Genetic architecture of quantitative traits in mice, flies, and humans

Jonathan Flint1,3 and Trudy F.C. Mackay2
1Wellcome Trust Centre for Human Genetics, Oxford OX3 7BN, United Kingdom; 2Department of Genetics, North Carolina State University, Raleigh, North Carolina 27695, USA

We compare and contrast the genetic architecture of quantitative phenotypes in two genetically well-characterized model organisms, the laboratory mouse, Mus musculus, and the fruit fly, Drosophila melanogaster, with that found in our own species from recent successes in genome-wide association studies. We show that the current model of large numbers of loci, each of small effect, is true for all species examined, and that discrepancies can be largely explained by differences in the experimental designs used. We argue that the distribution of effect size of common variants is the same for all phenotypes regardless of species, and we discuss the importance of epistasis, pleiotropy, and gene by environment interactions. Despite substantial advances in mapping quantitative trait loci, the identification of the quantitative trait genes and ultimately the sequence variants has proved more difficult, so that our information on the molecular basis of quantitative variation remains limited. Nevertheless, available data indicate that many variants lie outside genes, presumably in regulatory regions of the genome, where they act by altering gene expression. As yet there are very few instances where homologous quantitative trait loci, or quantitative trait genes, have been identified in multiple species, but the availability of high-resolution mapping data will soon make it possible to test the degree of overlap between species.

How much does genetic architecture vary between species?

Within a species, the genetic architecture of different phenotypes appears to be similar, but that similarity may not extend to comparisons between species. How much does knowing about one species inform us about another? Below, we review data on the number of loci, their effect sizes, and how they operate on quantitative phenotypes. For quantitative traits, we define the homozygous effect, \( a \), as one-half of the difference in the mean trait phenotype between individuals that are homozygous for alternative QTL alleles (Falconer and Mackay 1996). Homozygous effects are usually reported as either a fraction of the phenotypic standard deviation of the trait in the mapping population (\( u/\sigma_p \)), or as a fraction of the phenotypic variation (\( \sigma^2_F \)) of the mapping population. In this review, unless otherwise stated, a genetic effect size is meant in this latter sense as a proportion of the phenotypic variance. Heterozygous effects, \( d \), are estimated as the difference between the mean value of the trait for individuals that are heterozygous for the QTL alleles, expressed as a deviation from the average of the two homozygous genotypes. Heterozygous effects are generally expressed as \( d/a \), where \( d/a = 0 \) denotes strict additivity (Falconer and Mackay 1996).

Genetic architecture: Many loci of small effect

In the 1990s, with the availability of genome-wide marker sets, it became possible for the first time to map the chromosomal location of genetic effects that contribute to quantitative variation (Lander and Botstein 1989; Hilbert et al. 1991; Jacob et al. 1991; Todd et al. 1991). The methods adopted depended on the organism: Mouse and fly geneticists could use inbred strains, thereby considerably simplifying the genetics. By setting up a cross between two inbred strains, or by using the inbred derivatives of homozygous small-effect genes in one or more species.
such a cross (recombinant inbred [RI] lines), only two alleles segregated at any locus, and all alleles had the same frequency. In contrast, human geneticists had to make do with what the vagaries of evolution, population history, migration, and mutation had put together.

To map QTLs in rodents, about 100 markers and a few hundred animals were sufficient to provide robust statistical evidence for the location of genetic variants contributing to quantitative phenotypic variation. With falling costs and automated genotyping, plus the development of software packages that made the analyses easily accessible to biologists without statistical expertise, any biologist could, and did, map a phenotype. Consequently, a large literature emerged in the last decade of the 20th century, as biologists turned to genetic mapping to help them identify genes involved in the physiological process or disease that constitutes their special interest.

Mapping QTLs was also relatively easy in flies (and much cheaper than in rodents). Remarkably, the picture that emerged from both organisms was broadly similar. Most of the early mapping studies revealed that a few QTLs with relatively large effects accounted for the bulk of phenotypic variation in most traits. This observation was central to driving attempts to capture common genetic variants segregating in human populations. Critically, the estimates of effect size from model organisms suggested that QTLs explaining as much as 10% of the phenotypic variance might be found, in which case linkage studies using a few hundred families would be sufficient to detect them.

Subsequent studies in mice and flies using larger samples began to reveal a more complex picture. For instance, doubling the number of mice in the mapping population from 800 to 1600 more than doubled the number of QTLs found to influence a measure of fear-related behavior (Turri et al. 2001a, b). A study in Drosophila genotyped over 2000 markers on pools of flies selected from an initial sample of over 10,000 F2 individuals (Lai et al. 2007) and substantially increased the number of QTLs known to influence life span in the fly. Furthermore, high-resolution mapping of QTLs revealed more complex genetic architectures than implicated from the initial genome scans. Single QTLs identified in genome scans typically fractionated into multiple, closely linked QTLs, often with opposite effects. In mouse genetics this finding emerged from studies that chased single effects through linkage and QTL mapping efforts using a higher density of recombinant inbred strains. Estimates of effect size, when expressed as the percentage of phenotypic variance attributable to genetic variation at a locus, vary according to the experimental design. Suppose we identify a QTL that contributes to 1% of the genetic variation in an outbred population. In a model organism we could then set up a breeding program to isolate that chromosomal region so that all other genomic regions are inbred—consequently, the locus now contributes to 100% of the genetic variation. Conversely, QTLs identified in simple systems may have smaller relative effects in outbred populations, because the fraction of the total phenotypic variance explained by a QTL with a given effect depends on both allele and genotype frequencies.

