|---|---|---|---|---|
| *Mus musculus* | Ehrich *et al*. 2005 | growth rate, response to glucose challenge, organ weight, fat pad weight, serum lipids, insulin levels | LG/J x SM/J | $F_{16}$ | 1011 | 60 | VNTR | N | | Adjusted significance threshold for correlation between genotypes |
| *Mus musculus* | Fawcett *et al*. 2008 | organ weight, fat pad weight, body weight | LG/J x SM/J | $F_{3}$ | 1595 | 370 | SNP | Y | | Adjusted significance threshold for correlation between genotypes |
| *Mus musculus* | Fawcett *et al*. 2009 | organ weight, fat pad weight, body weight | LG/J x SM/J | $F_{9\text{-}10}$ | 1455 | 1,470 | SNP | Y | | Confidence intervals calculated using pedigree |
| *Mus musculus* | Hernandez-Valladares et al. 2004a | malaria resistance | C57BL/6J x A/J | $F_{11}$ | 340 | 63 | VNTR | N | MapMaker EXP/QTL, QTL Express, QTL Cartographer | NA |
| *Mus musculus* | Hernandez-Valladares et al. 2004b | malaria resistance | C57BL/6J x A/J | $F_{11}$ | 242 | 19 | VNTR | N | MapMaker EXP/QTL, QTL Express | NA |
| *Mus musculus* | Iraqi *et al.* 2000 | trypanosomiasis resistance | C57BL/6J x A/J | $F_{6}$ | 1986 | 68 | VNTR | N | MapMaker EXP/QTL | NA |
| *Mus musculus* | Iraqi *et al.* 2000 | trypanosomiasis resistance | C57BL/6J x BALB/cJ | $F_{6}$ | 912 | 88 | VNTR | N | MapMaker EXP/QTL | NA |
| *Mus musculus* | Jarvis and Cheverud 2010 | reproductive fat pad weight | LG/J x SM/J | $F_{10}$ | 1,298 | 1,470 | SNP | Y | | Adjusted significance threshold for correlation between genotypes |
| *Mus musculus* | Johannesson *et al*. 2005 | rheumatoid arthritis | $R3^{k}$ x $Eae2^{k}$ | $F_{2\text{-}8}$ | 676 | NA | SNP, VNTR | N | R/qtl | NA |
| *Mus musculus* | Kärst *et al*. 2013 | muscle mass, intramuscular fat content, water holding capacity | BMMI866 x BMMI806 | $F_{3}$ | 308 | 138 | SNP | Y | R/qtl, GRAMMAR | Random polygenic effect |
| *Mus musculus* | Kelly *et al*. 2010 | voluntary exercise levels, body composition, body weight | C57B6/J x $HR^{j}$ | $F_{4}$ | 815 | 530 | SNP | Y | R/qtl | GRAIP |
| *Mus musculus* | Kelly *et al*. 2009 | voluntary exercise levels, body composition | C57B6/J x $HR^{j}$ | $F_{4}$ | 815 | 530 | SNP | Y | PROC MIXED (SAS) | Random effect of family |
| *Mus musculus* | Kelly *et al*. 2012 | voluntary exercise levels, body composition, body weight, gene expression | C57B6/J x $HR^{j}$ | $F_{4}$ | 815 | 530 | SNP | Y | R/qtl | Adjusted significance threshold for correlation between genotypes |
| *Mus musculus* | Kraja *et al*. 2012 | obesity, lipid levels, blood pressure | LG/J x SM/J | $F_{16}$ | 1002 | 1,402 | SNP | Y | PROC MIXED (SAS) | Random effect of family |
| *Mus musculus* | Lawson *et al*. 2010 | serum cholesterol, free-fatty acids, triglyceride, glucose, insulin levels | LG/J x SM/J | $F_{16}$ | 1002 | 1,402 | SNP | Y | PROC MIXED (SAS) | Random effect of family |
| *Mus musculus* | Lawson *et al*. 2011a | blood lipid levels, body weight, organ weight, fat pad weight | LG/J x SM/J | $F_{16}$ | 1002 | 1,402 | SNP | Y | PROC MIXED (SAS) | Random effect of family |
| *Mus musculus* | Lawson *et al*. 2011b | blood lipid levels, glucose tolerance, body weight, organ weight, fat pad weight | LG/J x SM/J | $F_{16}$ | 1002 | 1,402 | SNP | Y | PROC MIXED (SAS) | Random effect of family |
| *Mus musculus* | Leamy *et al*. 2012 | body weight & composition, distance run, exercise traits | C57BL/6J x $HR^{j}$ | $F_{10}$ | 473 | 2,058 | SNP | Y | QTLrel | Random polygenic effect |
| *Mus musculus* | Leamy *et al*. 2013 | bone density and morphology | C57BL/6J x $HR^{j}$ | $F_{10}$ | 466 | 2,058 | SNP | Y | QTLrel | Random polygenic effect |

