nature genetics

Genome-wide genetic association of complex traits in heterogeneous stock mice

William Valdar[1], Leah C Solberg[1,4], Dominique Gauguier[1], Stephanie Burnett[1], Paul Klenerman[2], William O Cookson[1], Martin S Taylor[1], J Nicholas P Rawlins[3], Richard Mott[1] & Jonathan Flint[1]

Difficulties in fine-mapping quantitative trait loci (QTLs) are a major impediment to progress in the molecular dissection of complex traits in mice. Here we show that genome-wide high-resolution mapping of multiple phenotypes can be achieved using a stock of genetically heterogeneous mice. We developed a conservative and robust bootstrap analysis to map 843 QTLs with an average 95% confidence interval of 2.8 Mb. The QTLs contribute to variation in 97 traits, including models of human disease (asthma, type 2 diabetes mellitus, obesity and anxiety) as well as immunological, biochemical and hematological phenotypes. The genetic architecture of almost all phenotypes was complex, with many loci each contributing a small proportion to the total variance. Our data set, freely available at http://gscan.well.ox.ac.uk, provides an entry point to the functional characterization of genes involved in many complex traits.

The mouse is a key model organism for understanding gene function in mammals, yet many mouse phenotypes of interest to biomedical research have poorly understood and complex, polygenic origins. Despite new genomic resources such as access to dense maps of sequence variation and the ability to interrogate the expression levels of virtually every gene, molecular dissection of the loci that contribute to quantitative variation remains a challenge[1]. A central problem that impedes the cloning of QTLs is the difficulty of resolving genetic effects into sufficiently small intervals to make gene identification possible.

Successful strategies for high-resolution mapping ideally should be able to identify small genetic effects for any phenotype across the entire mouse genome. The classical approach begins by genetic mapping in a cross between two inbred strains or, more recently, in chromosome substitution strains[2], and it eventually results in the identification of a small number of independently segregating loci mapped into intervals larger than 20 Mb. Subsequent fine-mapping typically proceeds by repeatedly backcrossing one inbred strain onto another to isolate each locus. Such attempts are frequently frustrated when it is discovered that a single QTL segregating in inbred crosses fractionates into multiple smaller effects, each of which typically contributes less than 5% to the total phenotypic variance[1].

We have developed alternative methods for fine-mapping small-effect QTLs that use outbred mice of known ancestry[3–5]. By exploiting historical recombinants that have accumulated in a genetically heterogeneous stock of mice descended from eight inbred progenitor strains (A/J, AKR/J, BALBc/J, CBA/J, C3H/HeJ, C57BL/6J, DBA/2J and LP/J)[6], we have shown that QTLs explaining 5% or less of the phenotypic variation can be mapped into intervals of <1 cM. Although the heterogeneous stock has been used thus far for fine-mapping single-QTL intervals[4,7,8], it is theoretically ideal for high-resolution mapping of multiple QTLs across the genome: its derivation from multiple founders means it should contain more QTLs than any inbred cross, and the use of pseudorandom breeding for over 50 generations (in the case of the stock discussed in this paper) should result in an average distance between recombinants of <2 cM.

However, a number of potential obstacles must be tackled before the heterogeneous stock becomes a tool for genome-wide QTL mapping. First, the high costs of performing whole-genome association in the heterogeneous stock may preclude its use, because $\sim$100 times more markers and ten times more animals are required compared with an inbred strain cross[1]. Second, random fluctuations in allele frequencies and unrecognized selective pressures operating during the production and maintenance of the stock could reduce its heterozygosity, with consequent reductions in QTL resolving power. Third, a whole-genome analysis in the heterogeneous stock, which involves the simultaneous identification of multiple QTLs, poses unknown analytical problems that could seriously vitiate the outcome. The extensive repertoire of methods developed for whole-genome analysis of a classical intercross[9–12] are not directly applicable: in a heterogeneous stock, more loci are tested than individuals, so it is not possible to use methods that fit all markers simultaneously, and many more parameters are estimated at each locus than in a classical cross[4].

In this paper we demonstrate the utility of the heterogeneous stock for high-resolution whole-genome association analyses of quantitative

[1]Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK. [2]Peter Medawar Building for Pathogen Research, Nuffield Department of Medicine, University of Oxford, Oxford OX1 3SY, UK. [3]Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK. [4]Current address: Medical College of Wisconsin, Human and Molecular Genetics Center (HMGC), 8701 Watertown Plank Road, Milwaukee, Wisconsin 53226, USA. Correspondence should be addressed to J.F. (jf@well.ox.ac.uk) or R.M. (rmott@well.ox.ac.uk).

Received 18 April; accepted 13 June; published online 9 July 2006; doi:10.1038/ng1840

Table 1 Summary of phenotypes analyzed

Phenotype	Description	No. of animals	Mean age (d)
Weight, 6 weeks	Body weight at the beginning of testing	2,516	42
Immunology	CD4, CD3, CD8 and B220 antibody staining	1,872	42
Open field test	Model of anxiety. Measures distance in the perimeter, the center and total distance traveled in 5 min	2,504	45
Elevated plus maze (EPM)	Model of anxiety. Measures distance traveled, time spent and entries into closed and open arms	2,452	46
Food hyponeophagia (FN)	Model of anxiety. Measures time taken to sample a novel foodstuff after overnight food deprivation	2,474	47
Burrowing	Number of pellets removed from burrow in 1.5 h	2,455	48
Activity	Activity measured in a home cage in 30 min	2,445	48
Startle	Startle to a loud noise	1,948	52
Context freezing	Freezing to the context in which a tone is associated with a footshock	2,070	55
Cue freezing	Freezing to a tone after association with a foot shock	2,110	56
Plethysmography	Model of asthma. Animals sensitized by injection with ovalbumin inhale metacholine and changes in lung function are measured by plethysmography. Respiratory rate, tidal volume, minute volume, expiratory time, inspiratory time and enhanced pause are recorded with and without exposure to metacholine.	2,304	63
IPGTT	Model of type 2 diabetes mellitus. Glucose and insulin values taken at 0, 15, 30 and 75 min after intraperitoneal injection of glucose.	2,334	68
Weight, 10 weeks	Body weight at the end of testing	2,319	70
Full blood count	Hematocrit, hemoglobin concentration, mean cellular volume, mean cellular hemoglobin concentration, white cell count, platelet count	1,892	71
Tissue harvest	Adrenal weight	2,309	71
Wound healing	Reduction in size of a 2-mm ear punch hole	2,273	71
Biochemistry	Albumin, alkaline phosphatase, alanine transaminase, aspartate transaminase, calcium, chloride, creatinine, high density lipoprotein, low density lipoprotein, phosphorous, sodium, total cholesterol, total protein, triglycerides	1,890	71

traits. We have already shown that it is possible to overcome the first problem by measuring many phenotypes in parallel on each mouse, thereby reducing the cost of genotyping relative to the information gained per phenotype[13]. Our protocol for high-throughput phenotyping includes models of common diseases (anxiety, type 2 diabetes mellitus, obesity and asthma) as well as biochemistry, hematology and immunology profiles. We now deal with the two remaining issues: we show that the genetic structure of the heterogeneous stock meets the requirements of high-resolution mapping, and we present methods suitable for whole-genome QTL analysis in a heterogeneous stock. The QTL data set we have generated provides a starting point for gene identification in complex traits, and its analysis has implications for whole-genome association studies in humans.