The genetic variance ($\sigma_g^2$) of a strictly additive ($d = 0$) QTL with allele frequencies $p$ and $q$ and homozygous effect $a$ is $2pq^2$ in an outbred population (Falconer and Mackay 1996). In progeny derived from a cross of two inbred lines, the frequency of all segregating alleles is $p = q = 0.5$, whereas in an outbred population, allele frequencies can vary throughout the entire range. Thus, a QTL with homozygous effect $a = 0.5q^2$ will explain 12.5% of the phenotypic variance in an $F_2$ population but only 8% of the variance in an outbred population where the minor allele frequency is 0.2%, and 4.5% of the variance in an outbred population where the minor allele frequency is 0.1. Furthermore, only the two homozygous genotypes are present in a population of RI lines, whereas only one homozygous genotype and the heterozygote segregate in a backcross (BC) population. Thus, the genetic variances for a population of RI lines, $F_2$ individuals, and BC individuals are, respectively, $a^2$, $a^2/2$, and $a^2/4$ (Lynch and Walsh 1998); i.e., in the ratio 4:2:1 for RI:$F_2$:BC.

Nevertheless, although more QTLs were discovered, the results were still inconsistent with the “infinitesimal” model of quantitative genetic variation (Falconer and Mackay 1996; Lynch and Walsh 1998), which posits extremely large numbers of loci with very small effects. Rather, the allelic effects followed an exponential distribution, as proposed by Robertson (1967), with a few loci with moderate to large effects, and increasingly larger numbers of loci with increasingly smaller effects.

Despite the success and predictions of genetic architecture from work on model organisms, by the end of the 20th century human genetic studies of complex phenotypes were founding. A small number of high-profile successes continued to engender optimism (e.g., the identification of the epsilon4 allele of APOE [APOE*4] as a susceptibility gene for Alzheimer’s disease [Corder et al. 1993]), but the majority of genetic studies of human quantitative phenotypes became mired in failures to replicate. One simple and unavoidable interpretation of the relative failure was that human studies were underpowered (Risch and Merikangas 1996). Many of the human linkage studies were powered on the basis that the distribution of effect sizes would be approximately exponential, with a few at the top end explaining perhaps 10% of the phenotypic variation, as found in the model organism work. Could it be the case that genetic architecture of quantitative phenotypes in humans was different from mice and flies?

Discrepant findings are largely attributable to differences in the number and frequencies of alleles segregating in outbred populations compared with those in crosses between inbred strains. Estimates of effect size, when expressed as the percentage of phenotypic variance attributable to genetic variation at a locus, vary according to the experimental design. Suppose we identify a QTL that contributes to 1% of the genetic variation in an outbred population. In a model organism we could then set up a breeding program to isolate that chromosomal region so that all other genomic regions are inbred—consequently, the locus now contributes to 100% of the genetic variation. Conversely, QTLs identified in simple systems may have smaller relative effects in outbred populations, because the fraction of the total phenotypic variance explained by a QTL with a given effect depends on both allele and genotype frequencies.

The genetic variance ($\sigma_g^2$) of a strictly additive ($d = 0$) QTL with allele frequencies $p$ and $q$ and homozygous effect $a$ is $2pq^2$ in an outbred population (Falconer and Mackay 1996). In progeny derived from a cross of two inbred lines, the frequency of all segregating alleles is $p = q = 0.5$, whereas in an outbred population, allele frequencies can vary throughout the entire range. Thus, a QTL with homozygous effect $a = 0.5q^2$ will explain 12.5% of the phenotypic variance in an $F_2$ population but only 8% of the variance in an outbred population where the minor allele frequency is 0.2%, and 4.5% of the variance in an outbred population where the minor allele frequency is 0.1. Furthermore, only the two homozygous genotypes are present in a population of RI lines, whereas only one homozygous genotype and the heterozygote segregate in a backcross (BC) population. Thus, the genetic variances for a population of RI lines, $F_2$ individuals, and BC individuals are, respectively, $a^2$, $a^2/2$, and $a^2/4$ (Lynch and Walsh 1998); i.e., in the ratio 4:2:1 for RI:$F_2$:BC.

