nature genetics

Genome-wide association study of behavioral, physiological and gene expression traits in outbred CFW mice

Clarissa C Parker[1–3,16], Shyam Gopalakrishnan[1,4,16], Peter Carbonetto[1,5,16], Natalia M Gonzales[1], Emily Leung[1], Yeonhee J Park[1], Emmanuel Aryee[1], Joe Davis[1], David A Blizard[6], Cheryl L Ackert-Bicknell[7,8], Arimantas Lionikas[9], Jonathan K Pritchard[10–12] & Abraham A Palmer[1,13–15]

Although mice are the most widely used mammalian model organism, genetic studies have suffered from limited mapping resolution due to extensive linkage disequilibrium (LD) that is characteristic of crosses among inbred strains. Carworth Farms White (CFW) mice are a commercially available outbred mouse population that exhibit rapid LD decay in comparison to other available mouse populations. We performed a genome-wide association study (GWAS) of behavioral, physiological and gene expression phenotypes using 1,200 male CFW mice. We used genotyping by sequencing (GBS) to obtain genotypes at 92,734 SNPs. We also measured gene expression using RNA sequencing in three brain regions. Our study identified numerous behavioral, physiological and expression quantitative trait loci (QTLs). We integrated the behavioral QTL and eQTL results to implicate specific genes, including *Azi2* in sensitivity to methamphetamine and *Zmynd11* in anxiety-like behavior. The combination of CFW mice, GBS and RNA sequencing constitutes a powerful approach to GWAS in mice.

In the last decade, GWAS have demonstrated that common alleles influence susceptibility to virtually all common diseases[1–3]. The success of GWAS in elucidating the genetic determinants of disease in humans is due in part to the large number of recombinations among unrelated individuals, which permits high-resolution mapping across the genome. One important conclusion from these studies is that most causal loci seem to be due to regulatory rather than coding polymorphisms[4].

Mice offer a powerful tool for elucidating the genetic architecture of complex traits: environmental factors can be held constant or systematically varied; genome editing permits experimental testing of identified genotype–phenotype relationships; most mouse genes have a human homolog, allowing rapid translation to humans; and relevant tissues can be obtained under tightly controlled conditions and used to identify gene eQTLs. However, the mouse populations used in most previous studies lacked sufficient recombination to narrow the implicated loci to a tractable size and thus generally failed to identify specific genes[5,6].

In this study, we mapped QTLs and eQTLs using CFW mice, which are a commercially available outbred population[7]. Although CFW mice were not developed for genetic research, they have several attractive properties. CFW mice were derived from a small number of founders and have subsequently been maintained as an outbred population for more than 100 generations, thus degrading LD between nearby alleles[8–10]. Although CFW mice have longer-range LD than most human populations, they have less LD than other commercially available laboratory mice[9] and therefore should provide fine-scale mapping resolution. In comparison to humans, the more extensive LD in CFW mice means that fewer markers are needed to perform GWAS and correspondingly lower levels of significance are required because fewer independent hypotheses are tested. We used GBS to overcome another barrier to GWAS in mice, which is the high cost and limited coverage of extant SNP genotyping arrays. Finally, on the basis of the importance of regulatory variation suggested by human GWAS[4,11], we identified eQTLs that co-mapped with behavioral QTLs in an effort to identify the most likely causal genes.

RESULTS

We phenotyped 1,200 male CFW mice for conditioned fear, anxiety-like behavior, methamphetamine sensitivity, prepulse inhibition, fasting glucose levels, body weight, tail length, testis weight, the weight of five hindlimb muscles, bone mineral density (BMD), bone morphology, and gene expression in prefrontal cortex, hippocampus and striatum (**Fig. 1**, Online Methods, **Supplementary Figs. 1–4** and **Supplementary Note**).

[1]Department of Human Genetics, University of Chicago, Chicago, Illinois, USA. [2]Department of Psychology, Middlebury College, Middlebury, Vermont, USA. [3]Program in Neuroscience, Middlebury College, Middlebury, Vermont, USA. [4]Natural History Museum of Denmark, Copenhagen University, Copenhagen, Denmark. [5]AncestryDNA, San Francisco, California, USA. [6]Department of Biobehavioral Health, Pennsylvania State University, University Park, Pennsylvania, USA. [7]Center for Musculoskeletal Research, University of Rochester, Rochester, New York, USA. [8]Department of Orthopedics and Rehabilitation, University of Rochester, Rochester, New York, USA. [9]School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, Aberdeen, UK. [10]Department of Genetics, Stanford University, Palo Alto, California, USA. [11]Department of Biology, Stanford University, Palo Alto, California, USA. [12]Howard Hughes Medical Institute, Stanford University, Palo Alto, California, USA. [13]Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, Illinois, USA. [14]Department of Psychiatry, University of California, San Diego, La Jolla, California, USA. [15]Institute for Genomic Medicine, University of California, San Diego, La Jolla, California, USA. [16]These authors contributed equally to this work. Correspondence should be addressed to A.A.P. (aapalmer@ucsd.edu).

Received 1 February; accepted 8 June; published online 4 July 2016; doi:10.1038/ng.3609

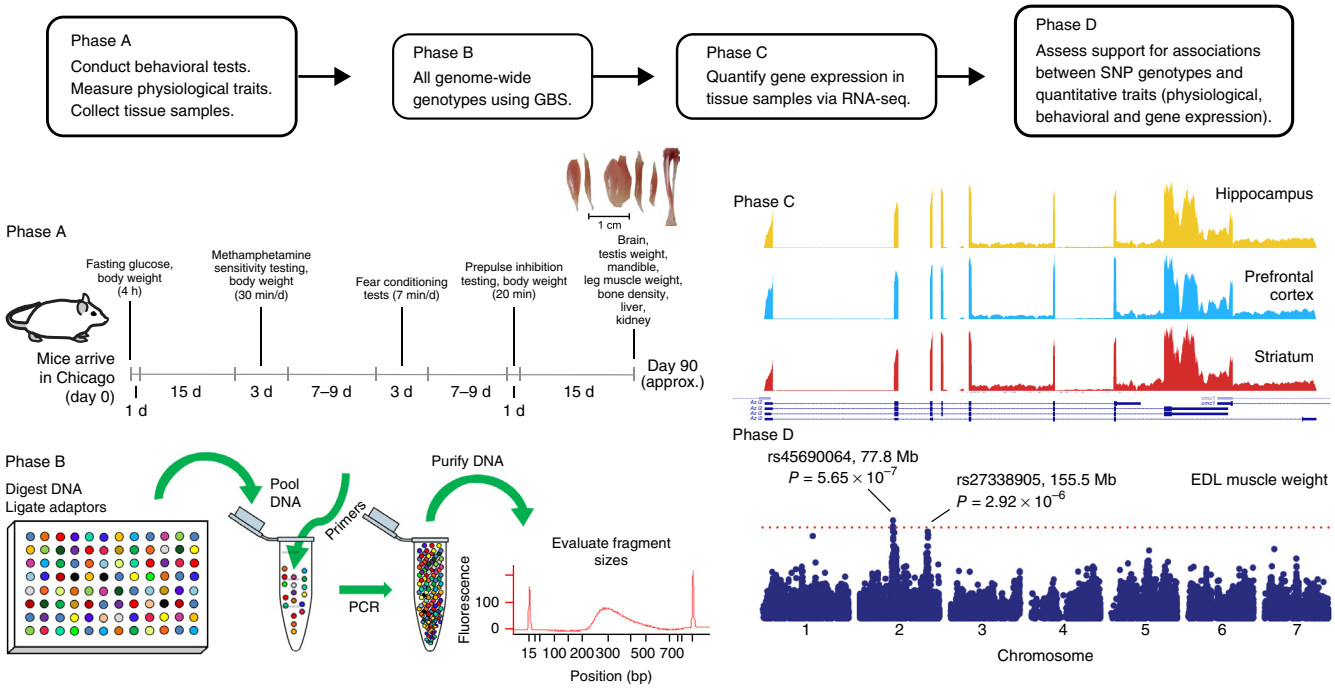

Figure 1 Components of the study. Our study consisted of four phases, including behavioral testing and measurement of physiological traits (phase A), GBS (phase B), measurement of gene expression in brain tissues using RNA-seq (phase C), and QTL mapping for physiological and behavioral traits and for gene expression (phase D). The red dotted line for phase D corresponds to $P = 2 \times 10^{-7}$.