RESULTS

Phenotypes and genotypes

We applied a high-throughput phenotyping protocol to collect data on heterogeneous stock mice as described[13] (available online at http://gscan.well.ox.ac.uk/ and summarized in **Table 1**). We obtained genotypes for 13,459 SNPs on 1,904 fully phenotyped mice and 298 parents with an accuracy of >99.9%. The heterogeneous stock we tested comprises a complex pedigree: over the 2-year course of the project, four generations of related animals, consisting of 85 families of average size 13.9 (s.d. 9.1), passed through the testing protocol. Because of this, linkage disequilibrium (LD) between pairs of markers is likely to be complex, with important consequences for mapping resolution and QTL detection. Mapping resolution depends on the extent of local LD, whereas QTL detection is affected by correlation between markers on different chromosomes.

We investigated the extent of LD between pairs of markers using R^2 as a measure of correlation. Within chromosomes, we found that for all the autosomes, mean R^2 fell to <0.5 within 2 Mb and is less than 0.2 within 8 Mb; 92% of marker pairs that are more than 8 Mb apart have $R^2 < 5\%$, indicating that the heterogeneous stock will deliver high-resolution mapping, although, as expected, LD extends over much larger physical distances on the X chromosome.

When we examined LD between markers on different chromosomes, we found that although 99.2% of such marker pairs have $R^2 < 5\%$, χ^2 tests of genotype association indicated the presence of statistically significant correlations between chromosomes. Under the assumption that pairs of markers on different autosomes are independent, we would expect to observe only one pair with a χ^2 logP (negative $\log_{10}$ of the P value) > 7.7, given the number of analyses performed. Despite the apparently low LD suggested by the R^2 analysis, we found that 17% of pairs exceeded this value. As family structure was ignored in the LD calculations, the R^2 and significance estimates are inflated, but this is justified because we are considering here the effects of LD on subsequent QTL analysis in the mapping population. The correlation between markers on different chromosomes complicates the detection of QTLs, as described below.

Genome scans

We performed genome scans for all 101 phenotypes. We used a multipoint probabilistic reconstruction of the heterogeneous stock haplotype mosaic[4] and a genetic map we built based on the three-generation heterogeneous stock pedigree (unpublished data). Consistent with previous results from heterogeneous stock analyses, we found that single-marker association generally detected far fewer significant loci[4,7,8], so we present results for the multipoint analyses only.

We used permutation to estimate significance thresholds for a model in which no QTL segregates. The 95% point of the distribution of the most significant peak in the permuted genome scans was used as a genome-wide threshold, or E value, equivalent to one false positive per 20 genome scans, or approximately five false-positive QTLs for the analysis of all phenotypes. The threshold varies only slightly between phenotypes: the mean threshold at $E < 0.05$

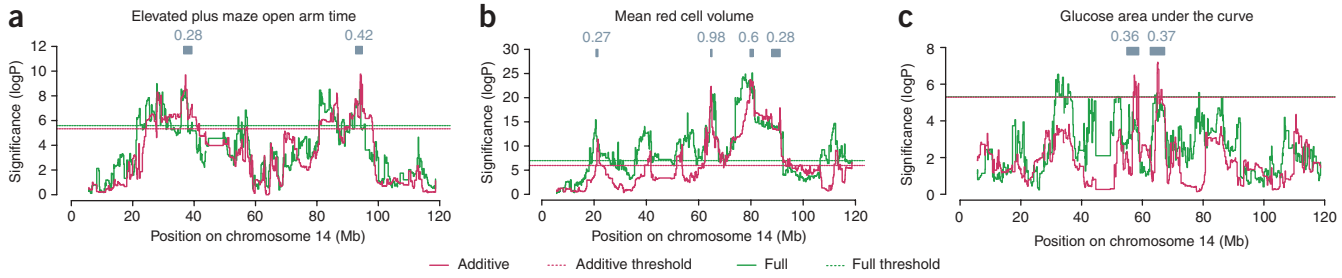

Figure 1 Fine-mapping of three phenotypes on chromosome 14. Phenotypes are (**a**) time spent in the open arm of the elevated plus maze; (**b**) the mean volume of red cells; and (**c**) area under the curve of glucose concentration. Each graph plots the strength of association between phenotype and estimated founder haplotypes for the additive (red) and full (green) genetic models. The vertical axes give the negative $\log_{10}$ of the *P* value (logP) of the test of association. The horizontal axes give the location in Mb. Horizontal dashed lines mark permutation-derived thresholds corresponding to $E = 0.01$. Dark gray boxes in the top margins show 95% c.i. for additive QTLs detected by multiple QTL analysis, with the corresponding bootstrap posterior probability (BPP) above each box.

is logP = 4.84 and at $E < 0.01$ is logP = 5.60. The median logP ($E < 0.5$) is 3.65, corresponding to 4,466 independent tests, so LD reduces the effective number of tests to approximately one-third.

We identified QTLs that exceeded an *E* value threshold of 0.05 for all phenotypes except two: the rates of decline of glucose and insulin from an intraperitoneal glucose tolerance test. We also mapped coat color as a test of the method. All known coat color loci mapped between adjacent markers that flank their known positions. To illustrate the complexity of the QTL landscape, **Figure 1** shows additive and full model analyses (in which the full model allows for interactions between alleles, including dominance) for three phenotypes on chromosome 14: mean cellular volume of red blood cells, time spent in the open arms of an elevated plus maze (a measure of anxiety) and a measure of glucose tolerance. The analyses uncover a large number of peaks exceeding the threshold, and, although full and additive models (**Fig. 1**) are broadly concordant, there are regions in which one model seems to do better than the other.

Multiple QTL models

A simple interpretation of the data from **Figure 1** is that many hundreds of QTLs underlie each phenotype. In fact, among all phenotypes, we found 9,269 candidate peaks that had $E < 0.05$ and that were separated by more than 8 Mb (at which distance the correlation between markers has fallen to an R^2 of < 0.2). We do not, however, believe that all these QTLs are genuine. If they were, then, given the effect size estimates we obtained, we would have explained more than 100% of the variation of many phenotypes, which is impossible.

The family structure in our sample causes genotype correlations (or LD) that inflate QTL effects and complicate the interpretation of QTL mapping results: the effect at one QTL could be attributable to the effect at a second locus, possibly on a different chromosome. Genotype correlations can also result in one QTL becoming nonsignificant given the presence of a second, correlated effect, even if both are genuine. Therefore, we need to consider the joint effects of QTLs and identify a set of QTLs that act independently to explain the phenotypic variation.

To deal with these issues, we used model-averaging techniques. There are many possible multiple-QTL models built from the candidate QTLs found in the genome scans that can explain the data. For simplicity, and to avoid the instability associated with excessive numbers of parameters, we considered only models built from autosomal additive QTLs identified in the genome scans. We created multiple data sets of 2,000 mice for analysis by randomly selecting animals from the genuine data set. We sampled with replacement so that some animals were resampled several times and others not at all. For each of these bootstrap data sets, we constructed a multiple QTL model iteratively by forward selection. Starting with a model with no QTLs, we tested the effect of adding a candidate QTL. The QTL with the most significant improvement in fit was added to the model, provided the improvement exceeded a 5% genome-wide significance threshold. The procedure was repeated until no more significant QTLs could be added. We then averaged across all models. We scored each QTL by its bootstrap posterior probability (BPP), the proportion of bootstraps in which it was selected[14,15], and reported its apparent effect size, defined as the percentage variance explained by the QTL after allowing for other QTLs and covariates in the model, averaged across all models containing that QTL.

