QUANTITATIVE TRAIT LOCI IN *DROSOPHILA*

Trudy F. C. Mackay

Phenotypic variation for quantitative traits results from the simultaneous segregation of alleles at multiple quantitative trait loci. Understanding the genetic architecture of quantitative traits begins with mapping quantitative trait loci to broad genomic regions and ends with the molecular definition of quantitative trait loci alleles. This has been accomplished for some quantitative trait loci in *Drosophila*. *Drosophila* quantitative trait loci have sex-, environment- and genotype-specific effects, and are often associated with molecular polymorphisms in non-coding regions of candidate genes. These observations offer valuable lessons to those seeking to understand quantitative traits in other organisms, including humans.

INTROGRESSION
Transfer of genetic material from one strain to another by repeated backcrosses. With marker-assisted introgression, markers that distinguish the parental strains are used to track the desired interval and select against the undesired genotype.

Department of Genetics, Box 7614, North Carolina State University, North Carolina 27695, USA. e-mail: mackay@unity.ncsu.edu

The ease with which Mendelian and quantitative traits give up their genetic secrets is inversely proportional to the relative importance of the two classes of trait for human health, agriculture, evolution and even functional genomics. Although devastating to the possessor, highly deleterious alleles that cause inborn errors of metabolism and other single gene disorders are rare in the general population. By contrast, susceptibility to common diseases such as atherosclerosis, arthritis, diabetes, hypertension and schizophrenia is affected by multiple genetic factors and by the environment. These diseases are therefore quantitative traits (FIG. 1), and affect a large proportion of the human population. Similarly, individuals vary quantitatively in their response to drug therapy. There is great excitement in the human genetics community and the pharmaceutical industry that susceptibility loci for common diseases and individual variation in drug response can be identified and the molecular basis for this variation determined. This knowledge will herald a new era of personalized medicine in which environment-specific risk factors for common diseases are assessed for individual genotypes (and hopefully avoided by the patient) and pharmaceutical treatment is genotype dependent.

Similar arguments apply to the agriculture industry, in which most characters of economic importance in domestic animal and crop species are quantitative. There is a long history of success in improving productivity traits by selective breeding for favourable phenotypes. Knowledge of the allelic status at each locus affecting these traits will greatly facilitate this process, and will enable INTROGRESSION of favourable alleles from other strains, while simultaneously eliminating deleterious alleles.

Variation for quantitative traits is the raw material on which the forces of evolution act to produce phenotypic diversity and adaptation. Major research efforts in evolutionary quantitative genetics are aiming to determine how genetic variation for adaptive quantitative traits is maintained in natural populations; whether the loci at which variation occurs within a population are the same as those that cause divergence between populations and species; and how the answers to these questions depend on the relationship of the trait to the ultimate quantitative trait — reproductive fitness. So a comprehensive understanding of the evolutionary process is contingent on a detailed description of the molecular genetic basis of variation for quantitative traits in natural populations.

The complete genome sequences of the yeast *Saccharomyces cerevisiae*[1], the nematode *Caenorhabditis elegans*[2] and the fruitfly *Drosophila melanogaster*[3] reveal that a large fraction of these genomes is uncharted phenotypic territory. In *Drosophila*, for example, only 2,500 of the 13,600 genes and predicted genes (18%) have been characterized by classic genetic and molecular methods[3]. An important challenge for the future is to devise ways of determining the phenotypic effects of

© 2001 Macmillan Magazines Ltd

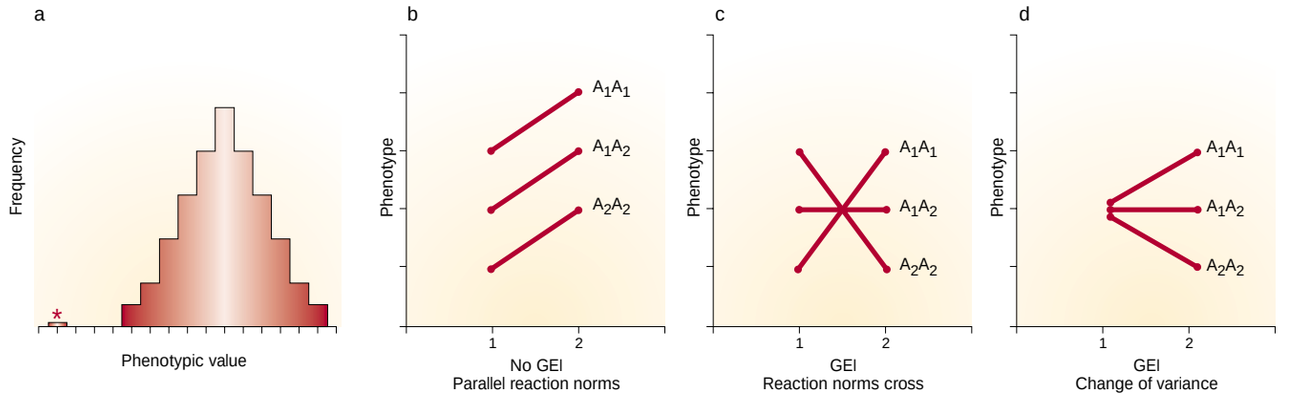


Figure 1 | **Characteristics of quantitative traits. a** | The plot of phenotypes of a quantitative trait forms a continuously graded series, often approximating a statistical normal distribution. The continuous variation in phenotypes is partly attributable to the joint segregation of alleles at multiple quantitative trait loci (QTL), and by truly continuous environmental variation. By contrast, variation for Mendelian characters, represented by the asterisk, is discrete and has a simple genetic basis. **b** | QTL effects are typically sensitive to changes in the environment. Here, the phenotypic value of each of three genotypes at a single QTL is plotted in two different environments (1 and 2). The environments can be the two sexes, physical environments (for example, temperature), or alternative genotypes at a second QTL that affect the trait. The line joining the phenotypes of the same genotype in different environments is the 'norm of reaction' of the genotype. Here, there are differences in the mean value of the quantitative trait between the two environments, but alternative genotypes at the QTL react in the same manner to the change in mean. The rank order and absolute magnitude of the difference between the genotypes remains constant, and the norms of reaction are parallel. In this case, there is no genotype-by-environment (GEI) interaction. **c** | Genotype-by-environment interactions occur when there is a change of rank order of the QTL effects in the two environments. Changes in rank order of QTL effects are attributable to ANTAGONISTIC PLEIOTROPY at the level of QTL. **d** | Interactions also occur when there is a change of variance of the QTL effects with sex, environment, or genetic background. Changes in variance of QTL effects are due to CONDITIONAL NEUTRALITY at the level of QTL.

uncharacterized and predicted genes. Conventional screens for mutations with large phenotypic effects can lead to the identification of function for a biased sample of genes — mutating one gene in a pathway in which there is functional redundancy might not cause a major effect on the phenotype. Furthermore, homozygous lethal mutations define loci that are essential for viability, but less severe mutations at these loci may have unknown and unexpected pleiotropic effects on morphology, physiology and behaviour. So, genetic screens for mutations with subtle, quantitative effects and genetic analysis of naturally occurring variation for quantitative traits will be important components of the functional genomics tool kit.

Until very recently, the genetic basis of variation for quantitative traits was inferred solely from statistical estimates of correlations between relatives, response to artificial selection and changes of mean and VARIANCE of the trait on inbreeding and crossing[4,5]. To reap the benefits of a thorough understanding of quantitative traits, we must lift this statistical fog[6] and describe quantitative genetic variation in terms of complex genetics (FIG. 1). Specifically, a full understanding of the genetic architecture of a quantitative trait will require answers to the following questions. What are the loci at which mutational variation affecting the trait occurs? What are the spontaneous mutation rates at these loci? What loci affect naturally occurring variation within and between populations of a single species, and between species? What are the homozygous and heterozygous effects of alleles at these loci? Are the effects of the individual loci on the final phenotype independent (additive), or is the effect of multiple loci on the phenotype nonlinear (epistasis)? What is the effect of quantitative trait locus (QTL) alleles on multiple quantitative traits, including reproductive fitness (pleiotropy)? How do the homozygous, heterozygous, epistatic and pleiotropic QTL effects vary between the sexes and in a range of ecologically relevant environments? What defines a QTL allele at the molecular level? What are QTL allele frequencies within and between populations?