Genotyping

Existing mouse SNP genotyping technologies, such as the Mouse Universal Genotyping Array (MUGA), MegaMUGA[12], the more recent GigaMUGA[13] and the Mouse Diversity Array (MDA)[14], were not designed to capture common genetic variation in the CFW population. Furthermore, we sought to reduce the cost of genotyping, which has been a barrier to GWAS in mice. Therefore, we adapted GBS, which was originally developed in maize[15], for use in mice. We used GBS to genotype 1,024 CFW mice and identified 92,734 autosomal biallelic SNPs after filtering, 79,284 (86%) of which were present in dbSNP (v137). The remaining 13,450 SNPs (14%) represent new SNPs that had not previously been reported. The distribution of GBS SNPs on autosomes is shown in **Figure 2a**. The non-uniform distribution of SNPs is likely due to differences in the numbers of polymorphic markers among all laboratory mice (**Fig. 2a** and **Supplementary Fig. 5**) and regions that are identical by descent among CFW mice. The non-uniform distribution of polymorphic SNPs seems to be a characteristic of CFW mice, as polymorphic SNPs identified by the MegaMUGA array showed a similar pattern ($r^2 = 0.43$ on a log scale; **Fig. 2a**).

To assess the quality of the GBS genotypes, we estimated the genotyping error rate in two ways. First, we compared genotypes for GBS SNPs against genotypes for SNPs that were also present on the MegaMUGA array in 24 CFW mice that were genotyped using both platforms. This comparison yielded an overall discordance rate of 3%. We obtained a second estimate of the error rate of 1.6% by comparing genotypes in pairs of haplotypes that were identical by descent. On the basis of these results, we concluded that GBS provided a larger number of polymorphic SNPs than were found using MegaMUGA.

Genetic architecture of the CFW population

Comparing LD in different populations is useful for gauging mapping resolution[16]. LD (r^2) decays rapidly in CFW mice in comparison to other populations, consistent with previous findings based on a much smaller number of SNPs[9,17] (**Fig. 2b**), supporting the suitability of the CFW population for high-resolution mapping. Notably, the majority of the SNPs we identified in CFW mice segregate among *Mus musculus domesticus*–derived laboratory strains (**Fig. 2c**). Unlike SNPs in the Collaborative Cross (CC) and Diversity Outbred (DO) lines, few of the SNPs found in the CFW population are derived from the *Mus musculus castaneous* and *Mus musculus musculus* subspecies[18–20]. When compared to a panel of inbred mice, CFW mice are most genetically similar to FVB/NJ mice (**Fig. 2c**).

Next, we considered the distribution of minor allele frequencies (MAFs) for SNPs genotyped in the CFW mice (**Supplementary Fig. 6**). The majority of SNPs (73%) had relatively high allele frequencies (MAF >0.05). This profile is consistent with the reported history of CFW mice, namely, a severe bottleneck at the inception of the CFW population, followed by expansion to create an outbred population with a modest effective population size[9]. The mean MAF for new SNPs was lower than for previously reported SNPs, consistent with the hypothesis that some of these new SNPs are unique to the CFW population.

Although we requested only one mouse from each litter, we were concerned that individuals in our study might have close familial relationships because they were sampled from a finite breeding population; however, we did not detect widespread population structure or cryptic relatedness in the CFW mice (**Supplementary Figs. 7–11**).

SNP heritability

Supplementary Table 1 shows SNP heritability[21,22], which is the proportion of variance in a trait explained by available SNP genotypes. SNP heritability estimates ranged from 9–60%, with a mean of 28%. The mean SNP heritability for physiological traits was slightly higher (32%) than for behavioral traits (27%).

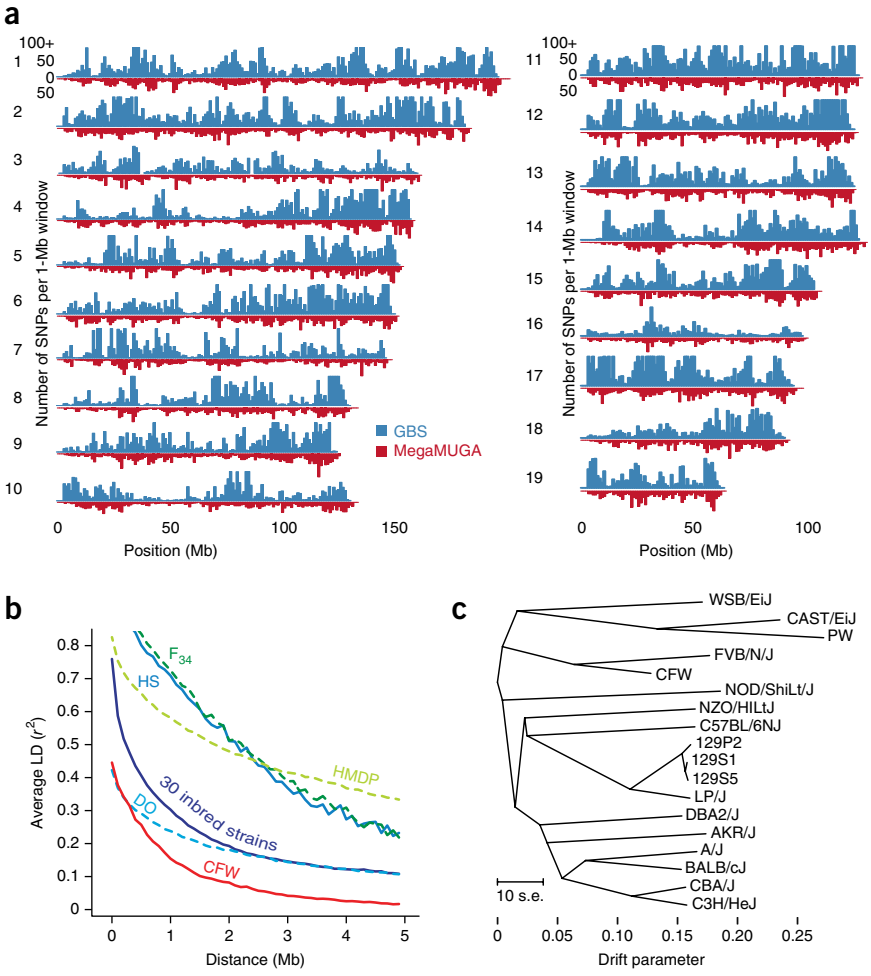

Figure 2 Genetic characteristics of the CFW mouse population. (**a**) Density of GBS SNPs on autosomal chromosomes. (**b**) Mean LD (r^2) decay rates estimated using frequency-matched SNPs[55], with MAF >20%, in a 34th-generation advanced intercross line (AIL) derived from LG/J and SM/J strains[43,46], heterogeneous stock (HS) mice bred for >50 generations[49], HMDP[83], a panel of 30 inbred laboratory strains[14,52], DO mice[12] and CFW mice. (**c**) Treemix analysis summarizing the genetic relationship between CFW mice and inbred strains in the Wellcome Trust sequencing panel.

Genome-wide association studies

We mapped QTLs for 66 behavioral and physiological phenotypes (**Fig. 3** and **Supplementary Tables 1** and **2**). We used GEMMA to fit a linear mixed model (LMM) and quantify support for an association at each SNP. We also used a simpler linear model that did not correct for population structure and observed that this model produced broadly similar results (**Supplementary Fig. 12**). However, we have presented the results from the LMM-based analysis because this approach may reduce subtle inflation of the test statistics due to close relationships or fine-scale population structure. We calculated a significance threshold via permutation, which is a standard approach for QTL mapping in mice that controls for the type I error rate[23,24] (**Supplementary Fig. 13**). This approach identified numerous QTLs for physiological and behavioral traits (**Fig. 3** and **Supplementary Figs. 14–18**) that exceeded $P = 2 \times 10^{-6}$ ($P < 0.1$). **Supplementary Table 2** contains more detailed information about all the most significant physiological and behavioral QTLs.