We used simulation to assess the performance of the model averaging method under conditions that reproduced the variance components structure observed in our pedigree-based analysis of the data[16]. In particular, we wished to take into account the consequences

Table 2 Performance of multiple QTL modeling

BPP threshold[a]	Numbers of QTL[b]	Proportion of detected QTLs that are true[c]	Proportion of true QTLs detected[d]	Expected number of false QTLs per genome scan[e]
0.05	3,127	0.26	0.90	0.47
0.10	2,119	0.41	0.90	0.43
0.20	1,105	0.63	0.89	0.32
0.25	843	0.70	0.89	0.25
0.30	633	0.75	0.89	0.21
0.40	364	0.82	0.89	0.17
0.50	251	0.85	0.88	0.12
0.75	58	0.91	0.87	0.04
0.90	30	0.91	0.85	0.03
1.00	13	0.96	0.60	0.00

[a]The bootstrap posterior probability (BPP) for a QTL peak is the proportion of bootstraps in which that peak was included in the multiple-QTL analysis. Each row gives summary statistics for QTL peaks that have a BPP exceeding the stated BPP threshold. [b]The number of QTL peaks found exceeding the BPP in multiple-QTL modeling of 101 phenotypes. [c]Estimated proportion of QTLs detected in a multiple-QTL analysis that are within 4 Mb of a true QTL and exceed the BPP, in a simulation with seven 5% QTLs. [d]Estimated proportion of true QTLs lying within 4 Mb of a detected QTL that exceed the BPP in the same simulations. [e]Expected number of false-positive QTLs per genome scan under a simulated null-QTL model with 35% infinitesimal genetic variance.

of genotype correlations, as well as large-family and environment effects that may cause false-positive QTLs. As we found an average of seven QTLs per model, for each simulation we selected seven SNPs at random across the genome as true QTL locations. Then, without altering the genotypes or family structure, we simulated QTLs at these SNPs, each with an effect size explaining 5% of the phenotypic variance. Finally, we added a 20% shared environment effect, resulting in a breakdown of variance components (approximately 35% genetic, 20% shared environment, 45% random) typical of the phenotypes scored[16]. The simulated effect sizes were chosen to be representative of those detected in the multiple QTL models, bearing in mind the effect size estimates obtained from the model averaging method are downwardly biased: a simulated 5% effect corresponds to an apparent 3% effect in the heterogeneous stock, typical of the apparent effect sizes we obtained in the genome scans. To investigate false-positive rates, we simulated a null QTL scenario that consisted solely of 35% genetic and 20% environment random effects, under the assumption that there are no detectable QTLs, but that the genetic variance arises from a large number of infinitesimally small effects[17].

Table 2 shows the results of the simulations and provides performance measures for BPP thresholds between 0.05 and 1. Each row collates the results for all peaks detected by model averaging with a BPP exceeding the given threshold. We report three measures: first, the proportion of QTLs detected after model averaging that exceed a given BPP threshold and match a true QTL; second, the proportion of true QTLs that are detected (this figure estimates power); and third, the expected number of false-positive QTLs detected per genome scan under the null QTL scenario. We see that over a wide range of BPP values, the power to detect a QTL is about 90%. Therefore, noting that a simulated 5% QTL corresponds to an apparent effect of about 3%, our experiment is powered to detect apparent effects $> \sim 2\%$. Furthermore, a QTL that exceeds a BPP threshold of 0.5 will be true in 85% of cases, and in 70% of cases for a threshold of 0.25. Finally, under the assumptions of the simulations consisting solely of infinitesimal effects, at a BPP threshold of 0.25, about one false-positive QTL occurs every four genome scans. This value can be thought of as the upper limit of the number of false-positive QTLs under an infinitesimal genetic model[17].

We identified 843 QTLs for 97 phenotypes at a BPP threshold of 0.25. There were four phenotypes for which no QTLs were identified at this threshold: entries into the closed arms of the elevated plus maze, defecation in the open-field arena and the rates of decline of glucose and insulin from an intraperitoneal glucose tolerance test. **Table 3** summarizes all major QTLs with BPP > 0.90, and **Supplementary Table 1** reports all QTLs with BPP > 0.25. **Table 3** also reports apparent effect sizes for all QTLs, a lower bound on the real QTL effect size. By contrast, genome scans report the percentage variance explained by each QTL in isolation after removing covariate effects, which are overestimates. The medians of the two measures are 2.5% and 13.7%, respectively.

The picture we see is of many loci each contributing a small proportion to the variance. Large-effect QTLs are rare: only ten QTLs have apparent sizes greater than 5%; these are all included in **Table 3**. One hundred and nine QTLs have apparent effect sizes less than 2%. On average, the QTL-based additive variance is 73% of the pedigree-based variance (calculated in a variance components framework from the last three generations of the heterogeneous stock pedigree[16]), and there is a high correlation between the two estimates ($R = 96\%$; **Fig. 2**). Thus, for the majority of phenotypes, the QTLs we have identified explain about three-quarters of the additive genetic variance.

High-resolution architecture

We mapped 575 QTLs into 50% confidence intervals (c.i.) of ≤ 1 Mb by bootstrapping, and 101 QTLs to 95% c.i. of < 1 Mb. The mean size of the 95% intervals is 2.78 Mb (s.d. 1.47) and 0.86 Mb (s.d. 0.80) for the 50% intervals. We calculated the number of annotated genes under each peak and report the results in **Supplementary Table 1** online. The distribution of gene counts is shown in **Supplementary Figure 1** online, for the 50% c.i. The distribution is approximately exponential, with 119 QTLs containing no annotated genes and one QTL containing 106 (a set of odorant receptors). As the true location of the QTL lies within the interval in only 50% of cases, this result indicates that about 60 QTLs contain no annotated gene.

We considered whether the QTLs have substructure, despite the small intervals into which they have been mapped. We asked whether a single diallelic variant could be responsible for each QTL. For each of the 127 possible diallelic founder strain distribution patterns, we compared the fit of a model in which the effects attributable to founder strains sharing the same allele are identical to the unconstrained case in which each heterogeneous stock progenitor has a distinct QTL allele (a merge analysis[5]). In 27% of cases, every possible diallelic strain distribution pattern performed significantly worse than

Table 3 Highly significant QTLs

					95% c.i.	
Phenotype	Chr	BPP	logP	Effect size	From	To
Alkaline phosphatase	4	1.00	82.72	14.23	136.10	137.89
Aspartate transaminase	11	1.00	9.22	3.28	56.16	59.60
High density lipoproteins	1	1.00	54.34	9.14	171.51	172.06
Total cholesterol	1	1.00	47.25	9.99	171.15	171.51
Time freezing during cue	15	1.00	28.05	10.10	91.37	92.23
Ear punch hole area	7	1.00	25.28	5.00	75.79	77.57
Startle response	15	1.00	24.49	5.74	91.37	92.62
Basophils	6	1.00	8.35	3.54	79.77	82.74
Hematocrit	1	1.00	9.21	3.41	188.03	189.80
Mean platelet volume	1	1.00	36.93	10.49	170.99	173.52
$CD4^+/CD8^+$	17	1.00	21.53	3.90	42.83	43.25
$CD4^+/CD8^+$	17	1.00	65.76	11.93	30.84	32.08
Percentage $CD4^+/CD3^+$	17	1.00	54.03	11.22	31.08	32.08
Red cell mean cellular volume	1	1.00	17.73	4.09	129.79	131.93
Percentage $CD8^+$ cells	17	1.00	48.61	8.00	31.08	32.08
Mean cellular hemoglobin	9	0.99	17.17	3.36	107.90	111.90
Mean cellular hemoglobin	14	0.99	20.04	4.05	65.09	66.12
Red cell mean cellular volume	14	0.98	17.37	3.73	64.70	65.09
Open field arena center time	5	0.97	12.82	3.37	48.80	51.62
Red cell distribution width	7	0.97	18.60	3.98	98.26	101.18
Mean cellular hemoglobin	1	0.97	15.25	3.30	129.79	132.16
Body mass index	2	0.97	10.70	3.34	125.95	130.20
Weight 6 weeks	7	0.96	17.02	3.24	51.38	54.75
Mean cellular hemoglobin conc.	9	0.95	14.54	3.69	107.77	109.30
Red cell distribution width	9	0.94	12.69	2.60	108.85	111.08
Startle response	11	0.93	13.10	3.15	118.17	120.25
Body length (cm)	11	0.92	10.85	3.14	79.27	83.14
Latency to enter EPM open arms	5	0.92	12.39	NA	50.57	52.09
Mean cellular hemoglobin	11	0.91	12.27	2.57	4.93	5.85