At present, detailed genetic dissection of quantitative traits is most feasible in genetically tractable and well-characterized model systems. *Drosophila melanogaster* is one of the model organisms that provides us with all the tools necessary for identifying QTL and characterizing them at the molecular level[7] (FIG. 2). Over eight decades of research on this organism have provided us with a library of stocks that bear mutations at single loci and deficiency chromosomes that cover around 70% of the genome. The *P* transposable element has been harnessed as a transformation vector and modified for efficient insertional mutagenesis, analysis of tissue-specific expression patterns, general and targeted overexpression, and, most recently, homologous recombination[8]. Highly polymorphic molecular markers with known physical map locations are available for recombination mapping of QTL. Finally, *Drosophila* has been used as a model organism for quantitative and molecular population genetic studies for over 40 years, providing a historical framework on which to juxtapose and interpret recent progress.

Quantitative trait locus mapping methods

In principle, QTL mapping is deceptively simple. All that is required are two inbred strains in which different alleles at loci affecting variation in the trait of interest are fixed, and a polymorphic molecular marker linkage map. (Usually the parental strains will have different mean values for the trait, but this is not necessary, as two strains with the same mean phenotypic value can vary

ANTAGONISTIC PLEIOTROPY
Alternative homozygous genotypes (A_1A_1, A_2A_2) have opposite phenotypic effects under different conditions.

CONDITIONAL NEUTRALITY
The difference between quantitative trait loci genotypes is only expressed under some conditions.

VARIANCE
A statistic to quantify dispersion about the mean. In quantitative genetics, the phenotypic variance, V_P, is the observed variation of the trait in a population. V_P is partitioned into components due to variation in the additive (V_A) dominance (V_D) and epistatic (V_I) genetic variance, the variance attributable to the environment (V_E), and gene–environment correlations and interactions.

© 2001 Macmillan Magazines Ltd

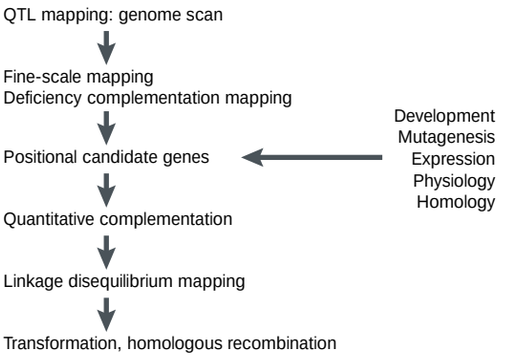


Figure 2 | **The road to genetic dissection of quantitative traits in *Drosophila*.**

genetically owing to complementary patterns of positive and negative allelic effects.) Then one creates a mapping population of BACKCROSS, F_2, recombinant inbred lines (RIL) or other segregating generations derived from the parental strains, and determines the phenotype and multi-locus genotype of each of the individuals in the mapping population (FIG. 3). At its simplest, QTL mapping involves going through the genome, one marker at a time, dividing the individuals into marker genotype classes, and doing a statistical test to determine whether there is a significant difference in phenotype between the marker genotype classes. If there is such a difference, then the QTL is linked to the marker. This procedure, as described, underestimates the effect of the QTL by an amount that is proportional to the distance of the QTL from the marker locus, but this problem is readily overcome by mapping the QTL relative to two flanking markers (interval mapping)[4,5].

The logic of QTL mapping is not new, and was used nearly 80 years ago[9] to map a QTL associated with seed size by linkage to a pigment locus in the bean *Phaseolus vulgaris*. The main practical limitation to implementing QTL mapping was the lack of a large number of closely spaced Mendelian marker loci in most species. It should not come as a surprise, then, to learn that *Drosophila*, which has an abundance of visible mutant stocks and for which we can easily construct designer genotypes, has led the way in the genetic dissection of QTL. Whole chromosome substitution lines were used in *Drosophila*[10] nearly four decades before this technique was proposed in the context of QTL mapping in the mouse[11]. Moreover, *Drosophila* was the first organism in which QTL were localized to sub-chromosomal segments by introgression[12] and interval mapping combined with PROGENY TESTING[12,13]. These pioneering studies were limited by the number of available visible markers and the problem that the markers used were deleterious and affected many quantitative traits, often including the one of interest. Two technical advances have opened the door for detailed characterization of the genetic architecture of quantitative traits in *Drosophila* and in other organisms: the discovery of abundant, polymorphic, neutral molecular markers and the development of sophisticated statistical methods for mapping QTL[14–17] (FIG. 3, BOX 1).

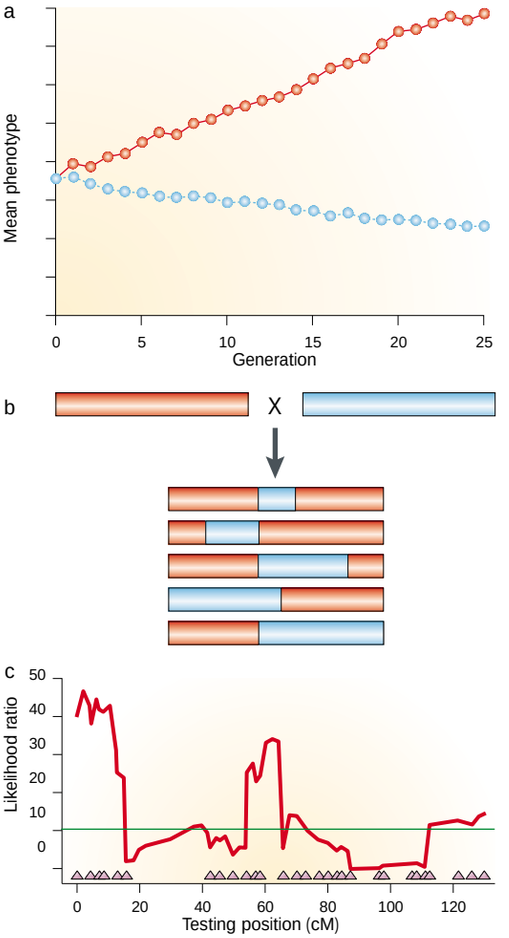


Figure 3 | **Quantitative trait locus mapping. a** | Quantitative trait locus (QTL) mapping requires parental strains (red and blue plots) that differ genetically for the trait, such as lines created by divergent artificial selection. **b** | The parental lines are crossed to produce individuals or strains that contain different fractions of the genome of each parental line. The phenotype for each of the recombinant individuals or lines is assessed, as is the genotype for multiple markers that are polymorphic between the parental strains. Recombinant inbred lines (RIL) are preferred over F_2 and backcross designs when traits with low HERITABILITIES are the focus of the study. Because multiple phenotypic measurements can be obtained for each RIL genotype, more accurate estimates of the genotypic values of the lines are possible than with the same number of F_2 or backcross individuals. Using RIL also facilitates the estimation of QTL by sex and QTL by environment interaction effects, because the same set of genotypes can be assessed in both sexes, and in different environments. **c** | Composite interval mapping[15] evaluates the probability that a marker or an interval between two markers is associated with a QTL affecting the trait, while simultaneously controlling for the effects of other markers on the trait. This method considerably increases the power to detect QTL, by decreasing the within-marker-class phenotypic variation[4]. The results of such an analysis are presented as a plot of the LIKELIHOOD RATIO test statistic against the chromosomal map position, in recombination units (cM). Positions of the markers are shown as triangles. The horizontal line marks the 5% significance threshold[80,81]. Likelihood ratios above this line are formally significant, with the best estimate of QTL positions given by the chromosomal position corresponding to the highest significant likelihood ratio.

BACKCROSS
The cross of the F_1 progeny of two parental strains to either of the two parents.

PROGENY TEST
Crossing an individual at random to a number of unrelated individuals in the population, and determining the mean phenotype of the progeny. The progeny mean phenotype is a more accurate measure of the genetic merit of the tested individual than the individual's own phenotype.

HERITABILITY
The fraction of the phenotypic variance attributable to additive genetic variance (V_A/V_P).

LIKELIHOOD RATIO
A method for hypothesis testing. The maximum of the likelihood that the data fit a full model of the data is compared with the maximum of the likelihood that the data fit a restricted model and the likelihood ratio (LR) test statistic is computed. If the LR test statistic is significant, the full model provides a better fit to the data than does the restricted model.

© 2001 Macmillan Magazines Ltd

PERMUTATION TEST
A statistical test in which the data are randomized many times to determine the statistical significance of the experimental outcome (in this case, the association of a quantitative trait phenotype and a multi-locus marker).