For testis weight, we found a strong association with rs6279141 on chromosome 13 ($P = 4.51 \times 10^{-18}$; **Fig. 3b**) that accounted for 7.5% of variation in this trait. The implicated region contained few genes (**Fig. 3d**), one of which was *Inhba*, a gene that has been shown to affect testis morphogenesis, testicular cell proliferation and testis weight in mice[25–27], and is therefore a promising candidate gene.

The strongest association for soleus muscle weight mapped to rs30535702 on chromosome 13 ($P = 8.33 \times 10^{-8}$) and explained 2.8% of trait variance. One of the genes in this interval, *Fst*, is known to influence muscle mass[28,29] and is a strong candidate to explain this association.

We identified several examples of pleiotropy. For example, two independently measured muscle weights, for tibialis anterior (TA) and extensor digitorum longus (EDL), were both associated with rs27338905 on chromosome 2, with this SNP accounting in each case for 2.3% of variation in weight. *Tp53inp2* is near the peak marker and is abundantly expressed in skeletal muscle[30], where it functions as a negative regulator of muscle mass[31]. Likewise, the weight of three muscles (gastrocnemius, EDL and soleus) mapped to the proximal end of chromosome 13; in each case, the minor allele was associated with increased muscle weight. Finally, on chromosome 12, we identified pleiotropic effects on tibia length and EDL weight.

Unexpectedly, we found that CFW mice seem to be predisposed to abnormally high BMD. This is a characteristic of CFW mice that does not seem to be shared with commonly used inbred laboratory strains (**Supplementary Fig. 3**). This 'abnormal BMD' phenotype was strongly associated with rs33583459 on chromosome 5 and rs29477109 on chromosome 11 ($P = 1.57 \times 10^{-9}$ and 1.12×10^{-14}, respectively). The locus on chromosome 5 contains a large number of genes, including *Abcf2* and *Slc4a2*. The human ortholog *ABCF2* has been associated with BMD in the largest GWAS of BMD completed thus far[32] and is highly expressed in osteoblasts[33]. *Slc4a2* has a critical role in osteoclasts in mice[34], and homozygous deletion of *SLC4A2* is associated with the osteopetrosis-like phenotype 'marble bone disease' in Red Angus cattle[35]. Thus, both *Abcf2* and *Slc4a2* are viable candidates for this region. The association on chromosome 11 contains the gene *Col1a1*. In humans, osteogenesis imperfecta type I can be caused by a null allele of *COL1A1* and results in gracile bones with decreased strength[36,37]. *COL1A1* is also associated with other bone size phenotypes[38], making *Col1a1* a likely causal gene for this locus.

Finally, we identified several associations for behavioral traits, including for methamphetamine sensitivity on chromosome 6 at rs22397909 ($P = 9.03 \times 10^{-7}$) and on chromosome 9 at rs46497021 ($P = 1.58 \times 10^{-6}$); these associations account for 2.6% and 2.1% of phenotype variance, respectively (**Fig. 4**). We also identified an association for anxiety-like behavior with rs238465220 on chromosome 13 ($P = 7.31 \times 10^{-8}$) that explains 3% of the trait variance. For prepulse inhibition (PPI; 12 dB), we identified associations with rs264716939 on chromosome 7 (**Fig. 3c**) and rs230308064 on chromosome 13 ($P = 1.18 \times 10^{-6}$ and 2.17×10^{-6}, respectively). There were many genes in the ~3-Mb region on chromosome 7 that were associated

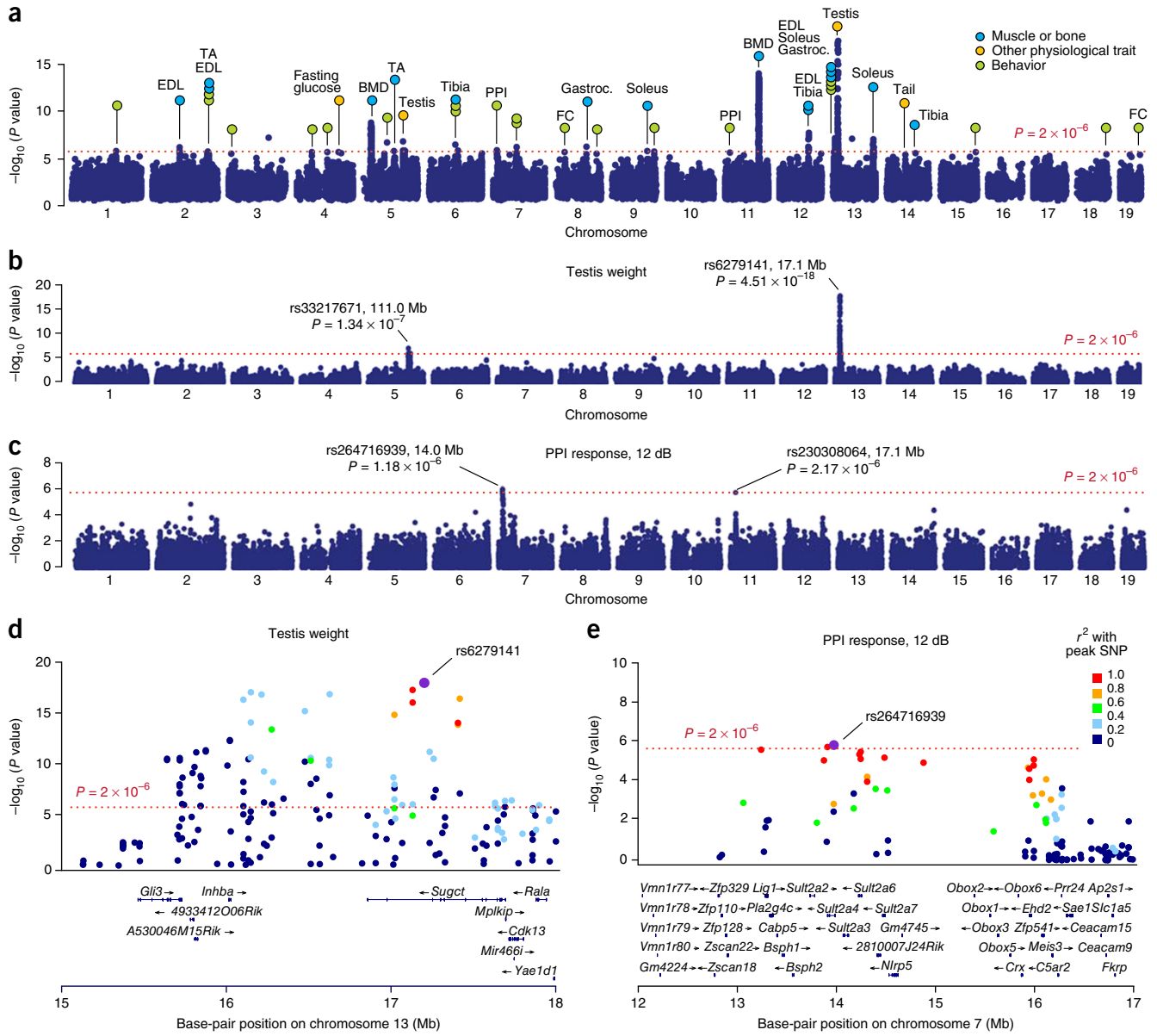

Figure 3 QTLs for physiological and behavioral traits. (**a**) Minimum *P* values for association across all tested behavioral and physiological phenotypes (see **Supplementary Tables 1** and **2** for details). FC, fear conditioning; gastroc., gastrocnemius. (**b**,**c**) Genome-wide scans for testis weight (**b**) and PPI in response to a 12-dB prepulse (**c**). (**d**) Association signal for testis weight near the QTL on chromosome 13. (**e**) Association signal for PPI near the QTL on chromosome 7. Dotted red lines correspond to significance thresholds ($P < 0.1$) estimated via permutation tests.

with PPI (**Fig. 3e**), making it difficult to identify the causal gene(s). Candidate genes for the associations with behavioral traits are discussed below.