The table shows all QTLs with BPP greater than 0.9. Chr: chromosome number. BPP: the bootstrap posterior probability of the QTL. LogP: negative $\log_{10}$ of the *P* value for the QTL, computed as the average logP for the QTL in all multiple QTL models that contained it. Effect size: the percentage variance explained by a given QTL after allowing for the other QTLs in the model, averaged across all models containing that QTL. The effect sizes for the latency measures are unavailable. 95% c.i.: the bootstrap 95% confidence interval for the QTL (in Mb).

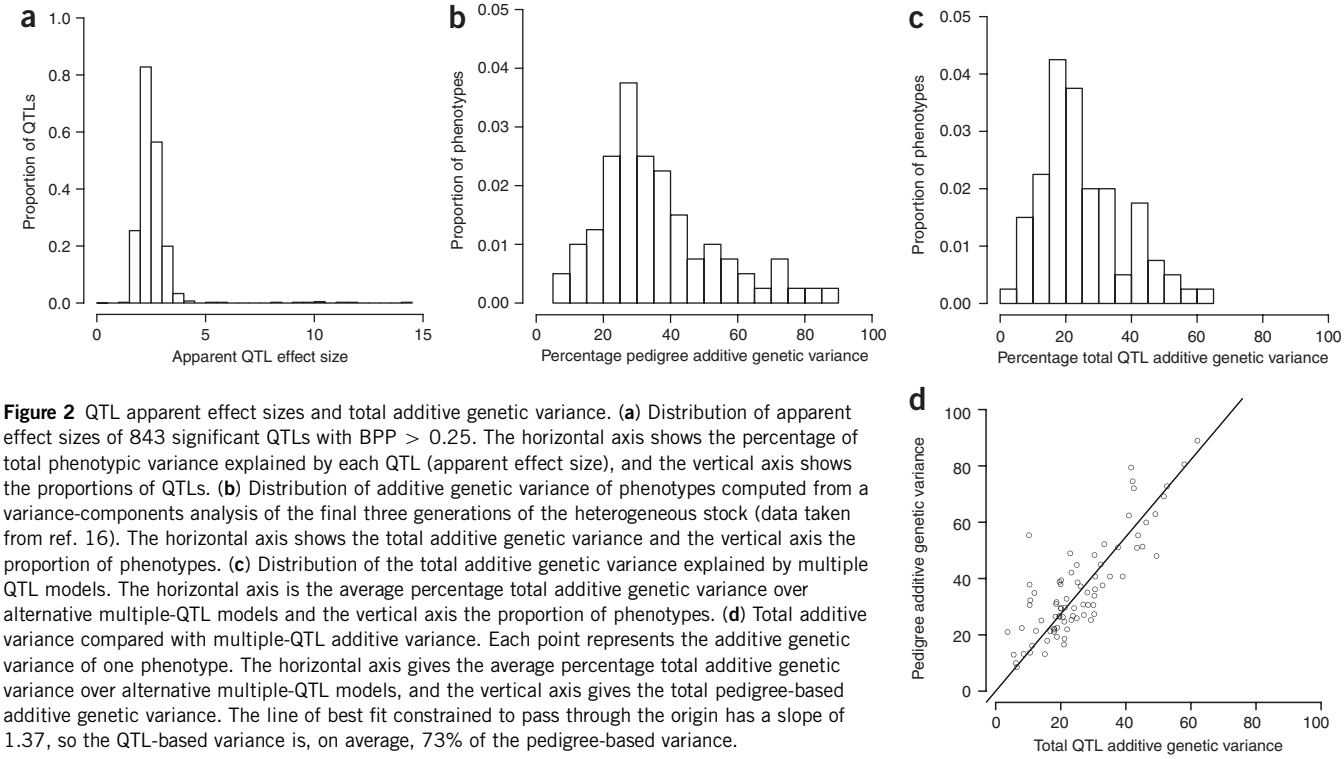

Figure 2 QTL apparent effect sizes and total additive genetic variance. (**a**) Distribution of apparent effect sizes of 843 significant QTLs with BPP > 0.25. The horizontal axis shows the percentage of total phenotypic variance explained by each QTL (apparent effect size), and the vertical axis shows the proportions of QTLs. (**b**) Distribution of additive genetic variance of phenotypes computed from a variance-components analysis of the final three generations of the heterogeneous stock (data taken from ref. 16). The horizontal axis shows the total additive genetic variance and the vertical axis the proportion of phenotypes. (**c**) Distribution of the total additive genetic variance explained by multiple QTL models. The horizontal axis is the average percentage total additive genetic variance over alternative multiple-QTL models and the vertical axis the proportion of phenotypes. (**d**) Total additive variance compared with multiple-QTL additive variance. Each point represents the additive genetic variance of one phenotype. The horizontal axis gives the average percentage total additive genetic variance over alternative multiple-QTL models, and the vertical axis gives the total pedigree-based additive genetic variance. The line of best fit constrained to pass through the origin has a slope of 1.37, so the QTL-based variance is, on average, 73% of the pedigree-based variance.

the unconstrained model at $P < 0.01$, indicating that at these QTLs, no single diallelic variant was responsible for the QTL effect. **Figure 3** gives one example in which the diallelic model did not perform as well as the multiallelic model.

Non-additive and sex-specific QTLs

We tested for the presence of non-additive effects, including dominance, at each of the 843 QTLs. We found 272 QTLs with significant non-additive effects at logP > 2, and 39 QTLs at logP > 5 (approximate genome-wide *E* threshold of 0.05). The most significant non-additive effects were for immunological phenotypes around the MHC on chromosome 17. Moreover, we identified 5,853 candidate loci in the genome scan data for which there was a significant non-additive, but no significant additive, effect.

We looked for sex-specific loci by testing the significance of a sex-by-locus interaction across the genome for all phenotypes. The total number of loci with significant sex effects irrespective of the presence of an additive QTL is small: only 190 candidate loci affecting a total of 63 phenotypes showed evidence for sex interaction at a genome-wide *E* threshold of 0.05 (**Supplementary Table 2** online). We would expect fewer than 20 sex-specific QTLs to remain significant in a multiple QTL analysis, given the observed tenfold reduction in the number of significant main-effect QTLs after applying model averaging. However, we identified two highly significant sex interactions, both coinciding with major additive QTLs: on chromosome 1 at 170 Mb for mean platelet volume (gene-by-sex logP = 38.07) and on chromosome 6 at 104 Mb for several related immunology measures, including percentage $CD8^+$ (gene-by-sex logP = 10.51).