| Organism | Study | Phenotype | Cross | Generation | Sample size | Marker number | Marker type | Genome-wide? | Analysis software | Controlled for relatedness? |
|---|---|---|---|---|---|---|---|---|---|---|
| *Mus musculus* | Lionikas *et al.* 2010 | hind limb muscle weight | LG/J x SM/J | $F_{34}$ | 695 | 3,144 | SNP | Y | QTLrel | Random polygenic effect |
| *Mus musculus* | Moradi Marjaneh *et al.* 2012 | atrial septal parameters (patent foramen ovale and flap valve length) | QSi5 × 129T2/SvEms | $F_{14}$ | 400 | 150 | SNP | N | QTL-MLE | NA |
| *Mus musculus* | Norgard *et al.* 2009 | long-bone length | LG/J x SM/J | $F_{9\text{-}10}$ | 1455 | 1,402 | SNP | Y | SYSTAT, R/qtl | Adjusted significance threshold for correlation between genotypes |
| *Mus musculus* | Norgard *et al.* 2010 | long-bone length | LG/J x SM/J | $F_{34}$ | 1424 | 2,842 | SNP | Y | SYSTAT, PROC MIXED (SAS) | Random effect of sibship |
| *Mus musculus* | Parker *et al.* 2011 | body weight over time | LG/J x SM/J | $F_{34}$ | 701 | 3,144 | SNP | Y | QTLrel | Random polygenic effect |
| *Mus musculus* | Parker *et al.* 2012 | saline and methamphetamine-induced locomotor activity | C57BL/6J x DBA/2J | $F_8$ | 552 | 1,060 | SNP | Y | QTLrel | Random polygenic effect |
| *Mus musculus* | Pavlicev *et al.* 2013 | covariance of fore- and hind limb bone length | LG/J x SM/J | $F_{34}$ | 1134 | 2,842 | SNP | Y | R/lme4 | Random effect of family |
| *Mus musculus* | Samocha *et al.* 2010 | prepulse inhibition, acoustic startle, habituation to tone | LG/J x SM/J | $F_{34}$ | 135 | 3,144 | SNP | Y | QTLrel | Random polygenic effect |
| *Mus musculus* | Wang *et al.* 2003a | pulmonary adenoma susceptibility | A/J x C57BL/6J | $F_{11}$ | 399 | 26 | VNTR | N | MapMaker EXP/QTL | NA |
| *Mus musculus* | Wang *et al.* 2003b | HDL cholesterol | C57BL/6J × NZB/BlNJ | $F_{11}$ | 345 | 97 | VNTR | Y | | NA |
| *Mus musculus* | Yu *et al.* 2006 | collagen-induced arthritis (severity, onset, antibody response) | DBA/1J x FVB/N | $F_{11\text{-}12}$ | 308 | 59 | SNP, VNTR | N | QTX Map Manager, R/qtl | NA |