Box 1 | Multiple tests

Mapping of quantitative trait loci (QTL) is plagued by the statistical problem of multiple tests. If only one test is done, 5% is conventionally accepted as a significance threshold. That is, the result of a statistical test is deemed 'significant' if its probability of occurring by chance alone is 5% or less. However, if many tests are done on the same data, as occurs when associations between multiple markers and the quantitative trait are considered, 5% of the associations are expected to be significant by chance. So, the significance threshold for the experiment must be adjusted downwards to correct for the number of independent tests. PERMUTATION TESTS[80,81] are typically used to determine empirical 5% significance thresholds for declaring significant QTL.

How many QTL affect variation in a quantitative trait? This simple question is not easy to answer. The number of QTL mapped in any one experiment is always a minimum number. It may be obvious, but is often forgotten, that QTL can only be mapped if there are allelic differences between the two parent strains used to construct the mapping population. To the extent that these strains are a limited sample of the existing genetic variation, it should not be surprising if different QTL are found in different studies. Designs that use parent strains derived from divergent artificial selection experiments will, however, contain a more representative fraction of segregating variation than will two random inbred lines. Furthermore, an axiom of QTL mapping is that the harder you look, the more QTL you find. There are two reasons why the number of QTL is expected to increase with the sample size (number of backcross or F_2 individuals, or RIL). First, the lower limit for the magnitude of QTL effect that is detectable in any mapping experiment is set by the sample size and increasing the sample size allows mapping of QTL with smaller effects. Second, the precision of mapping depends on the ability to separate linked QTL by recombination. In general, the larger the sample size, the more recombinant events, and the more QTL that are detectable, given a sufficiently dense marker map.

Table 1 | Variation for quantitative traits is due to multiple loci

Trait*	Chromosome(s)‡	Number of QTL	References
Sternopleural bristle number	3	17	13
Sternopleural bristle number	1,2,3	22	19,21
Abdominal bristle number	1,2,3	26	19,21
Longevity	1,2,3	19	23–25
Wing shape	3	11	22
Competitive fitness§	1,2,3	6	29
Reproductive success§	1,2,3	2	28
Morphology of male genital arch	1,2,3	19	27

*Most studies mapped quantitative trait loci (QTL) affecting variation within *D. melanogaster*. The exception was the study of QTL affecting variation in morphology of the posterior lobe of the male genital arch between *D. simulans* and *D. mauritiana*. ‡*Drosophila* has three major chromosomes; the tiny fourth chromosome represents approximately 1% of the genome and does not recombine. §The low number of QTL for these fitness traits is probably a consequence of the reduced power to detect QTL for traits with high environmental variance.

Quantitative trait locus mapping results

The results from several recent studies are summarized in TABLE 1 and show that quantitative traits are truly polygenic. Perhaps more interesting than QTL counts are inferences about their effects. In quantitative genetics theory, it is often assumed for mathematical convenience that genetic variation for quantitative traits is caused by a very large number of QTL with very small and equal allelic effects[4,5]. However, the 'infinitesimal' model makes little sense genetically, as it predicts all loci equally affect all conceivable quantitative traits. Alan Robertson[6] proposed that the distribution of allelic effects should be more nearly exponential, whereby a few loci have large effects and cause most of the variation in the traits, with increasingly larger numbers of loci with increasingly smaller effects making up the remainder. The distributions of QTL effects for *D. melanogaster* sensory bristle number[10,13,18–21], wing shape[22], longevity[23–25], and for the difference in size and shape of the posterior genital lobe between *D. simulans* and *D. mauritiana*[26,27] all clearly support Robertson's[1] model (FIG. 4a).

How do QTL effects vary between males and females, in different environments and in different genetic backgrounds? A recurring theme from studies of *Drosophila* QTL is that genotype-by-sex (GSI), genotype-by-environment (GEI) and epistatic interactions are common and sometimes complicated. *Drosophila* QTL are often sex-[18–25,28] and environment-[19,24,25,29] specific (FIG. 4b). Some QTL have more complicated effects, and are both sex- and environment-specific[19,24,25]. Sex- and environment-specific QTL effects are usually attributable to conditional neutrality[19,23,28,29]. However, longevity QTL often show antagonistic pleiotropy[24,25]. Detecting epistatic interactions between QTL is more difficult than estimating effects of each QTL separately. The need to control the false positive error rate by adjusting for the number of tests (for n QTL there are $n(n-1)/2$ possible pairwise tests for interactions) effectively rules out genome scans for epistasis among all possible markers; even if interactions between significant QTL are explored, only very large epistatic effects can be detected. Nevertheless, epistatic interactions between QTL are often found. Significant interactions have been observed for bristle number[13,18,20], lifespan[25] and wing shape[22]. Epistatic interactions between QTL can be as large as the main effects[18,20,25], and can be sex-[18,25] and environment-[25] specific. Epistatic interactions between wing shape QTL were individually large, but contributed negligibly to the total phenotypic variance[22], because the interactions were balanced between positive and negative effects. By contrast, QTL associated with divergence of the size and shape of the posterior genital arch between *D. simulans* and *D. mauritiana* act predominately additively[27].

Extensive GSI, GEI and epistasis have practical and theoretical consequences. First, estimates of QTL positions and effects are relevant only to the sex and environment in which the phenotypes were assessed, and may not replicate across sexes and in different environments. Second, estimates of main QTL effects that assume no epistatic interactions between QTL will be

© 2001 Macmillan Magazines Ltd

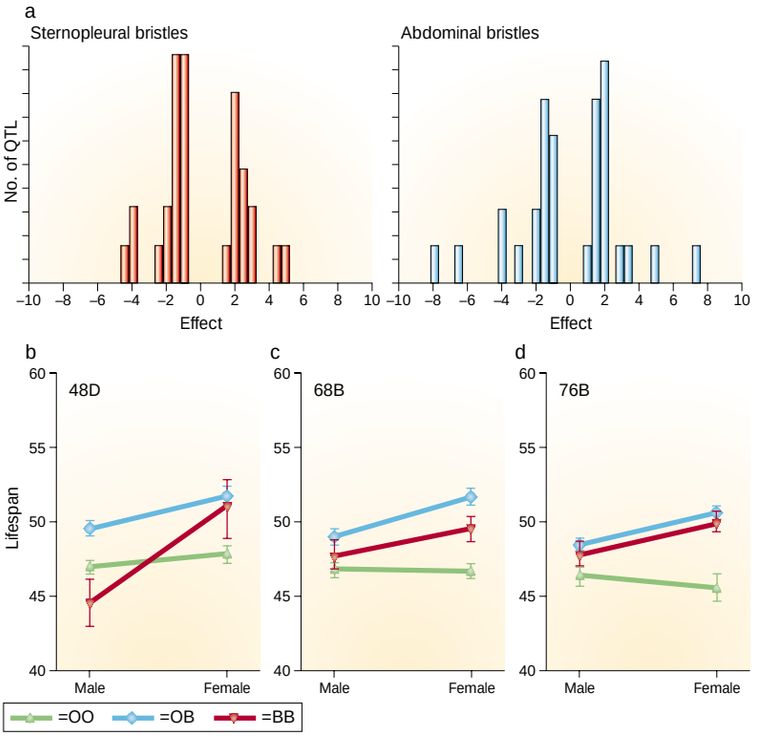


Figure 4 | **Quantitative trait locus effects. a** | The distribution of quantitative trait locus (QTL) effects is typically exponential, with a few QTL with large effects causing most of the difference between strains, and an increasingly larger number of QTL with smaller effects contributing to the remainder of the difference. (Data are from REF. 21.) **b** | Sex-specific effects of three lifespan QTL. The numbers and letters refer to the cytological position of the QTL. For each QTL, the mean lifespan of individuals homozygous for the Oregon (OO) or 2b (BB) strain QTL allele, and the heterozygous (OB) genotype, are shown, in males and females. The QTL at 48D shows antagonistic pleiotropy between the sexes, and seems OVERDOMINANT in males and strictly dominant in females. **c,d** | The QTL at 68B and 76B show conditional neutrality; they are not associated with significant variation in male longevity, but are associated with a difference in mean lifespan of females. (Data are from REF. 25.)

incorrect if there is epistasis. Third, GSI and GEI at the level of QTL are relevant to the puzzle of why there is extensive genetic variation for quantitative traits in natural populations, in spite of strong natural selection reducing variation[4]. Some fraction of genetic variation must be attributable to a balance between mutation and selection[30]. However, for many quantitative traits, the mutational variation seems to be too small to account for the levels of segregating variation observed[4], suggesting that additional mechanisms must be implicated. GEI or GSI caused by antagonistic pleiotropy can maintain variation in a heterogeneous environment (or the two sexes) in which directional selection favours alternative homozygous genotypes in the different environments or sexes[31]. GEI or GSI can also promote variation in heterogeneous environments in which the trait is under STABILIZING SELECTION, and the heterozygous genotype has an intermediate effect that is less sensitive to the environmental conditions than the more extreme homozygous genotypes[32]. Last, GEI can promote variation if the interaction is one of conditional neutrality, by reducing the overall selection acting at the locus[33].