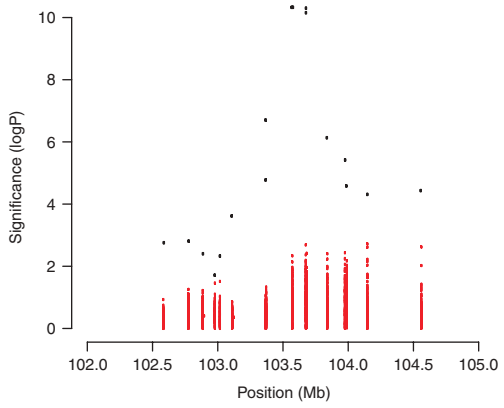

Figure 3 Test for the presence of a diallelic QTL on chromosome 12. The phenotype is time spent in the open arms of the elevated plus maze. The horizontal axis shows the chromosomal position in Mb and the vertical axis the significance of association, measured as the negative $\log_{10}$ of the *P* value (logP). The black dots give the logP of multipoint HAPPY analysis, which tests the significance of association at each marker interval as if there were eight QTL alleles representing the genetic effects of the eight heterogeneous stock progenitors. The red dots represent the results of constraining the genetic effect to each of the 127 possible diallelic strain distribution patterns in every marker interval. As shown here, all of the diallelic strain distribution patterns have lower logP values than that of the multiallelic analysis, indicating that the genetic effect cannot be attributable to a single diallelic effect.

DISCUSSION

We present here the first comprehensive fine-scale analysis of the genetic architecture of quantitative variation in the mouse. We have shown that it is possible to map QTLs at high resolution across the whole genome and in doing so we have dealt with issues general to all genome-wide association studies. Most importantly, we found small,

but significant, correlations between QTLs. We extended an earlier use of bootstrap aggregation analysis in plant genetics[15] that takes into account multiple QTLs and family and environmental effects, all of which inflate the genome scan test statistics. In a departure from using either a genome-scan significance threshold or a false-discovery rate to accept or reject QTLs as genuine, we give bootstrap posterior probabilities, which must be calibrated to the specific experimental design. For this experiment, we suggest that a BPP threshold of 0.25 is a suitable balance between false-positive and false-negative rates.

Table 4 Comparison of QTLs segregating in inbred crosses with those found in the heterogeneous stock (HS)

		Prior cross		Position (in Mb)		HS analysis	
Phenotype	Strain crossed with C57BL/6J	Ref.	Chr.	Prior cross	HS	BPP	LogP
Basal glucose	C3H/HeJ	35	11	12	6	0.18	5
Context conditioning	C3H/HeJ	36	1	85	75	0.15	7
Context conditioning	C3H/HeJ	36	1	158	192		6
Context conditioning	DBA/2J	37	10	32	15	0.18	5
Context conditioning	DBA/2J	37	16	100	88	0.21	7
Cue conditioning	DBA/2J	37	1	133	129		6
Cue conditioning	DBA/2J	37	10	32			
Enhanced pause	A/J	38	2	160	169	0.11	13
EPM (open arm entries)	BALB/cJ	39	1	156	160	0.58	9
EPM (open arm entries)	BALB/cJ	39	4	70	73	0.25	8
EPM (open arm entries)	BALB/cJ	39	7	60	67		5
EPM (open arm entries)	BALB/cJ	39	8	80	72	0.05	6
EPM (open arm entries)	BALB/cJ	39	15	30	25	0.11	8
EPM (open arm entries)	BALB/cJ	39	18	50	68		5
EPM (open arm entries)	A/J	40	5	123	132	0.28	7
Glucose, 30 min	C3H/HeJ	35	13	113	111	0.02	7
Glucose, 60 min	C3H/HeJ	35	13	109	111	0.02	8
HDL	C3H/HeJ	41	1	171	171	1.00	72
HDL	DBA/2	41	3	66	70	0.30	12
HDL	DBA/2	41	6	136	142	0.03	22
Open field activity	A/J	42,43	1	156	160	0.18	5
Open field activity	BALB/cJ	39	1	200	185	0.13	8
Open field activity	DBA/2J	44	3	142	145	0.07	12
Open field activity	BALB/cJ	39	4	70	56	0.04	7
Open field activity	BALB/cJ	39	7	60	58	0.45	8
Open field activity	BALB/cJ	39	8	80	76	0.03	7
Open field activity	BALB/cJ	39	15	30	21	0.02	6
Open field activity	BALB/cJ	39	18	48	58	0.14	9
Open field activity	A/J	43	6	10	49		6
Open field activity	DBA/2J	44	7	110	100	0.02	9
Open field activity	A/J	42	10	120	124	0.27	13
Open field activity	DBA/2J	44	15	26	22	0.06	7
Open field activity	DBA/2J	44	19	0	5	0.06	7
Pulmonary resistance	A/J	45	2	158	169	0.04	13
Pulmonary resistance	A/J	46	2	21	18	0.09	12
Weight, 10 weeks	DBA/2J	47	1	90	82	0.46	18
Weight, 10 weeks	AKR/J	48	2	91	88	0.34	16
Weight, 10 weeks	DBA/2J	47	6	11	4	0.83	7
Weight, 10 weeks	DBA/2J	47	15	87	92	0.64	18
Weight, 10 weeks	AKR/J	48	17	12	22	0.73	9

Data are from published crosses (prior cross) that used heterogeneous stock progenitor strains and reported QTLs that exceeded genome-wide significance. All prior experiments used a cross between C57BL/6J and the strain given in column 2. BPP is the bootstrap posterior probability of the QTL in the heterogeneous stock. LogP is the negative logarithm of the *P* value observed in the HS genome scans. Missing data represent QTLs not mapped in the HS.

At first glance, a low BPP threshold seems inappropriate, because it implies that most true QTLs are present in only a small fraction of multiple-QTL models. Moreover, **Table 2** shows that many genuine QTLs have BPP values as low as 0.05. The explanation is that genuine QTLs share variance because of genotype correlation. Fitting a multiple-QTL model by a single instance of forward selection is unlikely to identify all genuine QTLs, because those added to the model early on absorb most of the genetic variance. Bootstrapping perturbs the order in which QTLs are added to the model, and false-positive QTLs are included in repeated model building at much lower frequency. In other words, even if a QTL is included in a modest fraction of models, there is a good chance it is genuine. Therefore, whereas QTLs exceeding the 0.25 threshold should be prioritized for further investigation, a lower BPP threshold is appropriate if prior support for a QTL exists (for example, in data from another mapping experiment). Otherwise, QTLs detected at lower BPP thresholds will require further experimental confirmation.

An important issue is the extent to which we have replicated QTLs known to segregate in crosses that include the eight heterogeneous stock progenitor strains. We compared our results with those reported for eleven phenotypes (**Table 4**) mapped in crosses between C57BL/6J and one of the other eight strains. Forty QTLs have been reported that exceed genome-wide significance (references given in **Table 4**). In 39 cases, we found at least one QTL in the heterogeneous stock within 20 Mb of the peak position in the inbred cross (and hence likely to lie within the peak's 95% confidence interval) (**Table 4**). Twelve have a BPP > 0.25, and a further nine have a BPP > 0.1, so from **Table 2** we estimate we have mapped between one-third and one-half of the QTLs found in inbred strain crosses. We would not expect to find all previously identified QTLs because power depends on the QTL allele frequency, which is prone to genetic drift in the heterogeneous stock, and the effect size, which may fall below the experiment's limit of detection. Furthermore, our analysis does not provide unambiguous strain origins for alleles, so we cannot be certain we have identified the same QTLs.