| *Mus musculus* | Yu *et al.* 2007 | immune response during collagen induced arthritis (ROS production, T-cell subset proportions, anti-collagen II antibodies) | DBA/1J x FVB/N | $F_{11\text{-}12}$ | 308 | 107 | SNP, VNTR | N | QTX Map Manager | NA |
| *Mus musculus* | Yu et al. 2009 | collagen-induced arthritis(severity, onset, susceptibility) | DBA/1J x FVB/N | $F_{11\text{-}12}$ | 308 | 60 | SNP, VNTR | N | QTX Map Manager | NA |
| *Mus musculus* | Zhang *et al.* 2005 | excitatory and exploratory behavior (open field test and light-dark box) | A/J x C57BL/6J | $F_{12}$ | 1077 | 40 | VNTR | N | QTX Map Manager, QTL Express | NA |
| *Peromyscus maniculatus*, *Peromyscus polionotus* | Loschiavo *et al.* 2007 | pre- and post-natal growth | $BW^l$ x $PO^l$ | $F_3$, $F_{10}$, $F_{13}$ | 94, 96 | NA | VNTR | N | QTX Map Manager | NA |
| *Rattus norvegicus* | Backdahl *et al.* 2008 | rheumatoid arthritis | DA x PVG.1AV1 | $F_7$ | 422 | 54 | VNTR | N | R/qtl | NA |
| *Rattus norvegicus* | Becanovic *et al.* 2006 | experimental autoimmune encephalomyelitis | DA x PVG.1AV1 | $F_7$ | 314 | 17 | SNP, VNTR | N | MapMaker EXP/QTL, R/qtl | NA |
| *Rattus norvegicus* | Gillett et al. 2010 | TNF production | DA x PVG.1AV1 | $F_{12}$ | 463 | 17 | VNTR | N | R/qtl | NA |
| *Rattus norvegicus* | Huberle *et al.* 2009 | experimental autoimmune encephalomyelitis | DA x PVG.1AV1 | $F_{12}$ | 437 | NA | VNTR | N | R/qtl | NA |
| *Rattus norvegicus* | Jagodic *et al.* 2004 | experimental autoimmune encephalomyelitis | DA x PVG.1AV1 | $F_7$ | 314 | 56 | VNTR | N | MapMaker EXP/QTL, R/qtl | NA |
| *Rattus norvegicus* | Marta et al. 2010 | experimental autoimmune encephalomyelitis | DA x PVG.1AV1 | $F_{10}$ | 794 | 33 | VNTR | N | R/qtl | NA |
| *Rattus norvegicus* | Ockinger et al. 2006 | experimental autoimmune encephalomyelitis | DA x PVG.1AV1 | $F_{10}$ | 794 | 16 | VNTR | N | R/qtl | NA |
| *Rattus norvegicus* | Ockinger et al. 2010 | experimental autoimmune encephalomyelitis | DA x PVG.1AV1 | $F_{10}$ | 794 | 10 | VNTR | N | R/qtl | NA |