OVERDOMINANT
Heterozygote superiority. The phenotype of the heterozygote is greater than that of either homozygotes. Overdominance for fitness can lead to the maintenance of both alleles in the population.

STABILIZING SELECTION
Intermediate phenotypes have greater fitness than extreme high and low scoring phenotypes.

The degree of difficulty in achieving the ultimate goal of describing variation for quantitative traits in terms of all QTL affecting variation in the trait — or more reasonably, the QTL associated with 70% of the variance in the trait[6] — depends on the extent to which the same QTL actually affect trait variation in different populations. Assuming there are many loci at which mutations affecting any quantitative trait can occur, this issue is tied in with the question of maintenance of variation for the trait. If, for example, variation is maintained largely by mutation–selection balance and the mutation rate is similar at all loci, or by a balance between mutation and drift at some neutral loci[34], one might expect a different constellation of loci affecting the trait in different populations. The same genomic regions that contain QTL for *Drosophila* sensory bristle number recur in studies that span a period of nearly 40 years, for populations of diverse origin, and using different mapping methodologies[13,19,21,35,36], indicating that many of the same loci cause variation in bristle number in different geographical populations. This inference is supported by classical quantitative genetic analyses[37,38]. By contrast, the apparent similarities in map positions could be due to the large number of QTL identified and their broad genomic localizations. The answer to this fundamental question awaits the identification of genetic loci corresponding to QTL.

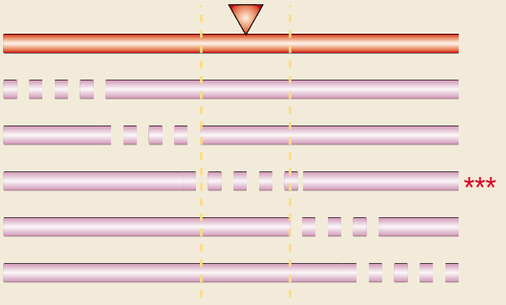
Fine-scale quantitative trait locus mapping

QTL, as defined in the above genome mapping studies, are not genetic loci, but chromosome regions that contain one or more loci affecting the trait. The size of the genomic region in which QTL are localized depends on the density of markers and the scale of the experiment, but are typically 3–10 cM. Furthermore, small recombination distances can correspond to large physical distances in regions of restricted recombination. It is important to remember that QTL mapping is an exercise in statistical model selection[17], and map positions and effects of significant QTL can vary according to the method of analysis[23,27,39]. The best fitting model, identifying the most QTL, is not necessarily the closest approximation to reality. So, the utility of QTL mapping for identifying human disease loci, as a functional genomics tool and for understanding evolutionary processes hinges on our ability to confirm the existence of QTL by an independent method that relies less on sophisticated statistical models, and to map QTL to the level of genetic loci[40,41]. The former is easier than the latter.

The classic method for confirming the existence of QTL is to introgress the putative QTL, one at a time, into a homozygous genetic background, by multiple generations of backcrossing. This method has been applied to three *Drosophila* bristle number QTL in the *achaete-scute* complex (*ASC*) — the *Notch* (*N*)[42], *scabrous* (*sca*)[43] and *Delta* (*Dl*)[44] gene regions. In all cases, the existence of QTL affecting bristle number in these regions was confirmed. However, the introgressed regions are expected to contain, on average, 20 cM of linked genome from the naturally derived chromosomes. These studies therefore confirm the existence of QTL but do not refine their map positions.

© 2001 Macmillan Magazines Ltd

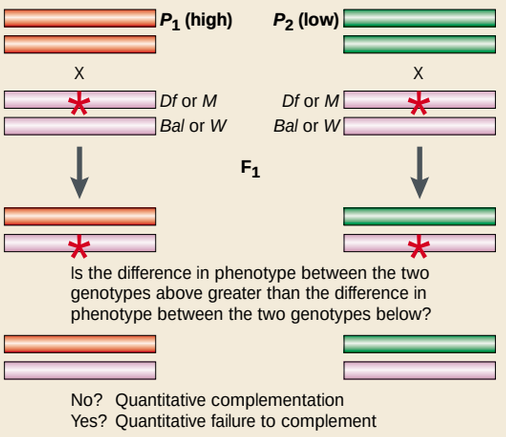
Box 2 | Deficiency and complementation mapping



Drosophila deficiency stocks (in pink) contain a chromosome missing a stretch of genome, the physical limits, or breakpoints, of which are known at the level of the cytological map (top figure). The breakpoints of the collection of deficiency stocks overlap; together they uncover 70–80% of the genome. To fine-map recessive, single gene mutations, the chromosome that contains the mutant allele is crossed to a set of deficiency stocks with overlapping breakpoints, and complementation (wild-type phenotype) or failure to complement (mutant phenotype) is recorded for the progeny that contain the deficiency chromosomes. The location of the mutation is then delineated by the region of non-overlap of adjacent deficiencies that complement the mutant phenotype with those that fail to complement the mutant (asterisks).

The principle of quantitative deficiency mapping is shown in the bottom figure. Homozygotes for the deficiency (*Df*) chromosome are not viable, so the stocks are maintained against a balancer (*Bal*) chromosome. The method requires that each of two parental strains (P_1, P_2) that contain different quantitative trait locus (QTL) alleles are crossed to the deficiency stock (*Df*/*Bal*) and the quantitative trait phenotype is evaluated for a number of F_1 individuals of each of the four resulting genotypes (*Df*/P_1, *Df*/P_2, *Bal*/P_1 and *Bal*/P_2). These data are analysed statistically to determine whether the difference in phenotype between the P_1 and P_2 strains is (quantitative complementation) or is not (quantitative failure to complement) the same in the *Df* and *Bal* chromosome backgrounds. As for all genetic complementation tests, there are two possible interpretations of quantitative failure to complement: the deficiency encompasses a QTL in the parental strains with different allelic effects on the trait, or there is epistasis between QTL in the parental strains with other QTL on the *Df* or *Bal* chromosome. To minimize the effects of epistatic failure to complement, one must impose the constraint that the difference between the parental strains is greater in the deficiency than in the balancer chromosome background, and, ideally, uncover the same genetic region with independent deficiencies.

The quantitative complementation test is logically analogous to quantitative deficiency complementation, and requires a mutant (*M*) and wild-type (*W*) allele at the candidate locus, and a minimum of two QTL alleles (Q_1 and Q_2). The strains that contain different QTL alleles are crossed to a strain that contains a mutant and a strain that contains the wild-type candidate gene allele, and the trait phenotype is measured in progeny of each of the four resulting genotypes. These data are analysed statistically to determine whether the difference in effect of the QTL alleles is (quantitative complementation) or is not (quantitative failure to complement) the same in the mutant and wild-type candidate gene background. As for quantitative deficiency complementation, failure to complement can be attributable to allelism or epistasis; thus, results of such tests are useful in nominating candidate genes for further study, but do not prove the QTL is allelic to the candidate gene.