Expression quantitative trait loci

In an effort to identify causal genes in our behavioral QTLs, we mapped eQTLs for three brain regions that are critical for the behaviors that we studied. We performed RNA-seq on mRNA from three brain regions: hippocampus ($n = 79$), striatum ($n = 55$) and prefrontal cortex ($n = 54$). In a *cis*-eQTL scan that was limited to the region flanking the gene being interrogated (**Supplementary Figs. 19** and **20**), we identified a total of 6,045 associations for 4,174 genes (**Fig. 4a**, **Supplementary Fig. 21** and **Supplementary Table 3**) at a permutation-derived significance threshold of $P < 0.05$ (this threshold reflects a per-gene, per-brain region). For 534 of these genes, we identified a *cis*-eQTL in all three tissues. For an additional 803 genes, we identified a *cis*-eQTL in two of the three tissues (**Fig. 4b**). The RNA-seq data were generated from a set of partially overlapping individuals; therefore, we did not perform a joint analysis of the three brain tissues[39].

In addition, we searched for *cis*-eQTLs by examining allele-specific expression (ASE), which measures the relative expression of the two possible RNA alleles derived from a heterozygous SNP[40,41]. We identified 655 genes with ASE in at least one of the three tissues. Of these, 380 (58%) were found only using ASE analysis and 275 (42%) were also identified in the conventional *cis*-eQTL scan, suggesting that there was more overlap than would be expected by chance. Overlap was likely limited by several factors, including type I errors in the ASE analysis and type II errors in both the ASE analysis and conventional *cis*-eQTL mapping.

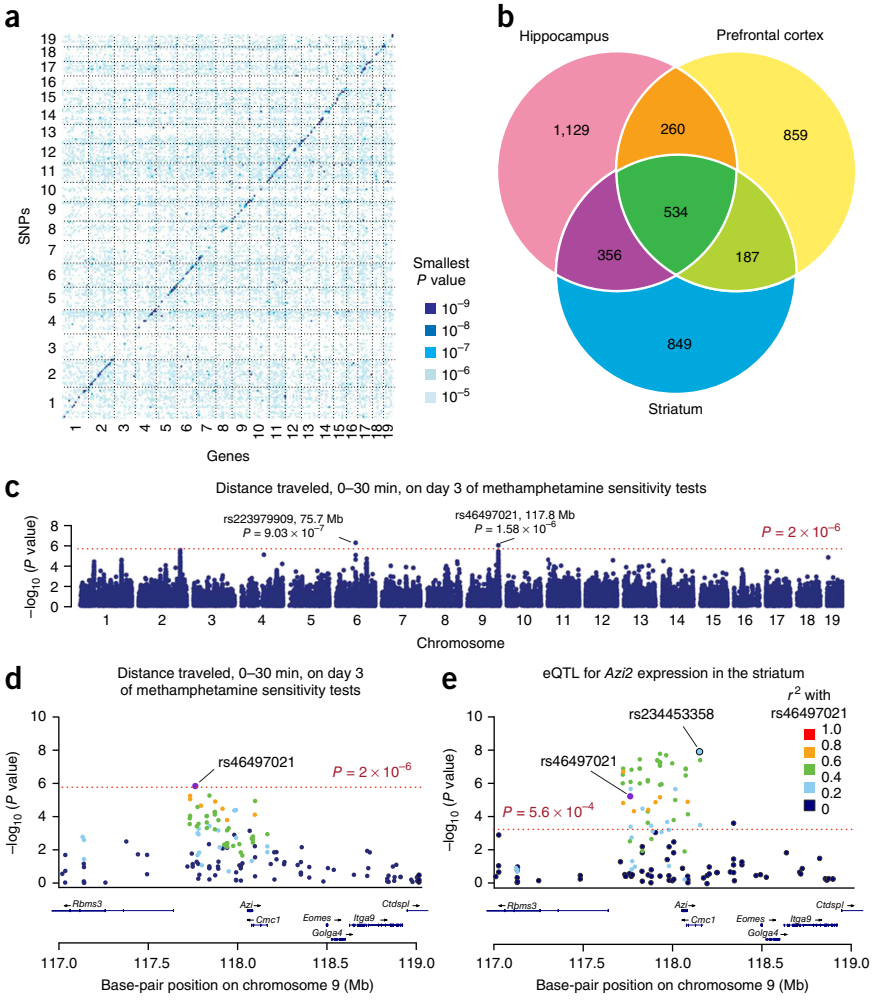

Figure 4 Overview of eQTL mapping. (**a**) The color of each pixel in the matrix corresponds to the lowest *P* value among all eQTLs using a 10 Mb × 10 Mb window. (**b**) Overlap of genes with eQTLs in the three brain tissues, detected using the traditional *cis*-eQTL mapping method (not ASE). The permutation-based *P*-value threshold for each eQTL is 0.05. (**c**) Genome-wide scan for total locomotor activity on day 3 of the methamphetamine sensitivity tests. (**d**) Association signal for total locomotor activity in the QTL region on chromosome 9. (**e**) Association signal for expression of *Azi2* in the striatum, in the same region as in **d**. Dotted red lines correspond to significance thresholds ($P < 0.1$) estimated via permutation tests.

We also mapped eQTLs across the genome for each gene in an effort to detect *trans*-eQTLs. We identified 2,278 *trans*-eQTLs that were significant ($P < 0.05$, permutation-based threshold) after testing 43,414 transcripts across the three brain regions. We expected almost that many tests to be positive under the null hypothesis. Consistent with this expectation, a quantile–quantile plot of these results suggested that only a small number of the results were true positives (**Supplementary Fig. 22**). As expected, most true positive results seem to be from the hippocampus, which had the largest sample size ($n = 79$).

Integration of behavioral QTLs with eQTLs

On the basis of evidence from human GWAS, we anticipated that heritable gene expression polymorphisms (eQTLs) would be responsible for most of the observed behavioral associations. Therefore, we tried to identify eQTLs that co-mapped with behavioral QTLs, under the assumption that the eQTL might be the molecular cause of the behavioral QTL. For example, we observed an association between methamphetamine sensitivity and rs46497021 on chromosome 9 ($P = 1.6 \times 10^{-6}$; **Fig. 4c** and **Supplementary Fig. 18**). The implicated region is small (<1 Mb) and contains only two genes: *Cmc1* and *Azi2* (**Fig. 4d**). We identified *cis*-eQTLs for both genes in the striatum, which is the tissue that is most relevant for methamphetamine sensitivity. However, rs46497021 was most strongly correlated with *Azi2* expression ($P = 1.2 \times 10^{-8}$; **Fig. 4e**). In addition, the patterns of SNPs associated with methamphetamine sensitivity and *Azi2* expression showed obvious overlap. Therefore, although both *Cmc1* and *Azi2* are credible positional candidates, the eQTL data suggest that *Azi2* is most likely to be the causative gene. To our knowledge, neither gene has previously been implicated in dopaminergic/striatal processes, suggesting that this observation may offer new insights into the biology of this drug-abuse-relevant trait.

Additionally, we identified an association between anxiety-like behavior and rs238465220 on chromosome 13 ($P = 7.3 \times 10^{-8}$; **Supplementary Fig. 17**). The implicated region spans ~1.5 Mb and contains four genes: *Chrm3*, *Larp4b*, *Dip2c* and *Zmynd11*. Among these genes, rs238465220 was also associated with expression of *Zmynd11* in the hippocampus, suggesting that this locus may influence anxiety-like behavior through regulation of *Zmynd11* expression (**Supplementary Fig. 23**). *Zmydn11* has not previously been implicated in anxiety; however, copy number variants in *ZMYND11* were recently shown to be associated with autistic tendencies and aggressive behaviors in humans[42]. These examples illustrate the usefulness of combining GWAS with eQTL data to identify the molecular mechanism by which a chromosomal region influences a complex trait.