Our results uncover notable complexity in the QTL landscape. First, the genetic component of almost all phenotypes arises from a large number of small effect loci. We were able to detect QTLs with apparent effect sizes $<2\%$, and yet on average, all QTLs accounted for only about three-quarters of the additive genetic variance (**Fig. 2**), indicating that there remain many even smaller-effect QTLs to be discovered. It is notable that, irrespective of the phenotype, QTLs explain about the same proportion of the additive variance, suggesting considerable similarity in the genetic architecture of all phenotypes. Second, genetic heterogeneity exists at many QTLs in the heterogeneous stock: failure to find a diallelic variant that could explain QTL action at almost one-third of the additive QTLs indicates they arise from the joint effect of molecular variants with different strain distribution patterns (**Fig. 3**) Therefore, many QTLs are likely to fractionate further. Indeed, we have already dissected one QTL on chromosome 1 into three effects[18]. Similar complexity has emerged from high-resolution mapping in other organisms[19,20]. It will be difficult to isolate such QTLs by making congenics or by testing for the QTL in a recombinant inbred segregation test[21] or by comparing the results of different inbred strain crosses[22]. At these loci, alternative approaches, such as mapping with further outbred stocks, will be needed[18].

We suspect that an analysis of interactions will uncover further complexity. We have found that environmental covariates are involved in an unexpectedly large number of significant interactions with genetic background and that interaction effects are more frequent

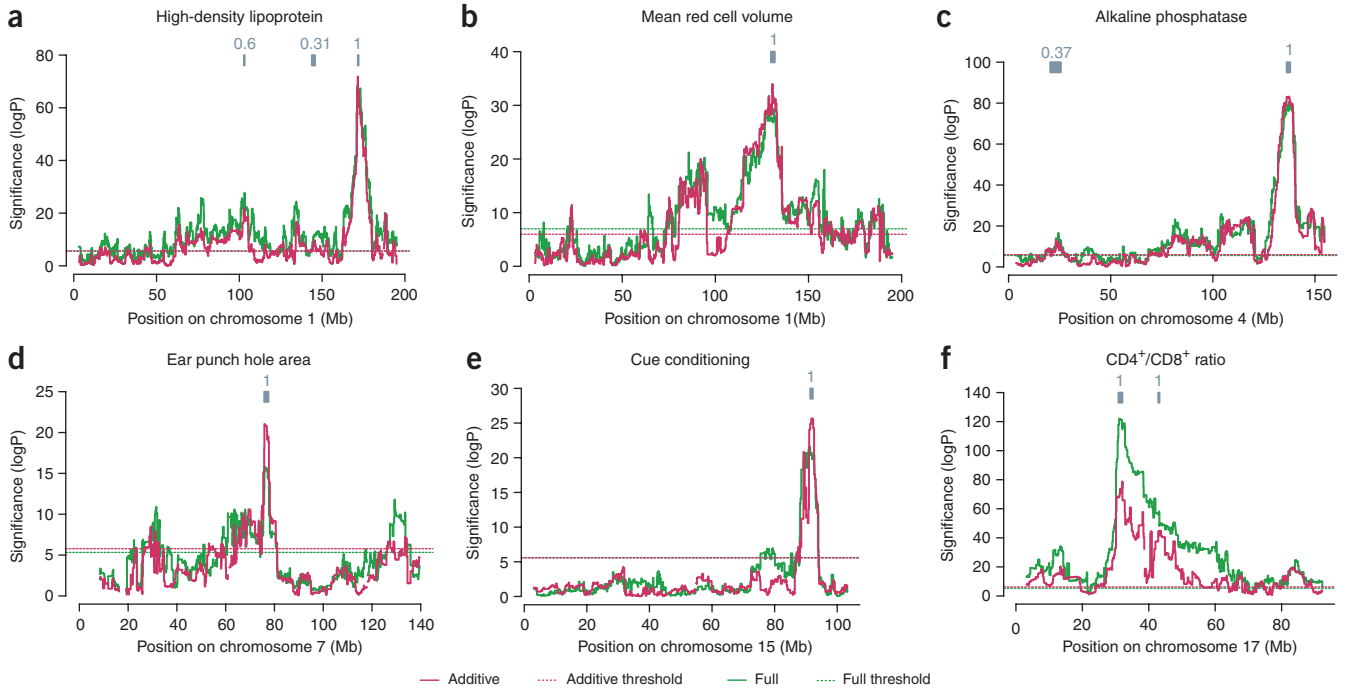

Figure 4 Large-effect QTLs that suggest a small number of candidates for molecular cloning. Graphs show chromosome scans for (**a**) serum high-density lipoprotein, (**b**) mean red cell volume, (**c**) serum alkaline phosphatase, (**d**) extent of wound healing following a 2-mm ear punch, (**e**) startle response in cue conditioning and (**f**) ratio of CD4 to CD8 lymphocytes. Axes are defined as for **Figure 1**. Each scan has one or more thin gray boxes labeled '1', indicating a 95% confidence interval for a QTL with 100% bootstrap posterior probability.

and larger than the main effects[16]. Mapping the QTLs responsible for these gene-environment interactions together with the mapping of epistatic loci remain important challenges that, once met, may increase the yield of QTLs significantly. However, our search uncovered little evidence for the existence of sex-specific QTLs.

Although we have not achieved single-gene resolution in all cases, over 300 QTLs contain fewer than three genes within their 50% c.i., which should accelerate the process of proving a gene's involvement at these loci. Additionally, a small number of QTLs have large effect sizes and narrow confidence intervals. These include HDL cholesterol (chromosome 1), mean red cell volume (chromosome 1), serum alkaline phosphatase (chromosome 4), wound healing from ear punch (chromosome 7), cue conditioning and startle response (chromosome 15) and ratio of CD4 to CD8 cells (chromosome 17) (**Fig. 4**). The HDL and the alkaline phosphatase QTLs have been discovered and attributed to variation in the *Apoa2* gene[23–25] and the alkaline phosphatase 2 gene (*Akp2*)[26], respectively. However, molecular characterization of the many small-effect QTLs we have identified remains an important challenge. Integration of expression, sequence and other functional data with mapping information from inbred strains has been shown to aid QTL identification[27,28] and should accelerate the functional characterization of genes involved in human diseases, when combined with the mapping results reported here. For this purpose, the raw data and our analyses are publicly available at http://gscan.well.ox.ac.uk/.

METHODS

Experimental design. Breeding pairs for the heterogeneous stock were obtained at 50 generations after the foundation of the stock[6]. The phenotyping protocol is described in detail in ref. 13. All experiments were conducted under the auspices of the UK Home Office Project and personal licenses held by the authors. Briefly, same-sex littermates were housed together in cages and subjected to a standardized battery of tests. For each phenotype, we recorded coincident environmental and physiological covariates (described fully in ref. 16 and briefly here). Mouse-specific covariates were gender, age, cage, litter size and weight before each test; test-specific covariates were experimenter, test order and apparatus (if multiple testing units were used); time-related covariates were year, season, month, hour and study day (number of days into the study).

Phenotypes. All analyses were performed using the R statistical analysis package version 2.1.1 (ref. 29). We ranked all animals by the number of phenotypes obtained and selected the top 1,904 animals for genotyping, together with 298 of their parents. Phenotypes were analyzed using multiple linear regression (for quantitative traits), survival analysis (for latency measures), multinomial generalized linear models (for coat color) or logistic regression. Quantitative phenotypes were converted to have approximately Gaussian distributions, by Box-Cox transformation, or by converting their ranks to Normal percentiles[16]. Covariates (such as gender, experimenter, month, etc.) that explained a significant fraction ($>1\%$) of each phenotype's variance (or deviance) with analysis of variance (ANOVA) $P < 0.01$ were identified and included in all subsequent statistical analyses of the phenotype.