The main problems with fine-scale QTL mapping are that very large numbers of recombination events are necessary to whittle away linked genomic regions to the level of genetic loci, and individual QTL effects are small. In *Drosophila*, fine-scale mapping of recessive mutations with large effects is usually accomplished using deficiency stocks, which does not require the production of recombinants in the general region to which the gene maps. Deficiency complementation mapping has been modified for fine-scale QTL mapping (BOX 2), as illustrated by a study[39] in which this method was used to refine the map positions of QTL that affected variation in lifespan between the Oregon (standard wild-type) and Russian 2b (derived from a line selected for low male mating activity) strains under standard culture conditions[23]. Deficiency mapping revealed that multiple linked factors contribute to each QTL detected by recombination mapping. One of the lifespan QTL mapped to the *Alcohol dehydrogenase* (*Adh*) gene region, which is the best annotated gene region of any higher eukaryote. The availability of many overlapping deficiency stocks uncovering this region, and the excellent integration of the cytogenetic, genetic and DNA sequence maps[45], allowed an ultra-fine scale dissection of longevity QTL in this region. A minimum of three tightly linked QTL were detected, and mapped to 50-, 185- and 200-kb regions, respectively. These three QTL together contain only six genetically defined loci and 24 confirmed and predicted genes that are positional candidates corresponding to lifespan QTL. Furthermore, two of these QTL are associated with quantitative variation in lifespan in a sample of lines recently derived from nature. So, the pessimism about the prospects of going from QTL to gene in the mouse[41] may not be valid for *Drosophila*.

From quantitative trait locus to gene

All genes and predicted genes in the region to which a QTL maps are candidate genes that could correspond to the QTL. A systematic method is needed to test which genetic locus or loci contribute to the QTL effect. One genetic method is a quantitative test for complementation of mutations at candidate genes that are defined genetically (that is, for which mutant stocks are available) with the QTL alleles.

The logic of the quantitative complementation test is the same as that of quantitative deficiency mapping, and the interpretation of the results is subject to the same caveats[20,46,47] (BOX 2). In principle, it is possible to do these tests for all genetically defined loci in the QTL region. In practice, it is prudent to prioritize candidate genes for testing on the basis of our understanding of which genes affect the development, physiology, and expression of the quantitative trait, if possible. For example, *Drosophila* bristles are external mechanosensory organs of the peripheral nervous system (PNS). Many loci affecting PNS development have been characterized genetically and at the molecular level, and placed in a developmental pathway[48,49]. Nearly 80% of bristle number QTL map to regions containing one or more candidate genes known to affect PNS development or to regions identified by mutations with adult bristle number phenotype[21].

© 2001 Macmillan Magazines Ltd

Indeed, QTL alleles on chromosomes selected for high or low bristle number fail to complement mutations at such candidate genes[20,46].

Some QTL map to regions where there are no obvious candidate genes[21], and for traits such as lifespan, one can implicate almost any locus as a candidate gene. For such regions, there is no option but to test all possible genetically defined loci. Ultimately, this must be done even for those regions containing obvious candidates, because some QTL may correspond to loci with undescribed and unexpected pleiotropic effects on the trait. At present, the main roadblock on the journey from QTL to gene is that tests for quantitative complementation can only be done if a stock containing a mutant allele at the candidate gene exists. Much of the *Drosophila* genome consists of loci with known function but no mutant allele, genetically defined loci that have not been mapped to the sequence and predicted loci with no known function or alleles. So, our quest to understand the genetic architecture of quantitative traits must include large-scale mutagenesis.

Mutagenesis

Single *P* element transposon mutagenesis has been the workhorse of the Berkeley *Drosophila* Genome Gene Disruption Project[50] in which over 1,000 unique *P* element insertions in genes that are essential for adult viability have been described. These methods can be extended to the analysis of quantitative effects of *P* element insertions by ensuring that the *P* element insertions are derived in a common, isogenic background, and by assessing the phenotypic effects of the insertions on a quantitative measurement scale, on multiple individuals from each insertion line.

Analysis of quantitative effects of single *P* element insertion lines is an efficient method for recovering mutations with hitherto unknown effects on quantitative traits. Over half of the single insertion lines screened for activities of enzymes involved in intermediary metabolism were significantly variant for at least one of the traits scored[51]. Although insertion sites were not determined in this study, statistical arguments suggest that the insertions were highly unlikely to be in enzyme-coding loci. Highly significant mutational variance and sex-specific mutational variance among single *P*-element insertion lines were observed for sensory bristle number[52] and olfactory behaviour[53]. Of the 50 insertion lines with significant effects on bristle number, nine were hypomorphic mutations at loci known to affect nervous system development, whereas the remaining 41 insertions did not map to cytogenetic regions containing loci with previously described effects on adult bristle number[52]. Most of the mutational variance for olfactory behaviour was attributable to *P* element insertions in 14 novel *smell-impaired* (*smi*) loci[53].

A further advantage of a collection of single insertional mutations in a co-isogenic background is that they can be used to determine how inter-locus epistatic interactions shape the phenotype. Quantitative tests for epistasis have been conducted for *P* element insertions affecting odour-guided behaviour[54], by crossing *smi* insertion lines in all possible pairwise combinations, and assessing olfactory behaviour on the doubly heterozygous progeny. The test for epistasis was to determine if the olfactory behaviour of the double heterozygotes could be predicted from the heterozygous effects of each mutant separately. If so, the loci act independently; if not, they show epistasis. Of the 12 *smi* loci tested, eight formed an interacting network, and two interacted with each other, independently of the others.

A different design was used to examine epistasis for metabolic activity[55]. All nine two-locus genotypes were constructed for pairs of *P* element insertion lines selected at random relative to the initial screen for mutational effects on metabolism, and measured for the same battery of traits as the initial study[51]. Significant and large epistatic interactions were found in 27% of the tests. These studies indicate that epistatic interactions for quantitative traits are common among *P* element insertion lines, and that these interactions are not confined to those between mutations that themselves have significant effects on the trait. With large numbers of insertion lines, however, tests for epistasis quickly become tedious, requiring the measurement of quantitative trait phenotypes on the order of n^2 genotypes, where n is the number of insertion lines. This problem may be circumvented in the future by simultaneous analysis of genome-wide changes in expression at all loci in response to single *P* element mutations that affect a common trait, using expression microarrays.

Linkage disequilibrium mapping

In our quest for genetic loci corresponding to QTL, we will ultimately arrive at a point where the QTL is mapped to a small chromosomal region, in which all genes and predicted genes are known. How can we infer which of these loci is associated with the QTL? Having determined the genetic locus corresponding to the QTL, how do we determine what molecular polymorphisms — single nucleotide polymorphisms (SNPs), small insertions/deletions or large insertions — define a functional QTL allele? Perhaps surprisingly, the population genetic concept of linkage disequilibrium can be used to answer both questions. Linkage disequilibrium is a measure of the correlation of allele frequencies at two polymorphic loci[4] (FIG. 5). Population genetic theory[56,57] predicts that a molecular marker locus will be in linkage disequilibrium with a molecular variant affecting the quantitative trait (the quantitative trait nucleotide, or QTN) only if they are very tightly linked — provided the population demography satisfies the theoretical assumptions (FIG. 5). In principle, linkage disequilibrium mapping could be applied as a genome screen paradigm, dispensing with traditional QTL mapping altogether, as has been proposed for human disease genes and other medically important traits[58]. If the population demography does not conform to the assumptions of the population genetic model, however, significant linkage disequilibrium does not necessarily imply close physical linkage between molecular markers and QTNs. For example, admixture (inter-population gene flow) between populations that have different gene frequen-

© 2001 Macmillan Magazines Ltd

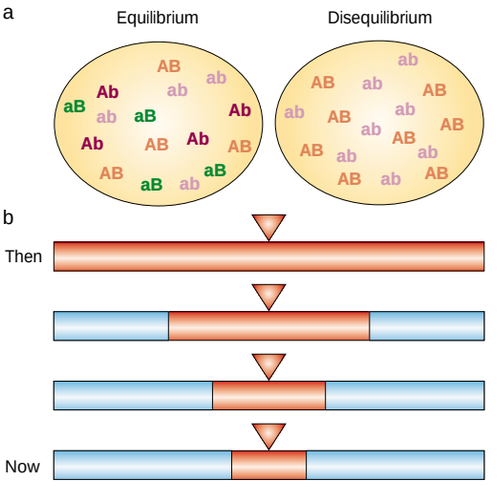


Figure 5 | **Linkage disequilibrium mapping. a** | Consider two loci (*A* and *B*) each with two alleles (*A*, *a*, *B*, *b*) and allele frequencies of p_1, p_2 at locus *A* and q_1, q_2 at locus *B*. These loci are said to be 'in linkage equilibrium' in a population if the observed frequencies of the two-locus gamete types — or 'haplotypes' (*AB*, *Ab*, *aB* and *ab*) — do not deviate significantly from those expected from the product of the frequencies of the constituent alleles (p_1q_1, p_1q_2, p_2q_1, p_2q_2, respectively). Otherwise, the alleles are nonrandomly associated, as would be the case if only *AB* and *ab* haplotypes were observed. **b** | Linkage disequilibrium mapping. When a new mutation occurs in a population at a locus affecting a quantitative trait, all other polymorphic alleles in that population will initially be in complete linkage disequilibrium with the mutation. Over time, however, recombination between the mutant allele and the other loci will create the missing haplotypes and restore linkage equilibrium between the mutant allele and all but closely linked loci. The length of the genomic fragment surrounding the original mutation in which linkage disequilibrium between the QTL and other loci still exists depends on the average amount of recombination per generation experienced by that region of the genome, the number of generations that have passed since the original mutation, and the population size, among other factors[56,57]. For old mutations in large equilibrium populations, strong linkage disequilibrium is only expected to extend over very short distances, of the order of kilobases or less.

cies at the marker and different values of the trait will cause spurious linkage disequilibrium, even between unlinked loci[4]. For this reason, association studies are optimally conducted on single populations.