Furthermore, the locus-specific effect on the phenotypic variance depends additionally on environmental variance and the number of other QTLs that contribute. In a cross between two inbred rodent strains a detectable locus contributes to about 5% of the phenotypic variation (Flint et al. 2005). In a complex quasi-outbred stock (known as the heterogeneous stock [HS]) (Valdar et al. 2006a), assuming the QTL allele frequencies remain comparable, the same locus is likely to explain about 2.5% of the variance (Flint et al. 2005). In a fully outbred population, the locus will account for much less.
In the last 2 yr, with the advent of adequately powered genome-wide association studies in humans, we have at last reliable estimates of the locus-specific effect sizes that contribute to

variation in complex phenotypes. Genome-wide association analyses of human case-control samples reveal that the genetic architecture of quantitative traits and disease consists of alleles of small effect size. With the exception of the HLA locus on some autoimmune diseases (such as type 1 diabetes) and a few outliers (such as a variant of the HITRA1 gene with an effect on age-related macular degeneration [Dewan et al. 2006]), the per-allele odds ratios are almost all less than 1.5. The genetic architecture of disease and quantitative traits appears to be the same: Small effect sizes characterize quantitative variation in humans. Good examples are height and weight: 54 loci have now been identified that influence height, a great advance on what we knew before the advent of genome-wide association studies (Gudbjartsson et al. 2008). However, each variant contributes less than half a percent to the total phenotype variance. Another way of putting this is that an allele that increases height does so by adding about 0.4 cm. Similarly, for the 10 loci identified that contribute to body mass (Willer et al. 2009), allelic effects range from 0.06 kg/m² to 0.33 kg/m².

The small effect sizes reported in human studies do not necessarily mean that the infinitesimal model is a useful description of quantitative genetic variation in human populations. The important parameter is the distribution of effect sizes. Figure 1 shows the distribution of effect sizes at 140 loci discovered for 20 disease phenotypes. In this case, effect size is expressed as a locus-specific odds ratio (so that an allelic odds ratio of two doubles the risk of developing the condition). Note the skewed distribution of the findings, and note also that the distribution is certainly truncated for smaller effects that studies do not have enough power to detect. Despite sample sizes in excess of 10,000, power to detect the QTLs that influence height and weight is about 10% (Willer et al. 2009). In fact, the loci discovered account for remarkably small amounts of the total phenotypic variance. For height, all known loci explain <5% of the variance. For a measure of obesity (body mass index) known loci account for <1% of the variance (Willer et al. 2009). It is quite possible that there is a long tail of much smaller effects remaining to be discovered.

### Genetic architecture: Context-dependent effects

While it is generally accepted that the effect of a QTL should be considered in the context of the genetic background on which it occurs, the sex of the animal, and within the environmental context in which the phenotype is measured, systematic measures of the extent of context-dependent effects are still lacking. Results from model organisms unambiguously reveal that these effects exist, largely through the ability to control both genotype and environment, but their detection is not always easy.

Epistatic interactions (defined as a statistical interaction between genotypes at two [or more] loci) are difficult to detect in QTL mapping studies, because the large number of pairwise tests for marker–marker interactions imposes a low experiment-wise significance threshold. Large mapping populations are required to sample individuals in the rarer two-locus genotype classes, and segregation of other QTLs can interfere with detecting epistasis between the pair of loci under consideration. Consequently, most QTL mapping studies report only main QTL effects, or evaluate epistasis only between QTLs with significant main effects.

In flies, epistatic interactions have been documented between QTLs affecting the quintessential additive traits, numbers of sensory bristles (Long et al. 1995; Gurganis et al. 1999; Dilda and Mackay 2002), and wing shape (Weber et al. 1999, 2001), as well as longevity (Leips and Mackay 2000, 2002; Mackay et al. 2006) and locomotor behavior (Jordan et al. 2006). Epistatic effects can be as large as main QTL effects, in opposite directions between different pairs of interacting loci, and between loci without significant main effects on the trait.

Epistatic interactions are more readily detectable in diadial crosses, which produce all possible doubly heterozygous genotypes from a collection of mutations affecting the same trait that have been generated in a common background, or by constructing all nine possible two-locus genotypes for two biallelic loci. The few studies using these designs all reveal extensive epistasis: for metabolic activity (Clark and Wang 1997), olfactory behavior (Fedorowicz et al. 1998; Sambandan et al. 2006), climbing behavior (van Swinderen and Greenspan 2005), and locomotor startle response (Yamamoto et al. 2008). Remarkably, epistatic interactions have also been observed between functional polymorphisms affecting alcohol dehydrogenase protein level within the Adh locus (Stam and Laurie 1996).

Where epistasis has been sought in mice, it has in most cases been found. For example, Cheverud and colleagues, in a series of studies on growth, body weight, and morphometry, have identified numerous interactions in a cross between a large inbred mouse strain (LG/f) and a small inbred strain (SM/f) (Cheverud et al. 2001; Klingenberg et al. 2001, 2004; Workman et al. 2002; Yi et al. 2006). Two locus interaction models indicated the presence of epistasis, as others have independently found for the same phenotypes (Brockmann et al. 2000). Methodological improvements have led to the identification of even more epistatic QTLs in the same cross (between LG/f and SM/f). For instance, a Bayesian method, developed to deal with the problem of testing multiple interactions, estimates epistatic effects by generating posterior samples from the joint posterior distribution of all unknowns, including the main and epistatic effects, given the observed data (Yi et al. 2007). Using this approach, Yi and colleagues detected

![Figure 1](http://www.genome.gov/gwastudies/).
several epistatic QTLs that were not found by frequentist methods (Ankra-Badu et al. 2009). However, this should not lead to the conclusion that epistasis is universal. In some circumstances, despite an extensive search, little evidence for epistasis emerges, as, for example, in the case of fear-related behavior in mice (Flint et al. 2004).