DISCUSSION

We performed a GWAS in a commercially available outbred mouse population, which identified numerous physiological, behavioral and expression QTLs. In several cases, the implicated loci were smaller than 1 Mb and contained just a handful of genes, including an obvious candidate. In addition, we used the eQTL results to further parse between the genes in the intervals that were implicated for the behavioral traits.

The goal of using CFW mice was to enhance our mapping resolution. CFW mice have shorter-range LD than other commercially available populations[9]. Using GBS genotypes, we estimated LD in CFW mice and compared it to that of other mapping populations (**Fig. 2b**). The F_{34} generation of the LG/J × SM/J advanced intercross line (LG × SM-AIL) that we have used in previous studies[43–46] showed more extensive LD than CFW mice. Various outbred heterogeneous stocks, typically made up of eight inbred strains, have also been used in previous mapping efforts[47–49]. We examined one heterogeneous stock[49] and found that it also had longer-range LD than CFW mice. The Hybrid Mouse Diversity Panel (HMDP)[50,51], which is a collection

of approximately 100 inbred mouse strains that has been used for QTL mapping, also showed greater LD than CFW mice, as did a smaller panel of 30 inbred strains[52]. DO mice[12,18,19,53] exhibited LD decay that was almost as degraded as that of CFW mice. Populations like the AIL and heterogeneous stocks (including the DO population) are expected to show decreased LD in the future because of the accumulation of additional recombinations (for example, the LG × SM-AIL is now at generation 62). The MF-1 population is another commercially available outbred population that has been used to map QTLs[50,54], but we were unable to obtain the data needed to estimate LD decay in this population. Comparing LD patterns in different populations is a common method for estimating mapping resolution[16]; however, additional factors including the allele frequency distribution[55], population structure[56], error rates and the number, effect size and frequency of causal variants all influence power and mapping resolution. Despite these limitations, our comparison of LD (**Fig. 2b**) and our mapping results (**Figs. 3** and **4**, **Supplementary Figs. 14**–**18** and **21**–**23**, and **Supplementary Table 2**) suggest that CFW mice are an attractive option for fine-mapping studies.

Another important parameter for GWAS is allele frequency, as power to identify associations increases with greater MAF. Laboratory mouse populations have higher average MAF than humans or wild mouse populations[17]. We found that 73% of SNPs genotyped in this study had MAF >0.05, although our SNP filtering steps may have underestimated the number of rare SNPs. Populations produced by crossing inbred strains, such as F_2 crosses, recombinant inbred lines, AILs and heterogeneous stocks, typically have even more desirable MAF distributions[43]. Because the ascertainment of SNPs included on genotyping platforms directly influences the estimated MAF distribution, we did not attempt to use publicly available data to compare MAFs in commonly used mapping populations.

We found that CFW mice lacked genetic variability in certain regions; for example, chromosome 16 had a low density of polymorphic markers, as measured using both GBS and MegaMUGA (**Fig. 2a**), and no significant QTLs (**Fig. 3a**). This region is an example of a previously described tendency for laboratory mouse populations to harbor regions that are identical by descent[43,57].

Several other advantages of CFW mice include their commercial availability, their low cost and the ability to acquire non-siblings upon request. We also found that the CFW mice were easy to handle, and their uniform coat color simplified automated scoring of certain behavioral traits.

One barrier to more widespread adoption of GWAS in mice has been the lack of universal and economical SNP genotyping platforms. In this report, we demonstrate the use of GBS to overcome this obstacle. GBS is a reduced-representation sequencing approach in which a small fraction of the genome is sequenced at moderate depth to obtain genotypes at a subset of markers. Although GBS shares some characteristics with low-coverage whole-genome sequencing[58–60], GBS yields high coverage for a subset of the genome, thus acquiring information about fewer SNPs but with greater confidence. Our GBS methods included a custom-designed library preparation protocol (which reduced per-sample costs) and used the standard software toolkits GATK[61] and IMPUTE2 (ref. 62). An advantage of GBS was that it did not require preselection of polymorphic SNPs. We chose conservative criteria for SNP calling, which yielded 92,734 SNPs, of which 14% were newly discovered and possibly unique to the CFW population. These 92,734 SNPs provided extensive coverage of the genome (**Fig. 2a**) and allowed for fine-mapping (**Figs. 3c**,**d** and **4d**,**e**, and **Supplementary Figs. 17** and **23**). The number of markers obtained using GBS can be titrated by varying the restriction enzymes used, the fragment sizes selected and the degree of sample multiplexing. GBS involves imputation to correct errors and to populate missing genotypes, requiring more expertise than analysis of SNP genotyping arrays. In comparison to conventional array-based SNP genotyping, GBS had a higher error rate, which is expected to modestly decrease power but should not produce false positive QTLs, as errors will not be correlated with the traits. We are currently improving genotype imputation methods for populations in which the founder haplotypes are known, such as the AIL and heterozygous stock populations[12,45,46,63,64]. Because the monetary advantage of GBS over array-based genotyping will continue to increase as sequencing prices decrease, we anticipate that GBS and other sequencing-based approaches will supplant array-based methods in the coming years.

The majority of human GWAS findings implicate regulatory rather than coding differences[4,11]. The identified haplotypes frequently contain several genes. It is now widely appreciated that, even when an association can be localized to a single gene, that gene may not be the cause of the association[65], meaning that proximity to the peak SNP is not sufficient to identify the causal gene. eQTLs can provide the crucial link between a region implicated by GWAS and the biological processes that underlie the association. Therefore, a major goal of our study was to integrate behavioral QTL and eQTL data. We used RNA-seq to examine gene expression in three brain regions that are known to be important for the behavioral traits that we studied. Although *Azi2* was not an obvious candidate for the behavioral QTL for methamphetamine sensitivity, our data showing the co-mapping of an eQTL for *Azi2* expression in the striatum provide an additional layer of evidence. Similarly, *Zmynd11* has not previously been implicated in anxiety-like behavior, but the eQTL for *Zmynd11* expression in the hippocampus suggests that it is the most promising of the four genes within the behavioral QTL. These examples demonstrate the power of integrating fine-mapping of behavioral QTLs and eQTLs, and extend on multiple previous mouse studies that have used similar approaches in conjunction with F_2 crosses[66], recombinant inbred lines[67–69], selected lines[70], heterogeneous stocks[71], outbred MF-1 mice[50] and the HMDP[51,72,73].

RNA-seq offers a number of advantages relative to array-based gene expression measurements[74–80]. In particular, we were able to map *cis*- and *trans*-eQTLs using a traditional mapping approach and simultaneously map *cis*-eQTLs by quantifying ASE. Because only a fraction of genes can be studied using ASE analysis, we did not anticipate complete overlap between genes identified using these two approaches. Using ASE, we identified 655 *cis*-eQTLs, of which 42% were also identified as *cis*-eQTLs using conventional mapping.

We found that physiological traits typically had slightly higher heritabilities than behavioral traits (**Supplementary Table 1**). We also found that the effect sizes of individual associations tended to be higher for physiological traits (**Supplementary Table 2**), consistent with findings from another recent study in rats[81]. However, it was not always true that traits with the highest heritabilities also showed the largest effect sizes for individual associations. Because the effect size of individual QTL alleles is of paramount importance for assessing power at a given sample size and because this parameter is never known in advance, it is not possible to provide general guidelines about the sample size needed for future studies. On the basis of our results, we suggest that a sample size of 1,000 or more CFW mice should be used for most traits, although traits like testis weight and abnormal BMD would have yielded significant results with just a few hundred mice. Although our use of CFW mice was intended to increase mapping precision, there is a direct tradeoff between mapping precision and statistical power[6]; therefore, sample sizes required

for studies using CFW mice will necessarily be larger than for those using F_2 crosses, recombinant inbred lines or other traditional mapping populations that offer less precision.