Linkage disequilibrium mapping in *Drosophila*

So far, linkage disequilibrium mapping has been applied in *Drosophila* on the finer scale of mapping the molecular variants within single candidate gene regions that are associated with quantitative variation in enzyme activity[59–64] and sensory bristle number[42,43,65–67]. Samples of chromosomes containing the candidate gene of interest were collected from natural populations, made homozygous and substituted into the homozygous genetic background of an unrelated strain. All genetic variation for the trait was therefore attributable to loci only on the chromosome containing the candidate gene. The power to detect associations was increased by limiting the contribution of chromosomally unlinked loci to trait variation, and making replicate measurements of the trait on several individuals from each chromosome line.

For example, efforts to associate variation for sensory bristle number with molecular polymorphism at the candidate bristle genes *achaete-scute* (*ASC*)[42,65], *scabrous* (*sca*)[43,66] and *Delta* (*Dl*)[67], have been successful. Alleles of *ASC* can be divided into two classes: those that contain large insertions and those that do not . Two independent studies showed that alleles with large insertions had fewer bristles than those without[65]. Furthermore, the effects of large insertions as a class[65], and of one particular polymorphic large insertion[42], were associated with female, but not male, abdominal bristle number. A small deletion was also associated with sternopleural bristle number in both sexes[42]. Marker associations at *sca* and *Dl* were consistent with expectations based on the results of the quantitative complementation tests. Three markers at *sca* were associated with female abdominal bristle number, but none were significantly associated with sternopleural bristle number[43]. One marker at *Dl* was associated with female abdominal bristle number and one with sternopleural bristle number in both sexes[67]. All kinds of molecular variation (SNPs, small insertions/deletions and large insertions) can therefore be associated with quantitative variation in phenotypes. So far, all significant associations with bristle number have been for molecular polymorphisms in introns and non-coding flanking regions of the candidate genes.

Caveats and future prospects

The power to detect associations between molecular polymorphisms and QTNs depends on the density of molecular polymorphisms and the sample size[68]. No study has yet used the optimal marker density, which could be every 200 base pairs in regions of high recombination in *Drosophila*[67]. Furthermore, the association between linkage disequilibrium and physical distance breaks down over very short physical distances[60,62], owing to statistical sampling error and variation in evolutionary history of the polymorphic molecular markers[56,69–71]. Difficulties in achieving the ultimate goal of defining QTL alleles at the molecular level arise because an association between a polymorphic site and quantitative trait does not mean the polymorphism is the cause of the phenotypic difference, and because multiple molecular markers in strong linkage disequilibrium with each other may all be associated with variation in the quantitative trait. Different technologies must be brought to bear to determine the causal relationship between molecular and phenotypic variation.

One method that is now available for the finest possible scale dissection of molecular polymorphism–trait associations is to combine *in vitro* mutagenesis with *P*-element-mediated germline transformation to test the functional effects of each putative polymorphic site, both one at a time and in combination. This method was first proposed, and executed, for the case of molecular polymorphism in the *Adh* gene region and its association with ADH activity and protein concentration, with some surprising and sobering results[72–76]. Initially, 92% of the

© 2001 Macmillan Magazines Ltd

difference in enzyme activity and protein concentration between the Fast (F) and Slow (S) allozyme classes was shown to map to a 2.3-kb genomic fragment containing all *Adh* coding sequences and the 3′ flanking region[72]. This region contains 13 polymorphic sites that distinguish between consensus F and S alleles, one of which is the amino-acid substitution causing the allozyme difference. This polymorphism is responsible for the difference in activity between the F and S alleles, but is not associated with the difference in protein concentration[73]. Efforts to map the sites responsible for the difference in protein concentration showed that a complex insertion/deletion polymorphism in the first intron accounted for about one-third of the difference in concentration between the allozyme classes[75]. The 2.3-kb region was then split into three fragments. All three fragments contributed significantly to the difference in protein concentration, with epistatic interactions between two fragments[76]. So, a minimum of three polymorphic sites in one 2.3-kb genomic region contribute to the naturally occurring variation in the concentration of ADH protein. In future, advances in homologous recombination technology[8] will enable substitution of engineered transgenes at the exact site of the endogenous allele, considerably reducing the number of independent transformant lines required to account for quantitative effects of random *P* element insertions.

Lessons from *Drosophila*

The *Drosophila* model system is relevant to the problem of genetic dissection of complex human diseases from two perspectives. First, there is direct homology between *Drosophila* genes and genes that affect human disease. Of all the genes known to affect human disease, 61% have *Drosophila* orthologues (see Homophila web site); and around half of all *Drosophila* protein sequences are similar to those of mammals[77]. It is therefore conceivable that the loci that affect variation in complex traits in *Drosophila* (for example, longevity) have human homologues. Identifying these loci in flies will suggest candidate genes to study in mammalian systems, where the process of identifying genetic loci that correspond to QTL is even more difficult than in *Drosophila*[41].

Second, lessons learned from studies in *Drosophila* should provide guidance as to experimental design for similar studies in other systems. *Drosophila* quantitative and population genetics has a rich theoretical and empirical history that has been largely overlooked in QTL mapping studies in mammals and humans, and more recently, in proposals for using linkage disequilibrium to map complex human diseases. Obvious examples where cross-fertilization from *Drosophila* could have motivated and expedited the development of methods for the genetic analysis of quantitative traits in other systems are the development of interval mapping combined with progeny testing[12,14] and the utility of chromosome substitution lines[10,11] for mapping QTL.

If key biological processes are conserved at the genetic level from *Drosophila* to humans, it is not unreasonable to assume that general properties of QTL are also conserved across taxa. If *Drosophila* QTL have variable effects depending on the sex, physical environment and genetic environment in which the QTL are expressed, similar properties are to be expected of QTL in other organisms. If molecular variation in introns and noncoding regions of *Drosophila* candidate genes is associated with quantitative variation in phenotypes, these regions should be examined for association with quantitative variation whenever such studies are done. And if multiple, linked molecular polymorphisms within *Drosophila* candidate genes are associated (sometimes nonlinearly) with quantitative variation in phenotypes, such complexity ought to be generally expected.

These QTL properties have practical implications for linkage disequilibrium mapping of genes that affect human complex traits. The lack of control over genotype and environment in natural populations, in combination with the likelihood that genetic and environmental contexts are important for the expression of QTL alleles, mean that the effect of any one QTL allele, expressed as a fraction of the phenotypic standard deviation of the trait, will be very small indeed. Operationally, this translates into a need for very large sample sizes to detect individual QTL alleles. Furthermore, the likelihood that multiple molecular polymorphisms within candidate genes affect variation in complex traits[78,79] suggests that an even higher density of SNP markers for association tests will be required than suggested by the most pessimistic current estimates[57]. Such challenges underscore the utility of model organisms for genetic dissection of medically important quantitative traits.