In these examples drawn from model organisms, using experimental resources that reduce genetic heterogeneity, the importance of epistasis has to be placed within the genetic context in which it is measured. The power to detect epistasis between two loci depends on the frequency of the rarest genotype, usually the double homozygous genotypes at both loci. In outbred populations, the least common double homozygote genotype may be rare or even absent if the minor allele frequency at both loci is low. Epistasis between two loci is easier to detect in inbred strain crosses, because the double homozygotes at both loci are present in ~12.5% of the individuals of an F2 population and 25% of the individuals in a population of RI lines. Methods that further simplify genetic backgrounds, such as the use of chromosome substitution strains (CSS) or genome-tagged mice (GTM), further enhance the ability to detect epistatic interactions in model organisms.

CSS have one chromosome derived from one inbred parental strain and the remainder derived from a second inbred strain; a CSS panel of mice is created by a laborious process of backcrossing so as to replace each chromosome from a host strain with its corresponding partner from a donor strain (Singer et al. 2004). Genetic differences between a panel member and the host arise from differences on the one chromosome they do not share, thereby allowing a relatively quick and simple test or the location of a QTL. Isolating each chromosome also permits a sensitive test of epistasis, and mapping of 90 traits revealed almost half to have significant deviation from an additive model, strong evidence for epistasis (Shao et al. 2008).

GTM are a panel of over 60 strains, each harboring an ~50 Mb donor segment on a uniform background. Thus, like the CSS, GTM represent a considerable reduction in genetic complexity from a cross between two inbred strains, and the relative contribution of each genetic variant is magnified. Smith and colleagues mapped 97 loci for a variety of behavioral traits, including hyperactivity, anxiety, prepulse inhibition, avoidance, and conditioned freezing (Gale et al. 2008). There was evidence for epistasis at about 60% of the QTLs they identified.

There are two senses in which one can discuss epistatic effects: in reference to the contribution of epistasis to the mean differences between genotypes (“physiological” epistasis) (Cheverud and Routman 1995), or in reference to the contribution of epistatic interactions to the population genetic variance (“statistical” epistasis) (Falconer and Mackay 1996). All of the model organism studies measured the contribution of epistasis in the former sense. Epistatic variance, like other components of genetic variance, depends on both effects and allele frequencies. Therefore, large effects on the mean difference between genotypes may not translate to large effects on the variance, if allele frequencies are low. We cannot compute the expected contributions of the epistasis observed from studies of mutations and QTLs in model organisms to the population genetic variance without knowledge of the allele frequencies in natural populations.

The remarkably large contribution of epistasis in some experimental designs (such as CSS and GTM) stands in stark contrast to the findings from human genome-wide association studies in which, to date, despite the success of finding main effects, there have been no replicated findings of interactions. We should be clear that this does not mean epistasis does not exist (see, e.g., the triallelic inheritance that gives rise to Bardet-Biedel syndrome [Katsanis et al. 2001] and the requirement for mutations at both RET and additional loci to bring about Hirschsprung disease [Emison et al. 2005]). It means, rather, that given the small main effects, the detection of epistasis in complex phenotypes in our own species will require astronomically large numbers of individuals.

Identical arguments apply to the detection of genotype by environment interactions. Again, it is abundantly clear from the model organism work that these context-dependent genetic effects are pervasive. In mice, quantitative analyses indicate that genotype by environment interaction (GEI) effects occur for all phenotypes, and often makes as large, or larger, contribution to the phenotypic variation than the environmental effects (Valdar et al. 2006b). In flies, whenever QTLs have been mapped in different environments, GEI has been observed: for numbers of sensory bristles (Gurkanus et al. 1998, 1999; Dilda and Mackay 2002; Geiger-Thornsberry and Mackay 2002), longevity (Leips and Mackay 2000, 2002; Vieira et al. 2000), competitive fitness (Fry et al. 1998), and immune response to different bacteria (Lazzaro et al. 2006). GEI is mostly attributable to environment-specific expression of QTL alleles, with few cases of opposite effects in different environments (Vieira et al. 2000).

On the face of it, human studies would appear to be one step ahead in having identified a number of variants involved in genes by environment interaction on behavioral phenotypes. One highly cited example is the effect of a functional variant in the promoter of the serotonin transporter gene (SLC6A4 [also known as S-HTT]) that is only, or primarily, manifest in people who had suffered stressful life events (such as the break-up of a long term relationship) (Caspi et al. 2003). However, meta-analysis of attempts at replication combined with a simulation study of the main effects of genotype and environmental effect indicate that, if the effect is real, it is much smaller than often claimed, and will again need very large sample sizes to detect robustly (Munafo et al. 2009).

Genetic architecture: Pleiotropy

Studies of phenotypic covariance between quantitative traits have long suggested the presence of pleiotropy, a finding supported by artificial selection experiments in which a direct response is frequently accompanied by changes in many genetically correlated traits (Wright 1977; Falconer and Mackay 1996). While strong genetic correlations are to be expected from directional pleiotropic effects among functionally related traits such as body weight and length, pleiotropic effects are also found for loci affecting traits that are less obviously functionally related. The importance of genes having pleiotropic effects on multiple quantitative traits is largely unappreciated and has emerged from the study of model organisms.