Marker selection and genotyping. We selected SNPs from across the genome that distinguish between the eight heterogeneous stock founders (A/J, AKR/J, BALBc/J, CBA/J, C3H/HeJ, C57BL/6J, DBA/2J and LP/J). Where possible, we used published data sets to select SNPs that are validated and polymorphic in at least some of the heterogeneous stock founders[30–32]. These SNPs were merged together and remapped onto the mouse genome (build 34). SNPs closer than 50 kb with identical strain distribution pattern (SDP) were discarded. All gaps without SNPs over 500 kb were filled from the mouse sequence variation data available from dbSNP. Genotyping was carried out at Illumina, using their proprietary BeadArray genotyping platform. These SNPs were generated as part of the Wellcome-CTC mouse strain SNP genotype set.

Genome scans. QTLs were mapped across the genome using the multipoint HAPPY dynamic-programming algorithm and regression models (described in

ref. 4 and at http://www.well.ox.ac.uk/happy). In brief, using the genotypes of the heterogeneous stock animals and of the heterogeneous stock founder strains, we calculated the probability $F_{Li}(s,t)$ that the individual i was descended from founder strains s, t at marker interval L. The presence of a QTL was tested in each marker interval by fitting the regression model $E(y_i) = \Sigma_{st} F_{Li}(s,t)\, T_L(s,t)$ where $T_L(s,t)$ is the phenotypic effect due to the strains s, t at L. We term this the 'full' model. Additive QTL effects were tested by fitting a model with the constraint $T(s,t) = R(s) + R(t)$. The null hypothesis of 'no QTL in the interval' is equivalent to testing if none of the parameters $T(s,t)$ is significantly different. The presence of non-additive (such as dominance) effects was tested by comparing the fit of the full and additive models. Sex-specific QTLs were detected by testing for a difference between the additive QTL model above and a model in which males and females were allowed to have different additive genetic effects.

The distribution of the most significant logP detected in a genome scan with no QTL (the null model) was estimated for each phenotype by permutation of phenotypes between subjects while fixing the genotypes. This took into account the LD structure of the genome and effects due to the phenotype's distribution. For each of 200 permutations, the genome scan was repeated and the most significant logP recorded. A generalized extreme value distribution was fitted to the 200 permuted maximum logP scores, and quantiles from this fitted distribution were used to estimate genome-wide significance thresholds as described[33]. Within each genome scan, we identified candidate QTLs, defined as local maxima that exceeded the permutation-based genome-wide 5% significance threshold and that were separated by at least 8 Mb.

Multiple QTL analysis. Multiple QTL analysis was carried out by bootstrap aggregation[14,15]. First, the animals were re-sampled with replacement to create a bootstrap sample of appropriate size (see **Table 1**). Then, a forward-selection algorithm[11] was used to search the candidate additive autosomal QTL peaks. At each step in the iteration, the effect of adding each candidate QTL was tested conditional upon the presence of the covariates and the previously selected QTLs using a partial F-test (or χ^2 test for latency measures). The QTL with the most significant partial logP was added to the QTL set, provided its partial logP exceeded 5 (corresponding to a genome-wide significance level of between 1% and 5%, depending on the phenotype). The procedure terminated when no more significant QTLs could be added. From 700 bootstraps, we counted the proportion of times each QTL was selected to give the bootstrap posterior probability (BPP) that the QTL is present.

We calibrated the BPP on simulated multiple QTL sets. In order to preserve the pedigree structure and pattern of shared environment of the heterogeneous stock, we fixed the observed heterogeneous stock genotypes and cage assignments. Seven SNPs were selected at random across the genome, and a QTL explaining 5% of the phenotypic variance was generated according to the genotypes at each locus. To model the effects of shared environment, a common random effect equivalent to 20% of the variance was added to all animals sharing a cage. Then, a random uncorrelated effect accounting for the remaining variance was added to each animal. Finally, we simulated an infinitesimal model with no QTL but 35% genetic and 20% environment effects. In our analyses, a simulated and detected QTL are treated as matching if they lie with 4 Mb of each other.

Estimation of QTL and model effect sizes. Apparent QTL effects were calculated as the average percentage residual variance attributed to a QTL after removing the effects of covariates and all other QTLs. The bootstrap apparent effect size for each QTL was estimated as the average over all multiple QTL models containing the QTL. The total variance explained by multiple QTL models (the QTL-based additive genetic variance) was estimated as the average over all multiple QTL models. The pedigree-based additive genetic variance was estimated by variance components analysis as described in ref. 16.

Merge analysis. Merge analysis was carried out as described in ref. 5 to test if a SNP's strain distribution pattern among the founder strains was consistent with the genetic effects estimated in the heterogeneous stock at a QTL.

QTL confidence interval. Confidence intervals were estimated by bootstrapping[34]. We first defined an 8-Mb window around each peak that encompassed the likely variability of the summit. The region was then rescanned using 200 bootstrap re-samples. For each bootstrap, we recorded the location of the most significant point to build up a spatial distribution of the most significant locus.

URLs. Mouse genome scan database: http://gscan.well.ox.ac.uk; the Wellcome-CTC mouse strain SNP genotype set: http://www.well.ox.ac.uk/mouse/INBREDS; dbSNP: http://www.ncbi.nlm.nih.gov/SNP/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

This work was supported by grants from The Wellcome Trust and the European Union (contract LHSG-CT-2003-503265). D.G. holds a Wellcome senior fellowship in basic biomedical science (057733). We are grateful to R. Hitzemann for providing heterogeneous stock mice, to M. Daly and R. Williams for help with SNP selection and to S. McCormick and A. Morris for comments on the manuscript.