Links

DATABASE LINKS *ASC* | *N* | *sca* | *Dl* | *Adh* | *smi* loci
FURTHER INFORMATION Berkeley *Drosophila* Genome Gene Disruption Project | Homophila
ENCYCLOPEDIA OF LIFE SCIENCES Quantitative genetics

1. Goffeau, A. *et al.* Life with 6,000 genes. *Science* **274**, 546–567 (1996).
2. The *C. elegans* sequencing consortium. Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* **282**, 2012–2018 (1998).
3. Adams, M. D. *et al.* The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195 (2000).
4. Falconer, D. S. & Mackay, T. F. C. *Introduction to Quantitative Genetics* (Addison Wesley Longman, Harlow, 1996).
5. Lynch, M. & Walsh, B. *Genetics and Analysis of Quantitative Traits* (Sinauer Associates, Sunderland, Massachusetts, 1998).
6. Robertson, A. in *Heritage From Mendel* (ed. Brink, A.) 265–280 (Univ. Wisconsin, Madison, Wisconsin, 1967).
The terms in which we must seek to describe genetic variation in quantitative traits are clearly set out in this article, which remain as valid today as they were over thirty years ago.
7. Ashburner, M. Drosophila, *A Laboratory Handbook* (Cold Spring Harbor Laboratory Press, New York, 1989).
8. Rong, Y. & Golic, K. G. Gene targeting by homologous recombination in *Drosophila*. *Science* **288**, 2013–2018 (2000).
9. Sax, K. The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics* **8**, 552–560 (1923).
This was the first attempt to map QTL by association with Mendelian markers.
10. Breese. E. L. & Mather, K. The organization of polygenic activity within a chromosome of *Drosophila*. I. Hair characters. *Heredity* **11**, 373–395 (1957).
11. Nadeau, J. H., Singer, J. B., Matin, A. & Lander, E. S. Analysing complex genetic traits with chromosome substitution strains. *Nature Genet.* **24**, 221–225 (2000).
12. Thoday, J. M. Location of polygenes. *Nature* **191**, 368–370 (1961).
The first paper to describe interval mapping in the context of QTL mapping, and summarize its

© 2001 Macmillan Magazines Ltd

application to mapping QTL for *Drosophila* bristle number.