In flies, evidence for pleiotropic effects of subtle mutations on multiple quantitative traits comes from evaluating the effects of new P-transposable element mutations in a common inbred genetic background (single genetically engineered P-elements can be induced to hop into new genomic locations by simple crosses to a stock containing a chromosomally stable source of transposase) (Venken and Bellen 2005). In mice, similar observations come from studies of effects of alleles created more energetically by targeted deletion of genes (Austin et al. 2004) or by chemical mutagenesis (Hrabe de Angelis et al. 2000; Nolan et al. 2000).
Comparative genetics of quantitative traits

Analyses of effects of P-element mutations on quantitative traits have revealed that many genes that play a major role in early development also affect adult quantitative traits. Mutations in the developmental loci neutralized, Semaphorin-5c, and Catecholamin affect numbers of sensory bristles as well as olfactory behavior (Sambandar et al. 2006; Rollmann et al. 2007). Mutations in the intergenic region between TreI (which affects transepithelial migration of germ cells) and Tre (a trehalase taste receptor also known as Gr5a) affect not only trehalose sensitivity, but also life span and resistance to starvation and heat stress (Rollmann et al. 2006). A mutation in muscleblind is associated with increased aggression (Edwards et al. 2006) and increased resistance to the inebriating effects of ethanol (Morozova et al. 2007), but reduced locomotor startle response (Jordan et al. 2007). Independent evidence for pervasive pleiotropy comes from the substantial overlap of transcripts for which there are correlated responses in expression to selection from the same base population for copulation latency, aggressive behavior, locomotor startle response, and ethanol resistance (Mackay et al. 2005; Edwards et al. 2006; Jordan et al. 2007; Morozova et al. 2007).

In mice, projects to deliver engineered knockouts for every gene in the genome (Collins et al. 2007) are beginning to provide new insights regarding the extent of pleiotropy. Phenotypic data from 250 knockout strains acquired by the knockout project from two companies (Deltagen and Lexicon) are publicly available from the MGI database (http://www.informatics.jax.org/external/ko/). Although not an unbiased selection of mutants, the data are the only instance where a large number of knockouts have been subjected to a broad range of measurements. Table 1, summarizing results for 12 phenotypes, shows that about 7% of mutants have an abnormal phenotype. There seems no reason to expect this figure to be any less for any of the hundreds of other traits that could, and eventually will be examined. If so, pleiotropy will be the rule, not the exception.

Demonstrating pleiotropy in QTL studies in mice and flies is more difficult since QTLs typically contain multiple genes: Disentangling coincidence of location from pleiotropic action is hard. In flies, this problem can be overcome by using genetic association. The structure of linkage disequilibrium in Drosophila is exceptionally fine grained and decays rapidly in regions of normal recombination; so rapidly, in fact, that polymorphic sites adjacent to each other can be in linkage equilibrium (Carbone et al. 2006). This makes it possible to differentiate the effects of single-nucleotide polymorphisms within the same gene. Several studies have evaluated associations of polymorphisms in candidate genes with more than one quantitative trait: the achaete-scute complex (Mackay and Langley 1990; Long et al. 2000), scabrous (Lai et al. 1994; Lyman et al. 1999), Delta (Long et al. 1998), and hairy (Robin et al. 2002) for abdominal and sternopleural bristle number; the epidermal growth factor receptor for wing shape (Pasloss and Gibson 2004) and cryptic variation for photoreceptor determination (Dworkin et al. 2003); Dopa decarboxylase for longevity (De Luca et al. 2003) and locomotor startle response (Jordan et al. 2006); and Catecholaminies up for locomotor behavior, longevity, starvation resistance, and abdominal and sternopleural bristle number (Carbone et al. 2006). In each case different polymorphic sites were independently associated with the different traits. These observations indicate that pervasive pleiotropy does not necessarily impose evolutionary constraints in the form of strong genetic correlations between traits.

Equivalent high-resolution association methodologies are not available in mice. Instead, statistical approaches have been applied to inbred cross data and typically find evidence for pleiotropy, for example, as in studies of weight and morphometry (Brockmann et al. 2000; Leamy et al. 2002; Wolf et al. 2005, 2006; Pavlicev et al. 2008). However, these methods leave open the possibility that pleiotropy at a locus is attributable to multiple genes of different function that happen to reside next to each on the genome.

In human genetic association studies, relatively low linkage disequilibrium makes it possible to argue that pleiotropy is due to the presence of a single gene influencing multiple conditions. In some cases, pleiotropy involves diseases, or physiological processes that have common features. For instance, type 1 diabetes and celiac disease share risk regions (Hunt et al. 2008; Smyth et al. 2008) and polymorphisms in IL23R contribute to susceptibility to three autoimmune disorders: psoriasis, Crohn’s disease, and ankylosing spondylitis (Duer et al. 2006; Burton et al. 2007; Cargill et al. 2007), and a SNP within the intron of CDKL1 has been associated with type 2 diabetes (T2D) and insulin secretion defects (Gudmundsson et al. 2007; Scott et al. 2007; Zeggini et al. 2007, 2008). In other cases, genetic relationships are beginning to emerge between traits that had not previously been associated: a SNP in HNF1B is associated with protection from type 2 diabetes (Winckler et al. 2007) and for prostate cancer (Gudmundsson et al. 2007); SNPs in TCF7L2 are associated with both type 2 diabetes (Grant et al. 2006) and colon cancer (Slattery et al. 2008); SNPs in JAZF1 are associated with both type 2 diabetes (Zeggini et al. 2008) and prostate cancer (Thomas et al. 2008).