Our data do not directly address the reasons that the effect sizes we observed are so much larger than the effect sizes observed in most human GWAS. We can speculate that the unique population history of laboratory mice (involving domestication, selection and repeated population bottlenecks) has increased the frequency of alleles that may have been rare in ancestral wild mouse populations. It is also true that, unlike many traits studied in human GWAS, the traits we are examining are not disease traits (and thus may not influence fitness); the underlying alleles may therefore not have been influenced by natural selection, even among ancestral wild mouse populations from which laboratory populations were originally derived. Furthermore, laboratory mice are drawn from a much more uniform environment, potentially diminishing gene-by-environment interactions that may reduce effect sizes in human GWAS. Finally, because LD in the CFW population is more extensive than in humans, we are effectively testing fewer hypotheses and therefore applied a lower (permutation-derived) significance threshold.

We have shown that use of CFW mice in conjunction with GBS and RNA-seq provides a powerful and efficient means for identifying genetic associations and for nominating candidate genes within the associated regions. In comparison to other outbred mouse populations, CFW mice showed rapid decay of LD (**Fig. 2b**), were less expensive and primarily allowed examination of *M. m. domesticus*–derived alleles (**Fig. 2c**). In comparison to human GWAS, this approach provided dramatically reduced costs, the ability to examine phenotypes that include experimental manipulations that would be impractical or unethical in humans, the ability to obtain tissue samples for expression analysis and the ability to exert exquisite control over environmental variables. Identified genes can be manipulated in future studies via genome engineering[82]. Thus, our approach can be used to rapidly generate specific and testable hypotheses for a wide array of complex traits. More broadly, our results demonstrate methods and principles that apply to a variety of other model systems.

URLs. Phenotype, genotype and RNA-seq gene expression data, http://dx.doi.org/10.5061/dryad.2rs41; R code implementing data analyses, http://github.com/pcarbo/cfw; Quanto program, http://biostats.usc.edu/software.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

The authors wish to acknowledge technical assistance from: D. Godfrey, S. Lionikaite, V. Lionikaite, A.S. Lionikiene and J. Zekos as well as technical and intellectual input from M. Abney, J. Borevitz, K. Broman, N. Cai, R. Cheng, N. Cox, R. Davies, J. Flint, L. Goodstadt, P. Grabowski, B. Harr, E. Leffler, R. Mott, J. Nicod, J. Novembre, A. Price, M. Stephens, D. Weeks and X. Zhou. This project was funded by NIH R01GM097737 and P50DA037844 (A.A.P.), NIH T32DA07255 (C.C.P.), NIH T32GM07197 (N.M.G.), NIH R01AR056280 (D.A.B.), NIH R01AR060234 (C.L.A.-B.), the Fellowship from the Human Frontiers Science Program (P.C.) and the Howard Hughes Medical Institute (J.K.P.).

AUTHOR CONTRIBUTIONS

A.A.P. conceived the study. C.C.P. and A.A.P. supervised the project. S.G. and P.C. designed and implemented the statistical and bioinformatics analyses with contributions from C.C.P., J.K.P. and A.A.P. N.M.G. designed and executed the RNA-seq and GBS protocols with assistance from E.A. and J.D. C.C.P. performed the behavioral phenotyping with assistance from E.L. and Y.J.P. A.L. performed the muscle and bone phenotyping with input from D.A.B. C.L.A.-B. performed the BMD phenotyping. C.C.P., S.G., P.C. and A.A.P. wrote the manuscript, with input from all co-authors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Manolio, T.A., Brooks, L.D. & Collins, F.S. A HapMap harvest of insights into the genetics of common disease. *J. Clin. Invest.* **118**, 1590–1605 (2008).
2. Manolio, T.A. Bringing genome-wide association findings into clinical use. *Nat. Rev. Genet.* **14**, 549–558 (2013).
3. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
4. Albert, F.W. & Kruglyak, L. The role of regulatory variation in complex traits and disease. *Nat. Rev. Genet.* **16**, 197–212 (2015).
5. Mott, R. & Flint, J. Dissecting quantitative traits in mice. *Annu. Rev. Genomics Hum. Genet.* **14**, 421–439 (2013).
6. Parker, C.C. & Palmer, A.A. Dark matter: are mice the solution to missing heritability? *Front. Genet.* **2**, 32 (2011).
7. Lynch, C.J. The so-called Swiss mouse. *Lab. Anim. Care* **19**, 214–220 (1969).
8. Rice, M.C. & O'Brien, S.J. Genetic variance of laboratory outbred Swiss mice. *Nature* **283**, 157–161 (1980).
9. Yalcin, B. *et al.* Commercially available outbred mice for genome-wide association studies. *PLoS Genet.* **6**, e1001085 (2010).
10. Chia, R., Achilli, F., Festing, M.F.W. & Fisher, E.M.C. The origins and uses of mouse outbred stocks. *Nat. Genet.* **37**, 1181–1186 (2005).
11. Gusev, A. *et al.* Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. *Am. J. Hum. Genet.* **95**, 535–552 (2014).
12. Gatti, D.M. *et al.* Quantitative trait locus mapping methods for diversity outbred mice. *G3 (Bethesda)* **4**, 1623–1633 (2014).
13. Morgan, A.P. *et al.* The Mouse Universal Genotyping Array: from substrains to subspecies. *G3 (Bethesda)* **6**, 263–279 (2015).
14. Yang, H. *et al.* A customized and versatile high-density genotyping array for the mouse. *Nat. Methods* **6**, 663–666 (2009).
15. Elshire, R.J. *et al.* A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* **6**, e19379 (2011).
16. Pritchard, J.K. & Przeworski, M. Linkage disequilibrium in humans: models and data. *Am. J. Hum. Genet.* **69**, 1–14 (2001).
17. Laurie, C.C. *et al.* Linkage disequilibrium in wild mice. *PLoS Genet.* **3**, e144 (2007).
18. Chesler, E.J. Out of the bottleneck: the Diversity Outcross and Collaborative Cross mouse populations in behavioral genetics research. *Mamm. Genome* **25**, 3–11 (2014).
19. Churchill, G.A., Gatti, D.M., Munger, S.C. & Svenson, K.L. The Diversity Outbred mouse population. *Mamm. Genome* **23**, 713–718 (2012).
20. Collaborative Cross Consortium. The genome architecture of the Collaborative Cross mouse genetic reference population. *Genetics* **190**, 389–401 (2012).
21. Lee, S.H. *et al.* Estimation of SNP heritability from dense genotype data. *Am. J. Hum. Genet.* **93**, 1151–1155 (2013).
22. Wray, N.R. *et al.* Pitfalls of predicting complex traits from SNPs. *Nat. Rev. Genet.* **14**, 507–515 (2013).
23. Cheng, R. & Palmer, A.A. A simulation study of permutation, bootstrap, and gene dropping for assessing statistical significance in the case of unequal relatedness. *Genetics* **193**, 1015–1018 (2013).
24. Churchill, G.A. & Doerge, R.W. Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963–971 (1994).
25. Mendis, S.H.S., Meachem, S.J., Sarraj, M.A. & Loveland, K.L. Activin A balances Sertoli and germ cell proliferation in the fetal mouse testis. *Biol. Reprod.* **84**, 379–391 (2011).
26. Mithraprabhu, S. *et al.* Activin bioactivity affects germ cell differentiation in the postnatal mouse testis *in vivo*. *Biol. Reprod.* **82**, 980–990 (2010).
27. Tomaszewski, J., Joseph, A., Archambeault, D. & Yao, H.H.-C. Essential roles of inhibin βA in mouse epididymal coiling. *Proc. Natl. Acad. Sci. USA* **104**, 11322–11327 (2007).
28. Lee, S.-J. Quadrupling muscle mass in mice by targeting TGF-β signaling pathways. *PLoS One* **2**, e789 (2007).
29. Lee, S.-J. *et al.* Regulation of muscle mass by follistatin and activins. *Mol. Endocrinol.* **24**, 1998–2008 (2010).
30. Lionikas, A. *et al.* Resolving candidate genes of mouse skeletal muscle QTL via RNA-Seq and expression network analyses. *BMC Genomics* **13**, 592 (2012).
31. Sala, D. *et al.* Autophagy-regulating TP53INP2 mediates muscle wasting and is repressed in diabetes. *J. Clin. Invest.* **124**, 1914–1927 (2014).
32. Estrada, K. *et al.* Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat. Genet.* **44**, 491–501 (2012).