AUTHOR CONTRIBUTIONS

The heterogeneous stock study was designed by J.F., R.M., W.O.C., L.C.S., D.G. and J.N.P.R. Phenotype and genotype assessments were performed by S.B., P.K., J.F. and L.C.S. W.V., M.S.T., J.F. and R.M. were responsible for the analysis. All authors contributed to the writing of the paper.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Flint, J., Valdar, W., Shifman, S. & Mott, R. Strategies for mapping and cloning quantitative trait genes in rodents. *Nat. Rev. Genet.* **6**, 271–286 (2005).
2. Singer, J.B. *et al.* Genetic dissection of complex traits with chromosome substitution strains of mice. *Science* **304**, 445–448 (2004).
3. Mott, R. & Flint, J. Simultaneous detection and fine mapping of quantitative trait loci in mice using heterogeneous stocks. *Genetics* **160**, 1609–1618 (2002).
4. Mott, R., Talbot, C.J., Turri, M.G., Collins, A.C. & Flint, J. A method for fine mapping quantitative trait loci in outbred animal stocks. *Proc. Natl. Acad. Sci. USA* **97**, 12649–12654 (2000).
5. Yalcin, B., Flint, J. & Mott, R. Using progenitor strain information to identify quantitative trait nucleotides in outbred mice. *Genetics* **171**, 673–681 (2005).
6. Demarest, K., Koyner, J., McCaughran, J. Jr., Cipp, L. & Hitzemann, R. Further characterization and high-resolution mapping of quantitative trait loci for ethanol-induced locomotor activity. *Behav. Genet.* **31**, 79–91 (2001).
7. Talbot, C.J. *et al.* High-resolution mapping of quantitative trait loci in outbred mice. *Nat. Genet.* **21**, 305–308 (1999).
8. Talbot, C.J. *et al.* Fine scale mapping of a genetic locus for conditioned fear. *Mamm. Genome* **14**, 223–230 (2003).
9. Balding, D.J. Discussion on the meeting on 'Statistical modelling and analysis of genetic data'. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64**, 737–775 (2002).
10. Sillanpaa, M.J. & Corander, J. Model choice in gene mapping: what and why. *Trends Genet.* **18**, 301–307 (2002).
11. Broman, K.W. & Speed, T.P. A model selection approach for the identification of quantitative trait loci in experimental crosses. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64**, 641–656 (2002).
12. Ball, R.D. Bayesian methods for quantitative trait loci mapping based on model selection: approximate analysis using the Bayesian information criterion. *Genetics* **159**, 1351–1364 (2001).
13. Solberg, L.C. *et al.* A protocol for high-throughput phenotyping, suitable for quantitative trait analysis in mice. *Mamm. Genome* **17**, 129–146 (2006).
14. Breiman, L. Bagging predictors. *Mach. Learn.* **24**, 123–140 (1996).
15. Hackett, C.A., Meyer, R.C. & Thomas, W.T.B. Multi-trait QTL mapping in barley using multivariate regression. *Genet. Res.* **77**, 95–106 (2001).
16. Valdar, W. *et al.* Genetic and environmental effects on complex traits in mice. *Genetics* (in the press).
17. Visscher, P.M. & Haley, C.S. Detection of putative quantitative trait loci in line crosses under infinitesimal genetic models. *Theor. Appl. Genet.* **93**, 691–702 (1996).
18. Yalcin, B. *et al.* Genetic dissection of a behavioral quantitative trait locus shows that Rgs2 modulates anxiety in mice. *Nat. Genet.* **36**, 1197–1202 (2004).
19. Deutschbauer, A.M. & Davis, R.W. Quantitative trait loci mapped to single-nucleotide resolution in yeast. *Nat. Genet.* **37**, 1333–1340 (2005).
20. Mackay, T.F. The genetic architecture of quantitative traits: lessons from *Drosophila*. *Curr. Opin. Genet. Dev.* **14**, 253–257 (2004).
21. Darvasi, A. Experimental strategies for the genetic dissection of complex traits in animal models. *Nat. Genet.* **18**, 19–24 (1998).
22. Dipetrillo, K., Wang, X., Stylianou, I.M. & Paigen, B. Bioinformatics toolbox for narrowing rodent quantitative trait loci. *Trends Genet.* **21**, 683–692 (2005).

23. Doolittle, M.H., LeBoeuf, R.C., Warden, C.H., Bee, L.M. & Lusis, A.J. A polymorphism affecting lipoprotein A-II translational efficiency determines high density lipoprotein size and composition. *J. Biol. Chem.* **256**, 16380–16388 (1990).
24. Warden, C.H., Hedrick, C.C., Qiao, J.H., Castellani, L.W. & Lusis, A.J. Atherosclerosis in transgenic mice overexpressing apoliprotein A-II. *Science* **261**, 469–472 (1993).
25. Wang, X., Korstanje, R., Higgins, D. & Paigen, B. Haplotype analysis in multiple crosses to identify a QTL gene. *Genome Res.* **14**, 1767–1772 (2004).
26. Foreman, J.E. *et al.* Serum alkaline phosphatase activity is regulated by a chromosomal region containing the alkaline phosphatase 2 gene (*Akp2*) in C57BL/6J and DBA/2J mice. *Physiol. Genomics* **23**, 295–303 (2005).
27. Chesler, E.J. *et al.* Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. *Nat. Genet.* **37**, 233–242 (2005).
28. Schadt, E.E. *et al.* An integrative genomics approach to infer causal associations between gene expression and disease. *Nat. Genet.* **37**, 710–717 (2005).
29. R Development Core Team. *A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, 2004).
30. Cervino, A.C. *et al.* Integrating QTL and high-density SNP analyses in mice to identify *Insig2* as a susceptibility gene for plasma cholesterol levels. *Genomics* **86**, 505–517 (2005).
31. Wiltshire, T. *et al.* Genome-wide single-nucleotide polymorphism analysis defines haplotype patterns in mouse. *Proc. Natl. Acad. Sci. USA* **100**, 3380–3385 (2003).
32. Petkov, P.M. *et al.* Evidence of a large-scale functional organization of mammalian chromosomes. *PLoS Genet.* **1**, e33 (2005).
33. Valdar, W., Flint, J. & Mott, R. Simulating the collaborative cross: power of QTL detection and mapping resolution in large sets of recombinant inbred strains of mice. *Genetics* **172**, 1783–1797 (2006).
34. Visscher, P.M., Thompson, R. & Haley, C.S. Confidence intervals in QTL mapping by bootstrapping. *Genetics* **143**, 1013–1020 (1996).
35. Toye, A.A. *et al.* A genetic and physiological study of impaired glucose homeostasis control in C57BL/6J mice. *Diabetologia* **48**, 675–686 (2005).
36. Caldarone, B. *et al.* Quantitative trait loci analysis affecting contextual conditioning in mice. *Nat. Genet.* **17**, 335–337 (1997).
37. Wehner, J.M. *et al.* Quantitative trait locus analysis of contextual fear conditioning in mice. *Nat. Genet.* **17**, 331–334 (1997).
38. Ackerman, K.G. *et al.* Interacting genetic loci cause airway hyperresponsiveness. *Physiol. Genomics* **21**, 105–111 (2005).
39. Henderson, N.D., Turri, M.G., DeFries, J.C. & Flint, J. QTL Analysis of multiple behavioral measures of anxiety in mice. *Behav. Genet.* **34**, 267–293 (2004).
40. Cohen, R.M., Kang, A. & Gulick, C. Quantitative trait loci affecting the behavior of A/J and CBA/J intercross mice in the elevated plus maze. *Mamm. Genome* **12**, 501–507 (2001).
41. Wang, X. & Paigen, B. Genetics of variation in HDL cholesterol in humans and mice. *Circ. Res.* **96**, 27–42 (2005).
42. Gershenfeld, H.K. *et al.* Mapping quantitative trait loci for open-field behavior in mice. *Behav. Genet.* **27**, 201–210 (1997).
43. Singer, J.B., Hill, A.E., Nadeau, J.H. & Lander, E.S. Mapping quantitative trait loci for anxiety in chromosome substitution strains of mice. *Genetics* **169**, 855–862 (2005).
44. Steinberger, D. *et al.* Genetic mapping of variation in spatial learning in the mouse. *J. Neurosci.* **23**, 2426–2433 (2003).
45. De Sanctis, G.T. *et al.* Quantitative locus analysis of airway hyperresponsiveness in A/J and C57BL/6J mice. *Nat. Genet.* **11**, 150–154 (1995).
46. Ewart, S.L. *et al.* Quantitative trait loci controlling allergen-induced airway hyperresponsiveness in inbred mice. *Am. J. Respir. Cell Mol. Biol.* **23**, 537–545 (2000).
47. Morris, K.H., Ishikawa, A. & Keightley, P.D. Quantitative trait loci for growth traits in C57BL/6J x DBA/2J mice. *Mamm. Genome* **10**, 225–228 (1999).
48. Taylor, B.A. & Phillips, S.J. Obesity QTLs on mouse chromosomes 2 and 17. *Genomics* **43**, 249–257 (1997).