Table 1. Percentage of abnormal phenotypes (at P < 0.05) for 250 mouse knockouts

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open field activity</td>
<td>18.9</td>
</tr>
<tr>
<td>Albumin</td>
<td>7.1</td>
</tr>
<tr>
<td>Alkaline phosphatase</td>
<td>7.1</td>
</tr>
<tr>
<td>Blood urea nitrogen</td>
<td>9.1</td>
</tr>
<tr>
<td>Calcium</td>
<td>6.3</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>10.2</td>
</tr>
<tr>
<td>Glucose</td>
<td>3.9</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>4.5</td>
</tr>
<tr>
<td>Phosphorous</td>
<td>7.1</td>
</tr>
<tr>
<td>Platelets</td>
<td>5.8</td>
</tr>
<tr>
<td>Triglycerides</td>
<td>6.7</td>
</tr>
<tr>
<td>White blood cells</td>
<td>6.2</td>
</tr>
</tbody>
</table>

Data are taken from http://www.informatics.jax.org/external/ko/.

How much does genetic architecture vary between phenotypes?

Evidence from model organisms strongly argues that the distribution of effect size of common variants is the same for all phenotypes. Figure 2 shows a summary of the effect sizes for 98 phenotypes mapped in 2000 heterogeneous stock mice (Valdar et al. 2006a). The phenotypes are classified into six categories: behavior, biochemistry, hematology, immunology, metabolism, and respiratory physiology (as measured in a plethysmograph). There is no significant variation among the categories (F = 1.8, df = 5, P = 0.109). Mapping multiple traits in chromosome substitution strains in mice gives a similar picture of the genetic architecture of quantitative phenotypes (Shao et al. 2008). CSS studies of 90 traits in a mouse panel and 54 traits in a rat panel identified multiple QTLs for each phenotype (Shao et al. 2008). The average
phenotypic effects for a CSS were similar across all phenotypes in both mouse and rat panels.

One of the surprises of the Wellcome Trust Case Control Consortium analysis of seven common diseases (Wellcome Trust Case Control Consortium 2007) was that although the same number of individuals was used for each disease (2000) and a common set of controls (3000), the findings for each phenotype varied. While for Crohn’s disease (an inflammatory bowel disorder), nine loci exceeded an association signal at \( P < 5 \times 10^{-7} \), only one locus was found for coronary artery disease and bipolar disease. In part, the difference in detection can be blamed on experimental design. The WTCCC uses a set of common controls that are not screened for disease. Consequently, for common disorders, such as coronary artery disease, controls include a proportion of affected individuals, and this misclassification will reduce power to detect association (Wellcome Trust Case Control Consortium 2007; McCarthy et al. 2008). But this does not rule out the possibility that failure to detect effects might reflect a different genetic architecture. It could be that the effect sizes of common variants are much less in psychological measures than in anthropometric traits.

It is also possible that the model organism gives an unrepresentative picture. Almost all of the work carried out in mice and flies on quantitative traits uses experimental designs that detect common alleles, not rare variants. In populations derived from crosses between inbred strains, there are a small number of alleles (in most designs only two) at each locus, and the frequencies are large and vary little (in an intercross, each allele will have a frequency of about 0.5). In outbred populations used by human geneticists, this is not the case. Genome-wide association studies detect common variants, so the comparison with the model organism results is probably fair. The impact of rare variants on the genetic architecture of quantitative phenotypes remains relatively unexplored in both model organisms and humans. However, there are clear indications from resequencing studies in humans that it will turn out to be important (Cohen et al. 2004, 2005; Kotowski et al. 2006; Romeo et al. 2007; Ji et al. 2008).

Could the contribution of rare variants differ between phenotypes? There is some evidence from human studies to support this view, coming from studies of the role of copy number variants (CNVs). The extent and distribution of CNVs in the human genome is still being investigated: Current estimates indicate large-scale (>50 kb) CNVs are rare, affecting <5% of the genome, but about 1300 smaller CNVs have minor allele frequencies >1% (McCarroll et al. 2008). Genome-wide association studies typically use array-based methods that can also detect CNVs. Rare CNVs are now known to be common in autism (Sebat et al. 2007) and schizophrenia (International Schizophrenia Consortium 2008; Stefansson et al. 2008; Walsh et al. 2008; Xu et al. 2008) than in controls; in contrast, few common variants have been found in schizophrenia and autism. A similar contribution of rare CNVs has not been reported for other diseases, suggesting that there may be important differences between some phenotypes. Possibly the lower reproductive fitness of schizophrenia and autism have contributed to the concentration of rare variants as a predisposing cause of these psychiatric illnesses, a finding that implies that it may be difficult to model these illnesses genetically using engineered mice.