33. Zheng, H.-F. *et al.* Whole-genome sequencing identifies *EN1* as a determinant of bone density and fracture. *Nature* **526**, 112–117 (2015).
34. Coury, F. *et al.* SLC4A2-mediated $Cl^-/HCO3^-$ exchange activity is essential for calpain-dependent regulation of the actin cytoskeleton in osteoclasts. *Proc. Natl. Acad. Sci. USA* **110**, 2163–2168 (2013).
35. Meyers, S.N. *et al.* A deletion mutation in bovine *SLC4A2* is associated with osteopetrosis in Red Angus cattle. *BMC Genomics* **11**, 337 (2010).
36. Sillence, D.O., Senn, A. & Danks, D.M. Genetic heterogeneity in osteogenesis imperfecta. *J. Med. Genet.* **16**, 101–116 (1979).
37. Sykes, B., Wordsworth, P., Ogilvie, D., Anderson, J. & Jones, N. Osteogenesis imperfecta is linked to both type I collagen structural genes. *Lancet* **2**, 69–72 (1986).
38. Long, J.-R. *et al.* Association between *COL1A1* gene polymorphisms and bone size in Caucasians. *Eur. J. Hum. Genet.* **12**, 383–388 (2004).
39. Flutre, T., Wen, X., Pritchard, J. & Stephens, M. A statistical framework for joint eQTL analysis in multiple tissues. *PLoS Genet.* **9**, e1003486 (2013).
40. Serre, D. *et al.* Differential allelic expression in the human genome: a robust approach to identify genetic and epigenetic *cis*-acting mechanisms regulating gene expression. *PLoS Genet.* **4**, e1000006 (2008).
41. Pickrell, J.K. *et al.* Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature* **464**, 768–772 (2010).
42. Coe, B.P. *et al.* Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat. Genet.* **46**, 1063–1071 (2014).
43. Cheng, R. *et al.* Genome-wide association studies and the problem of relatedness among advanced intercross lines and other highly recombinant populations. *Genetics* **185**, 1033–1044 (2010).
44. Samocha, K.E., Lim, J.E., Cheng, R., Sokoloff, G. & Palmer, A.A. Fine mapping of QTL for prepulse inhibition in LG/J and SM/J mice using F_2 and advanced intercross lines. *Genes Brain Behav.* **9**, 759–767 (2010).
45. Parker, C.C. *et al.* Fine-mapping alleles for body weight in LG/J × SM/J F and F_{34} advanced intercross lines. *Mamm. Genome* **22**, 563–571 (2011).
46. Parker, C.C. *et al.* High-resolution genetic mapping of complex traits from a combined analysis of F_2 and advanced intercross mice. *Genetics* **198**, 103–116 (2014).
47. Talbot, C.J. *et al.* High-resolution mapping of quantitative trait loci in outbred mice. *Nat. Genet.* **21**, 305–308 (1999).
48. 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).
49. Valdar, W. *et al.* Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nat. Genet.* **38**, 879–887 (2006).
50. Ghazalpour, A. *et al.* High-resolution mapping of gene expression using association in an outbred mouse stock. *PLoS Genet.* **4**, e1000149 (2008).
51. Orozco, L.D. *et al.* Unraveling inflammatory responses using systems genetics and gene–environment interactions in macrophages. *Cell* **151**, 658–670 (2012).
52. Sittig, L.J., Carbonetto, P., Engel, K.A., Krauss, K.S. & Palmer, A.A. Integration of genome-wide association and extant brain expression QTL identifies candidate genes influencing prepulse inhibition in inbred F_1 mice. *Genes Brain Behav.* **15**, 260–270 (2016).
53. Svenson, K.L. *et al.* High-resolution genetic mapping using the Mouse Diversity outbred population. *Genetics* **190**, 437–447 (2012).
54. 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).
55. Eberle, M.A., Rieder, M.J., Kruglyak, L. & Nickerson, D.A. Allele frequency matching between SNPs reveals an excess of linkage disequilibrium in genic regions of the human genome. *PLoS Genet.* **2**, e142 (2006).
56. Mangin, B. *et al.* Novel measures of linkage disequilibrium that correct the bias due to population structure and relatedness. *Heredity* **108**, 285–291 (2012).
57. Yang, H. *et al.* Subspecific origin and haplotype diversity in the laboratory mouse. *Nat. Genet.* **43**, 648–655 (2011).
58. CONVERGE Consortium. Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* **523**, 588–591 (2015).
59. Le, S.Q. & Durbin, R. SNP detection and genotyping from low-coverage sequencing data on multiple diploid samples. *Genome Res.* **21**, 952–960 (2011).
60. Li, Y., Sidore, C., Kang, H.M., Boehnke, M. & Abecasis, G.R. Low-coverage sequencing: implications for design of complex trait association studies. *Genome Res.* **21**, 940–951 (2011).
61. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
62. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
63. Parker, C.C., Sokoloff, G., Cheng, R. & Palmer, A.A. Genome-wide association for fear conditioning in an advanced intercross mouse line. *Behav. Genet.* **42**, 437–448 (2012).
64. Parker, C.C., Cheng, R., Sokoloff, G. & Palmer, A.A. Genome-wide association for methamphetamine sensitivity in an advanced intercross mouse line. *Genes Brain Behav.* **11**, 52–61 (2012).
65. Smemo, S. *et al.* Obesity-associated variants within *FTO* form long-range functional connections with *IRX3*. *Nature* **507**, 371–375 (2014).
66. Schadt, E.E. *et al.* Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422**, 297–302 (2003).
67. Chesler, E.J., Lu, L., Wang, J., Williams, R.W. & Manly, K.F. WebQTL: rapid exploratory analysis of gene expression and genetic networks for brain and behavior. *Nat. Neurosci.* **7**, 485–486 (2004).
68. 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).
69. Bystrykh, L. *et al.* Uncovering regulatory pathways that affect hematopoietic stem cell function using 'genetical genomics'. *Nat. Genet.* **37**, 225–232 (2005).
70. Palmer, A.A. *et al.* Gene expression differences in mice divergently selected for methamphetamine sensitivity. *Mamm. Genome* **16**, 291–305 (2005).
71. Huang, G.-J. *et al.* High resolution mapping of expression QTLs in heterogeneous stock mice in multiple tissues. *Genome Res.* **19**, 1133–1140 (2009).
72. Farber, C.R. *et al.* Mouse genome-wide association and systems genetics identify *Asxl2* as a regulator of bone mineral density and osteoclastogenesis. *PLoS Genet.* **7**, e1002038 (2011).
73. Calabrese, G. *et al.* Systems genetic analysis of osteoblast-lineage cells. *PLoS Genet.* **8**, e1003150 (2012).
74. de Klerk, E. & 't Hoen, P.A.C. Alternative mRNA transcription, processing, and translation: insights from RNA sequencing. *Trends Genet.* **31**, 128–139 (2015).
75. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).
76. Mane, S.P. *et al.* Transcriptome sequencing of the Microarray Quality Control (MAQC) RNA reference samples using next generation sequencing. *BMC Genomics* **10**, 264 (2009).
77. Tang, F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* **6**, 377–382 (2009).
78. Trapnell, C., Pachter, L. & Salzberg, S.L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
79. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).
80. Walter, N.A. *et al.* High throughput sequencing in mice: a platform comparison identifies a preponderance of cryptic SNPs. *BMC Genomics* **10**, 379 (2009).
81. Baud, A. *et al.* Combined sequence-based and genetic mapping analysis of complex traits in outbred rats. *Nat. Genet.* **45**, 767–775 (2013).
82. Sander, J.D. & Joung, J.K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* **32**, 347–355 (2014).
83. Bennett, B.J. *et al.* A high-resolution association mapping panel for the dissection of complex traits in mice. *Genome Res.* **20**, 281–290 (2010).