13. Shrimpton, A. E. & Robertson, A. The isolation of polygenic factors controlling bristle score in *Drosophila melanogaster*. II. Distribution of third chromosome bristle effects within chromosome sections. *Genetics* **118**, 445–459 (1988).
14. Lander, E. S. & Botstein, D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**, 185–199 (1989).
A landmark paper, proposing the use of interval mapping with molecular markers to localize QTL in species that do not contain multiple morphological markers, and developing a maximum likelihood statistical framework to analyse such data.
15. Zeng, Z. -B. Precision mapping of quantitative trait loci. *Genetics* **136**, 1457–1468 (1994).
16. Jiang, C. & Zeng, Z. -B. Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* **140**, 1111–1127 (1995).
17. Kao, C. -H., Zeng, Z. -B. & Teasdale, R. Multiple interval mapping for quantitative trait loci. *Genetics* **152**, 1203–1216 (1999).
References 15–17 refine the precision of QTL mapping by including marker cofactors to account for multiple QTL, and extend QTL mapping analysis to include multiple traits and epistasis.
18. Long, A. D. *et al.* High resolution mapping of genetic factors affecting abdominal bristle number in *Drosophila melanogaster*. *Genetics* **139**, 1273–1291 (1995).
19. Gurganus, M. C. *et al.* Genotype-environment interaction at quantitative trait loci affecting sensory bristle number in *Drosophila melanogaster*. *Genetics* **149**, 1883–1898 (1998).
20. Gurganus, M. C., Nuzhdin, S. V., Leips, J. W., & Mackay, T. F. C. High-resolution mapping of quantitative trait loci for sternopleural bristle number in *Drosophila melanogaster*. *Genetics* **152**, 1585–1604 (1999).
21. Nuzhdin, S. V., Dilda, C. L. & Mackay, T. F. C. The genetic architecture of selection response: inferences from fine-scale mapping of bristle number quantitative trait loci in *Drosophila melanogaster*. *Genetics* **153**, 1317–1331 (1999).
22. Weber, K. *et al.* An analysis of polygenes affecting wing shape on chromosome 3 in *Drosophila melanogaster*. *Genetics* **153**, 773–786 (1999).
23. Nuzhdin, S. V., Pasyukova, E. G., Dilda, C. L., Zeng, Z. -B. & Mackay, T. F. C. Sex-specific quantitative trait loci affecting longevity in *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **94**, 9734–9739 (1997).
24. Vieira, C. *et al.* Genotype-environment interaction for quantitative trait loci affecting lifespan in *Drosophila melanogaster*. *Genetics* **154**, 213–227 (2000).
References 23 and 24 illustrate that QTL effects can depend on the sex, physical environment and genetic environment in which they are estimated.
25. Leips, J. & Mackay, T. F. C. Quantitative trait loci for lifespan in *Drosophila melanogaster*: interactions with genetic background and larval density. *Genetics* **155**, 1773–1788 (2000).
26. Liu, J. *et al.* Genetic analysis of a morphological shape difference in the male genitalia of *Drosophila simulans* and *D. mauritiana*. *Genetics* **142**, 1129–1145 (1996).
27. Zeng, Z.-B. *et al.* Genetic architecture of a morphological shape difference between two Drosophila species. *Genetics* **154**, 299–310 (2000).
28. Wayne, M. L. *et al.* Quantitative trait loci for fitness–related traits in *Drosophila melanogaster*. *Genet. Res*. (in the press).
29. Fry, J. D., Nuzhdin, S. V., Pasyukova, E. G. & Mackay, T. F. C. QTL mapping of genotype-environment interaction for fitness in *Drosophila melanogaster*. *Genet. Res.* **71**, 133–141 (1998).
30. Houle, D., Morikawa, B. & Lynch, M. Comparing mutational variabilities. *Genetics* **143**, 1467–1483 (1996).
31. Levene, H. Genetic equilibrium when more than one ecological niche is available. *Am. Nature* **87**, 331–333 (1953).
32. Gillespie, J. H. & Turelli, M. Genotype-environment interactions and the maintenance of polygenic variation. Genetics **121**, 129–138 (1989).
33. Fry, J. D., Heinsohn, S. L. & Mackay, T. F. C. The contribution of new mutations to genotype-environment interaction for fitness in *Drosophila melanogaster*. *Evolution* **50**, 2316–2327 (1996).
34. Lynch, M. & Hill, W. G. Phenotypic evolution by neutral mutation. *Evolution* **40**, 915–935 (1986).
35. Wolstenholme, D. R. & Thoday, J. M. Effects of disruptive selection. VII. A third chromosome polymorphism *Heredity* **10**, 413–431 (1963).
36. Spickett, S. G. & Thoday, J. M. Regular responses to selection. 3. Interactions between located polygenes. *Genet. Res.* **7**, 96–121 (1966).
37. López-Fanjul, C. & Hill, W. G. Genetic differences between populations of *Drosophila melanogaster* for a quantitative trait. Laboratory populations. *Genet. Res*. **22**, 51–68 (1973).
38. López-Fanjul, C. & Hill, W. G. Genetic differences between populations of *Drosophila melanogaster* for a quantitative trait. Wild and laboratory populations. *Genet. Res.* **22**, 69–78 (1973).
39. Pasyukova, E. G., Vieira, C. & Mackay, T. F. Deficiency mapping of quantitative trait loci affecting longevity in *Drosophila melanogaster*. *Genetics* **156**, 1129–1146 (2000).
The first attempt to fine-map QTL using deficiency complementation.
40. Darvasi, A. Experimental strategies for the genetic dissection of complex traits in animal models. *Nature Genet.* **18**, 19–24 (1998).
41. Nadeau, J. H. & Frankel, D. The roads from phenotypic variation to gene discovery: mutagenesis versus QTLs. *Nature Genet.* **25**, 381–384 (2000).
42. Long, A. D., Lyman, R. F., Morgan, A. H., Langley, C. H. & Mackay, T. F. C. Both naturally occurring insertions of transposable elements and intermediate frequency polymorphisms at the *achaete-scute* complex are associated with variation in bristle number in *Drosophila melanogaster*. *Genetics* **154**, 1255–1269 (2000).
43. Lyman, R. F., Lai, C. & Mackay, T. F. C. Linkage disequilibrium mapping of molecular polymorphisms at the *scabrous* locus associated with naturally occurring variation in bristle number in *Drosophila melanogaster*. *Genet. Res.* **74**, 303–311 (1999).
44. Lyman, R. F. & Mackay, T. F. C. Candidate quantitative trait loci and naturally occurring phenotypic variation for bristle number in *Drosophila melanogaster*: The *Delta-Hairless* gene region. *Genetics* **149**, 983–998 (1998).
45. Ashburner, M. *et al.* An exploration of the sequence of a 2.9-Mb region of the genome of *Drosophila melanogaster*: the *Adh* region. *Genetics* **153**, 179–219 (1999).
46. Long, A. D., Mullaney, S. L., Mackay, T. F. C., & Langley, C. H. Genetic interactions between naturally occurring alleles at quantitative trait loci and mutant alleles at candidate loci affecting bristle number in *Drosophila melanogaster*. *Genetics* **14**, 1497–1510 (1996).
47. Mackay, T. F. C. & Fry, J. D. Polygenic mutation in *Drosophila melanogaster*: Genetic interactions between selection lines and candidate quantitative trait loci. *Genetics* **144**, 671–688 (1996).
48. Campos-Ortega, J. A. in *The Development of* Drosophila melanogaster Vol. 2 (eds Bate, M. & Martinez Arias, A.) 1091–1129 (Cold Spring Harbor Laboratory, Plainview, New York, 1993).
49. Jan, Y. N. & Jan, L. Y. in *The Development of* Drosophila melanogaster Vol. 2 (eds Bate, M. & Martinez Arias, A.) 1207–1244 (Cold Spring Harbor Laboratory, Plainview, New York, 1993).
50. Spradling, A. C. *et al.* The Berkeley *Drosophila* genome gene disruption project: Single *P* element insertions mutating 25% of vital Drosophila genes. *Genetics* **153**, 135–177 (1999).
51. Clark, A. G., Wang, L & Hulleberg, T. *P*-element-induced variation in metabolic regulation in *Drosophila*. *Genetics* **139**, 337–348 (1995).
52. Lyman, R. F., Lawrence, F. Nuzhdin, S. V. & Mackay, T. F. C. Effects of single *P* element insertions on bristle number and viability in *Drosophila melanogaster*. *Genetics* **143**, 277–292 (1996).
53. Anholt, R. R. H., Lyman, R. F. & Mackay, T. F. C. Effects of single *P* element insertions on olfactory behavior in *Drosophila melanogaster*. *Genetics* **143**, 293–301 (1996).
An example of the utility of screening for quantitative effects of insertional mutations to identify new loci that affect quantitative phenotypes.
54. Fedorowicz, G. M., Fry, J. D., Anholt, R. R. H. & Mackay, T. F. C. Epistatic interactions between *smell-impaired* loci in *Drosophila melanogaster*. *Genetics* **148**, 1885–1891 (1998).
55. Clark, A. G. & Wang, L. Epistasis in measured genotypes: *Drosophila P* element insertions. *Genetics* **147**, 157–163 (1997).
56. Weir, B. S. *Genetic Data Analysis II* (Sinauer Associates, Sunderland, Massachusetts, 1996).
57. Kruglyak, L. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nature Genet.* **22**, 139–144 (1999).
58. Risch, N. & Merikangas, K. The future of genetic studies of complex human diseases. *Science* **273**, 1516–1517 (1996).
59. Aquadro, C. F., Desse, S., Bland, M., Langley, C. H. & Laurie-Ahlberg, C. C. Molecular population genetics of the *alcohol dehydrogenase* gene region of *Drosophila melanogaster*. *Genetics* **114**, 1165–1190 (1986).
60. Miyashita, N. & Langley, C. H. Molecular and phenotypic variation of the *white* locus region in *Drosophila melanogaster*. *Genetics* **120**, 199–212 (1988).
61. Langley, C. H. *et al.* Naturally occurring variation in the restriction map of the *Amy* region of *Drosophila melanogaster*. *Genetics* **119**, 619–629 (1988).
62. Miyashita, N. Molecular and phenotypic variation of the *Zw* locus region in *Drosophila melanogaster*. *Genetics* **125**, 407–419 (1990).
63. Aquadro, C. F., Jennings, R. M., Bland, M., Laurie, C. C. & Langley, C. H. Patterns of naturally occurring restriction map variation, Dopa decarboxylase activity variation and linkage disequilibrium in the *Ddc* gene region of *Drosophila melanogaster*. *Genetics* **132**, 443–452 (1992).
64. Game, A. Y. & Oakeshott, J. G. The association between restriction site polymorphism and enzyme activity variation for Esterase 6 in *Drosophila melanogaster*. *Genetics* **126**, 1021–1031 (1990).
65. Mackay, T. F. C. & Langley, C. H. Molecular and phenotypic variation in the *achaete-scute* region of *Drosophila melanogaster*. *Nature* **348**, 64–66 (1990).
66. Lai, C., Lyman, R. F., Long, A. D., Langley, C. H. & Mackay, T. F. C. Naturally occurring variation in bristle number and DNA polymorphisms at the *scabrous* locus of *Drosophila melanogaster*. *Science* **266**, 1697–1702 (1994).
67. Long, A. D., Lyman, R. F., Langley, C. H. & Mackay, T. F. C. Two sites in the *Delta* gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster*. *Genetics* **149**, 999–1017 (1998).
An example of the application of linkage disequilibrium mapping at a candidate gene that affects quantitative variation in sensory bristle number, in which limitations of this approach and future directions are discussed.
68. Long, A. D. & Langley, C. H. Power of association studies to detect the contribution of candidate genetic loci to complexly inherited phenotypes. *Genome Res*. **9**, 720–731 (1999).
69. Weir, B. S. & Hill, W. G. Nonuniform recombination within the human β-globin gene cluster. *Am. J. Hum. Genet*. **38**, 776–778 (1986).
70. Hill, W. G. & Weir, B. S. Variances and covariances of squared linkage disequilibria in finite populations. *Theor. Pop. Biol*. **33**, 54–78 (1988).
71. Hill, W. G. & Weir, B. S. Maximum likelihood estimation of gene location by linkage disequilibrium. *Am. J. Hum. Genet.* **54**, 705–714 (1994).
72. Laurie-Ahlberg, C. C. & Stam, L. F. Use of *P*-element-mediated transformation to identify the molecular basis of naturally occurring variants affecting *Adh* expression in *Drosophila melanogaster*. *Genetics* **115**, 129–140 (1987).
73. Choudhary, M. & Laurie, C. C. Use of *in vitro* mutagenesis to analyse the molecular basis of the difference in *Adh* expression associated with the allozyme polymorphism in *Drosophila melanogaster*. *Genetics* **129**, 481–488 (1991).
74. Laurie, C. C., Bringham, J. T. & Choudhary, M. Associations between DNA sequence variation and variation in gene expression of the *Adh* gene in natural populations of *Drosophila melanogaster*. *Genetics* **129**, 489–499 (1991).
75. Laurie, C. C. & Stam, L. F. The effect of an intronic polymorphism on alcohol dehydrogenase expression in *Drosophila melanogaster*. *Genetics* **138**, 379–385 (1994).
76. Stam, L. F. & Laurie, C. C. Molecular dissection of a major gene effect on a quantitative trait: the level of alcohol dehydrogenase expression in *Drosophila melanogaster*. *Genetics* **144**, 1559–1564 (1996).
A summary of several experiments combining *in vitro* mutagenesis with *P* element transformation to determine the causal relationship between naturally occurring DNA sequence variation and quantitative variation in enzyme activity and protein stability. Evidence for multiple, interacting polymorphic sites within a single locus is presented.
77. Rubin, G. M. *et al.* Comparative genomics of the eukaryotes. *Science* **287**, 2204–2215 (2000).
78. Sing, C. F. & Moll, P. P. Genetics of atherosclerosis. *Ann. Rev. Genet.* **24**, 171–187 (1990).
79. Nickerson, D. A. *et al.* DNA sequence diversity in a 9.7-kb region of the human lipoprotein lipase gene. *Nature Genet.* **19**, 233–240 (1998).
80. Churchill, G. A. & Doerge, R. W. Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963–971 (1994).
81. Doerge, R. W. & Churchill, G. A. Permutation tests for multiple loci affecting a quantitative character. *Genetics* **142**, 285–294 (1996).
References 80 and 81 elaborate the use of permutation testing to set the appropriate threshold above which to declare significant QTL, given multiple tests and correlated markers.

Acknowledgements

Work in the author's laboratory is supported by the National Institute of Health and by the W. M. Keck Center for Behavioral Biology of North Carolina State University.

© 2001 Macmillan Magazines Ltd