| Organism | Study | Phenotype | Cross | Generation | Sample size | Marker number | Marker type | Genome-wide? | Analysis software | Controlled for relatedness? |
|---|---|---|---|---|---|---|---|---|---|---|
| *Rattus norvegicus* | Sheng *et al.* 2005 | experimental autoimmune encephalomyelitis | DA x PVG.1AV1 | $F_7$ | 314 | 19 | VNTR | N | R/qtl | NA |
| *Rattus norvegicus* | Stridh et al. 2010 | experimental autoimmune encephalomyelitis | DA x PVG.1AV1 | $F_{10}$ | 794 | 20 | VNTR | N | R/qtl | NA |
| *Sus domesticus* | Terenina *et al.* 2013 | stress hormone levels, meat production and quality | LW[g] x MS[g] | $F_{10\text{-}12}$ | 100 | 22 | SNP | N | GenABEL | Random polygenic effect, genomic control |

[a] *D2S3* is a selected line derived from an $F_1$ cross of *Aedes aegypti* and *Aedes aegyptiformosus*.

[b] *D2MEB* is a selected line derived from *D2S3* and an outbred Houston strain

[c] *Ibo11* and *DS3* are selected lines derived form outbred founders

[d] IMU Is a wild-caught selected line.

[e] LTI and STI are selected lines derived from partially inbred founders.

[f] HWS and LWS are selected lines derived from outbred founders.

[g] Genetic background unspecified.

[h] Leg1 and Leg2 are selected lines derived from partially inbred founders.

[i] Bro1 and Bro2 are outbred lines.

[j] HR is a selected line derived from outbred founders.

[k] R3 and Eae2 are congenic lines.

[l] In this study, an AIL derived from captive strains of P. maniculatus (BW) and P. polionotus (PO) was backcrossed to each of the founder strains, creating two advanced backcross m lines for fine-mapping QTLs.

NA is used when information was ambiguous or not included in the original article. AILs have also been used to study social behavior (Wiren et al. 2009), ocular characteristics (Prashar et al. 2009; Chen et al. 2011) and gene expression (Ka *et al.* 2013) in chickens, fear-related behavior in mice (McGuire *et al.* 2013) and recombination in yeast (Illingworth *et al.* 2013). However, these studies did not map QTLs and were omitted from the table.

**Table 2**

Mixed model association software.

| Software | Authors | Website |
|---|---|---|
| DOQTL | Aulchenko et al. 2007; Amin et al. 2007; Svishcheva et al. 2012; Belonogova et al. 2013 | http://cgd.jax.org/apps/doqtl/DOQTL.shtml |
| GRAMMAR | | http://www.genabel.org/packages/GenABEL |
| Qtlrel | Cheng et al. 2011 | http://cran.r-project.org/web/packages/QTLRel/ |
| dlmap | Huang et al. 2012 | http://cran.r-project.org/web/packages/dlmap/index.html |
| MASTOR | Jakobsdottir and McPeek 2013 | http://www.stat.uchicago.edu/~mcpeek/software/MASTOR/index.html |
| EMMA | Kang et al. 2008 | http://genetics.cs.ucla.edu/emma/ |
| EMMAX | Kang et al. 2010 | http://genetics.cs.ucla.edu/emmax/ |
| MTMM | Korte et al. 2012 | https://cynin.gmi.oeaw.ac.at/home/resources/mtmm (ASREML) |
| Mendel | Lange et al. 2013 | http://www.genetics.ucla.edu/software/mendel |
| FaST-LMM | Lippert et al. 2011; Listgarten et al. 2013 | http://fastlmm.codeplex.com/ |
| WOMBAT | Meyer 2007; Meyer and Tier 2011 | http://didgeridoo.une.edu.au/km/wombat.php |
| Qxpak.5 | Pérez-Enciso and Misztal 2011 | http://www.icre.cat/Web/OtherSectionViewer.aspx?key=485&titol=Software:%20Qxpak&researcher=255 |
| LMM-Lasso | Rakitsch et al. 2013 | http://webdav.tuebingen.mpg.de/u/karsten/Forschung/research.html?page=research&topic=LMM-Lasso&html=text |
| MLMM | Segura et al. 2012 | https://github.com/bvilhjal/mixmogam (Python), https://cynin.gmi.oeaw.ac.at (R) |
| wgaim | Taylor and Verbyla 2011 | http://cran.r-project.org/web/packages/wgaim/index.html |
| ROADTRIPS | Thornton and McPeek 2010 | http://www.stat.uchicago.edu/~mcpeek/software/ROADTRIPS/index.html |
| GCTA | Yang et al. 2011 | http://www.complextraitgenomics.com/software/gcta/ |
| TASSEL | Yu et al. 2005; Zhang et al. 2010 | http://www.maizegenetics.net/index.php?option=com_content&task=view&id=89&Itemid=119 |
| GEMMA | Zhou and Stephens 2012 | http://home.uchicago.edu/xz7/software.html |

Links to download currently available mixed model association software and corresponding references.