Are the same genes involved in the same phenotype in different species?

Despite substantial advances in mapping QTLs, the identification of the QTGs and ultimately the QTN sequence variants has proved more difficult. Thus, our information on the overlap at the molecular level is relatively limited. But, the information that we have indicates that sequence positions giving rise to phenotypic variation are not homologous between species. The relevant finding is that, unlike coding regions, many regulatory regions are not conserved between species. One of the surprises of the Encyclopedia of DNA Elements (ENCODE) Project, which set out to catalog functional elements in 1% of the genome, was the excess of functional elements lying outside regions of sequence conservation (The ENCODE Project Consortium 2007).

In flies, the requirement for dense molecular marker genotypes and, ideally, complete DNA sequence data, for alleles of candidate genes sampled from a natural population has limited the number of association studies performed. However, in all studies reported to date, both common and rare QTNs have been associated with quantitative traits and occur in presumed regulatory regions, as well as noncoding, noncoding polymorphisms in coding regions (Mackay and Langley 1990; Long et al. 2000; Robin et al. 2002; Palsson and Gibson 2004; Jordan et al. 2006). Although the total number of studies is small, there is a tendency for less common polymorphisms associated with quantitative traits to...
occur in coding regions and have large effects, while intermediate frequency polymorphisms often occur in presumed regulatory regions (or synonymous SNPs in coding regions) and have smaller effects. All types of polymorphisms have been associated with variation in fly quantitative traits: SNPs, small insertions/deletions, complex SNP/copy number variants, and transposable elements. A more comprehensive picture of the nature and extent of polymorphisms associated with quantitative traits in the fly model will emerge in the near future, when many traits are assessed for a panel of 192 inbred strains for which complete sequences have been determined.

If the position of the variant is not conserved, the gene may well be. Claims that variation within the same gene could contribute to phenotypic variation in both rodents and humans have sustained attempts to map phenotypes in model organisms. However, given the assumption that many physiological processes are conserved between these species, and that, as we have shown, a very large number of loci are likely to be contributing to human phenotypic variation, it is surprising that so few examples of common QTLs have been reported.

One of the earliest examples is the detection of a locus that modifies the severity of cystic fibrosis. A locus was identified on mouse chromosome 7 that modified intestinal disease in homozygous mutant mice (Rozmahel et al. 1996) and subsequently a similar modifier was mapped to human chromosome 19, in the region syntenic to the mouse locus (Zielenski et al. 1999). Unfortunately, the gene involved remains unknown. Polymorphisms in the same gene, cytotoxic T-lymphocyte-associated protein 4 gene (CTLA4), have been shown to contribute to autoimmune disorders in humans and mice (Ueda et al. 2003). A noncoding 3′ region of CTLA4 increases risk of the Graves’ disease, autoimmune hypothyroidism, and type 1 diabetes; in a mouse model of type 1 diabetes, susceptibility is associated with variation in Cita4 gene splicing with reduced production of one splice form. Finally, QTL mapping identified the same susceptibility gene (Fgrs3) contributing to immunologically mediated glomerulonephritis in humans and rats (Aitman et al. 2006). Remarkably, in both rats and humans, copy number variation in the homologous gene was associated with the disease, suggesting that copy number variation may be a conserved molecular mechanism underlying complex and quantitative phenotypes.

A number of other genes have been identified in rodents and then tested by association in humans, but results have often been equivocal. For example, evidence from QTL mapping and analyses of knockouts implicates Trnf4/4 in diet-induced atherosclerosis in mice (Wang et al. 2005b); however, genetic association studies have not provided consistent evidence for its involvement in humans (Koch et al. 2008). Similar difficulties are found for the involvement of type II SH2 domain-containing inositol 5-phosphatase in the metabolic syndrome (Kaisaki et al. 2004; Marcano et al. 2007) and regulators of G protein signaling in anxiety (Fullerton et al. 2008).

The more general question, of the extent of overlap between QTLs identified in different species, remains unclear. Data from mouse QTL and human linkage studies were used by Paigen and colleagues to show that more than half of human atherosclerosis QTLs are located in regions homologous to mouse QTLs (Wang et al. 2005a). Human linkage results are, however, notoriously unreliable and, as Risch and colleagues pointed out, concordance is unlikely given that the genetic effect in humans depends on disease allele frequencies and such allele frequencies are unpredictable (Risch et al. 1993).

The more recent GWAS data present a new opportunity to revisit this question. The relatively large number of genes now identified in humans, and progress in mapping QTLs at high resolution in mice (Valdar et al. 2006a) should allow comparisons of the extent of overlap. These analyses have yet to be reported.

Overlaps between mammals and Drosophila at a QTL level have not been reported. At a mechanistic level there are examples of genes involved in similar processes, including learning and memory (Dubnau and Tully 1998). More intriguing is the possibility that, at a systems level, there may be commonalities. Could networks of coregulated genes, representing physiological processes, be common between species, and networks accessed by their genetic determinants? A recent study of liver and adipose gene expression in human and mouse tissue suggested that this could be the case: A common macrophage-enriched network was found (Chen et al. 2008; Emilsson et al. 2008). While it is not clear to what extent this may simply reflect common histological characteristics of fatty liver infiltrated by macrophages, the idea of searching for commonality at the systems level is intriguing. As transcriptomic data become more widely available, it will become possible to explore this idea.

Conclusions

Our review of the genetic architecture of quantitative traits has highlighted considerable agreement between analyses carried out in flies, rodents, and our own species. A picture emerges of large numbers of loci of small effect, although the distribution of allelic effects is not consistent with the “infinitesimal” model of quantitative genetic variation (Falconer and Mackay 1996; Lynch and Walsh 1998). Discrepant findings between species are largely attributable to differences in experimental design. For example, geneticists working with model organisms frequently exploit the opportunities provided to analyze inbred strains or derived crosses, while human geneticists are restricted to working with outbred populations. Differences in the number and frequencies of alleles segregating in outbred populations, compared with those in crosses between inbred strains, contribute to differences in inferred quantitative genetic architecture.

Systematic measures of the extent of context-dependent effects in human populations are still lacking, despite robust evidence from model organisms of their importance. Again, extrapolating findings from one species to another is problematic: We cannot compute the expected contributions of the epistasis observed from studies of mutations and QTLs in model organisms to the population genetic variance without knowledge of the allele frequencies in natural populations. To restate, epistatic variance depends on both the size of genetic effects and allele frequencies. Large effects on the mean difference between genotypes may not translate to large effects on the variance if allele frequencies are low. The same arguments apply to studies of genes by environment interaction.

We have highlighted what we believe to be the largely unappreciated importance of pleiotropy. We do not refer here to the expected correlation of genetic effects between related phenotypes, such as height, weight, or morphology. Nor do we refer to pleiotropy consequent upon linkage, that is to say, when mapping of two apparently unrelated traits identifies the same genetic locus because of the close proximity on the genome of two genes that have separate and independent functions. We refer, rather, to the observation from studies of mutations in flies and mice, and now, increasingly, from high-resolution mapping studies in humans,
that one gene may have a role in two or more phenotypes previously thought to be unrelated. An appreciation of pleiotropy, of its functional analysis, and how it works, will be critical for interpretations of the functional analysis of every gene in the vertebrate genome, accessible through the creation of a complete set of mouse knockouts. Finally, when we consider the molecular basis of quantitative traits, we find an important role for noncoding DNA. Many quantitative trait nucleotides lie in what are assumed to be regulatory regions of the genome, though whether this is necessarily the regulation of gene expression or of other features of genome function remains unknown. Given the differences in gene function and population structure between the organisms studied, it is not surprising that there is little evidence for conservation of the position of QTLs between species; however, there has been no thorough search for such loci, largely because until recently the necessary resources of high-resolution high-quality QTL maps were not present. Finally, as more data accumulate and, particularly, with access to whole-genome sequences of multiple individuals in any species, the coming years will undoubtedly see remarkable advances in our understanding of the molecular basis of the genetic architecture of quantitative traits.

Acknowledgments

J.F. is supported by the Wellcome Trust, and T.F.C.M. by grants from the National Institutes of Health.

References


Alemayehu, A., Breen, L., Krenova, D., and Printz, M.P. 2002. Reciprocal rat larly, with access to whole-genome sequences of multiple indi- position of QTLs between species; however, there has been no not surprising that there is little evidence for conservation of the


Marcano, A.C., Burke, B., Gungadoo, J., Wallace, C., Kaisaki, P.J., Woon, P.Y.,
analysis of an insertion polymorphous phosphate-1 like (INPP1_PL, SHIP2) variants with
McCarroll, S.A., Kuruvilla, F.G., Korn, J.M., Cawley, S., Nemesh, J., Wysoker,
detection and population-genetic analysis of SNPs and copy number variation.
McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J.,
studies for complex traits: Consensus, uncertainty and challenges. Nat.
McPherson, R., Fertemilidis, A., Kavaslar, N., Stewart, A., Roberts, R., Cox,
quantitative trait loci affecting wing shape of Drosophila melanogaster.
Moehring, A.J. and Mackay, T.F. 2004. The quantitative genetic basis of male
studies: Application to nonobese diabetic (NOD) mouse and human
trait locus mapping in laboratory mice derived from a replicated
study of anxiety in laboratory mice.
environment interactions at the serotonin transporter locus. Biol.
Nolan, P.M., Peters, J., Strivens, M., Rogers, D., Hagan, J., Spurr, N., Gray,
genome-wide association study of antibody production on mouse chromosome 8.
Noller, M.I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J.,
studies for complex traits: Consensus, uncertainty and challenges.
environment interactions at the serotonin transporter locus. Biol.
environment interactions at the serotonin transporter locus. Biol.
environment interactions at the serotonin transporter locus. Biol.
haplotypes in a black-footed ferret: A population genetic tool for
haplotypes in a black-footed ferret: A population genetic tool for
haplotypes in a black-footed ferret: A population genetic tool for
haplotypes in a black-footed ferret: A population genetic tool for
haplotypes in a black-footed ferret: A population genetic tool for