ONLINE METHODS

Animal models. We phenotyped 1,200 male CFW mice (*Mus musculus*) that were obtained from the Charles River Laboratories facility in Portage, Michigan, USA (strain code, CRL:CFW(SW); facility code, P08). We performed a power analysis using the program Quanto (see URLs). This analysis indicated that 1,200 mice would provide 80% power to detect QTLs that accounted for ~3% of total trait variance with $P < 5 \times 10^{-7}$. Since our study was completed, the Portage colony has been relocated to Kingston, New York, USA (new code K92). It has been reported that the ancestors of the CFW mice were obtained from a large colony of Swiss mice in 1926 and maintained by Leslie Webster at the Rockefeller Institute. A single pair of highly inbred albino mice was later acquired by Carworth Farms and used to initiate an outbred mouse stock. Several mice from this colony were later acquired in 1974 by Charles River Laboratories and were subsequently maintained as an outbred population[8–10].

Every 2 weeks, 48 male CFW mice were shipped from Charles River Laboratories in Portage to our laboratory in Chicago. We requested that Charles River Laboratories send only one mouse from each litter to avoid obtaining siblings, as the use of close relatives reduces power to map QTLs and complicates analysis. The average age of the mice upon arrival in our laboratories was 35 d (ranging from 34 to 46 d), and the average weight of the mice was 25.5 g (ranging from 13.4 to 38.7 g). Mice were housed four per cage and given ~15 d to adapt to their new environment (**Supplementary Fig. 1**). Standard laboratory chow and water were available *ad libitum*, except during the behavioral procedures and before testing for fasting glucose levels. Mice were maintained on a standard 12 h light/12 h dark cycle (lights on at 6:30 a.m.). All phenotyping occurred during the light phase between 8:00 a.m. and 4:00 p.m., over the period of August 2011 to December 2012. All procedures were approved by the University of Chicago Institutional Animal Care and Use Committee (IACUC) in accordance with National Institutes of Health guidelines for the care and use of laboratory animals.

Phenotyping. The order of phenotyping was identical for each mouse and is shown schematically in **Supplementary Figure 1**. One day after arrival, mice were fasted for 4 h before measurement of blood glucose levels. Fourteen days later, we assessed the response to a new environment and to administration of 1.5 mg/kg methamphetamine in a 3-d paradigm[64]. Twelve days later, we tested mice for conditioned fear[46]. Nine days after that, we tested mice for PPI[44]. Finally, after 15 d, we weighed and sacrificed the mice. Immediately after sacrifice, we weighed testes and collected one leg for measurement of muscle phenotypes and the other leg for measurement of bone phenotypes. We also measured tail length at this time (**Supplementary Note**).

RNA sequencing. We collected brain tissue from a subset of the mice as a source of mRNA from the hippocampus (n = 79), striatum (n = 55) and frontal cortex (n = 54). We used RNA-seq[84,85] to quantify gene transcript abundance in these brain tissues. Library preparation was performed with the TruSeq RNA Sample kit (Illumina). Samples were multiplexed five per lane and sequenced on an Illumina HiSeq 2000 sequencer, using single-end 100-bp reads. We processed the RNA-seq short reads using the Tuxedo software suite[79]: (i) first, we aligned the short reads to the reference genome assembly (NCBI release 38, mm10) with Bowtie 2 (ref. 86); (ii) next, we used TopHat2 (ref. 79) to align the short reads to known splice junctions; and (iii) finally, we used Cufflinks[87] to calculate, for each gene, a gene-level measure of expression based on the mapped reads. This measure is reported in RPKM. This measure does not depend on the length of the coding sequences or the sequencing depth of each sample (so mapping eQTLs will not be biased by these factors). We focused on this gene-level measurement for subsequent investigation, including eQTL mapping and assessment of ASE. See the **Supplementary Note** for further details.

Genotyping by sequencing. GBS is a reduced-representation genotyping method for obtaining genotyping information by sequencing only regions that are proximal to a restriction enzyme cut site[15]. Our protocol was adapted from previously described procedures[88]. GBS libraries were prepared by digesting genomic DNA with a restriction enzyme, PstI, and annealing oligonucleotide adaptors to the resulting overhangs. Samples were multiplexed 12 per lane and sequenced on an Illumina HiSeq 2000 sequencer, using single-end 100-bp reads. We obtained an average of 4.8 million reads per sample. By focusing the sequencing effort on PstI restriction sites, we obtained high coverage (~15×; **Supplementary Fig. 24**) at a subset of genomic loci, although the reads were very non-uniformly distributed. We aligned the 100-bp single-end reads to Mouse Reference Assembly 38 from the NCBI database (mm10) using BWA[89]. We used GATK[61,90] to discover variants and to obtain genotype probabilities. For the VQSR step, we calibrated variant discovery against (i) whole-genome sequencing data that we ascertained from a small set of CFW mice and (ii) SNPs and indels from the Wellcome Trust Sanger Institute Mouse Genomes Project[91] and SNPs available in dbSNP release 137. We used IMPUTE2 (ref. 62) to improve low-confidence genotypes or genotypes that were not called in individual mice. **Supplementary Table 4** and the **Supplementary Note** detail our efforts to estimate the error rate of GBS in this study. The **Supplementary Note** also contains a description of a small number of SNPs that were discarded because a large proportion of the genotypes were imputed with low certainty. Finally, the **Supplementary Note** details our identification of 110 DNA samples that seemed to be mislabeled and were therefore excluded from our study (**Supplementary Figs. 25–28**).

Treemix analysis. We estimated phylogenetic relationships between CFW mice and different laboratory strains sequenced as part of the Wellcome Trust Sanger Institute Mouse Genomes Project using Treemix[92]. We used the genotypes for the laboratory strains sequenced by the Wellcome Trust to obtain the locations of SNPs that were identified in the CFW mice using our GBS pipeline. We excluded the *Mus spretus* strain from the Wellcome Trust data, as this strain was included as an outgroup. Because the laboratory strains were all inbred, we assumed that allele frequency was 1 or 0. We represented each strain by only a single individual. We used a subset of 100 CFW mice to compute allele frequencies from the genotype likelihoods of GBS SNPs in our sample. Treemix was used to fit a maximum-likelihood tree to all the laboratory strains and CFW samples.

QTL mapping for behavioral and physiological traits. We performed a GWAS for the behavioral and physiological phenotypes using all SNPs with MAF >2% and good imputation quality (defined as 95% of the samples having a maximum-probability genotype greater than 0.5). Although our analyses did not suggest the presence of close relatives or population structure, we used the LMM implemented in the program GEMMA[93]. GEMMA is similar to a standard linear regression, in which a quantitative trait (y) is modeled as a linear combination of genotype (x) and covariates (z), except that it includes an additional 'random' or 'polygenic' effect capturing the covariance structure in the phenotype that is attributed to genome-wide genetic sharing.

$$y_i = \mu + z_{i1}\alpha_1 + \cdots + z_{im}\alpha_m + x_{ij}\beta_j + u_i + \varepsilon_i$$

The notation in this expression is defined as follows: y_i is the ith phenotype sample; z_{ik} is ith sample of covariate k, in which k ranges from 1 to the number of covariates included in the regression (m); α_k is the coefficient corresponding to covariate k; x_{ij} is the genotype of sample i at SNP j; β_j is the coefficient corresponding to SNP j; u_i is the polygenic effect for the ith sample; ε_i is the residual error; and μ is the intercept. The genotype, x_{ij}, is represented as the expected allele count, in which 0 represents homozygosity for the major allele and 2 represents homozygosity for the minor allele, and β_j is the additive effect of the expected allele count on the phenotype. The residuals ε_i are assumed to be *i.i.d.* normal with mean zero and covariance σ^2, whereas the polygenic effect $u = (u_1, \ldots, u_n)^T$ is a random vector drawn from the multivariate normal distribution with mean zero and $n \times n$ covariance matrix $\sigma^2\lambda K$, where n is the number of samples.

We estimated the relatedness matrix K from the genotype data. We specified the covariance matrix using the realized relationship matrix $K = XX^T/p$, where p is the number of SNPs and X is the $n \times p$ genotype matrix with entries x_{ij}. This formulation was derived from a polygenic model of the phenotype in which all SNPs helped explain variance in the phenotype and the contributions of individual SNPs were *i.i.d.* normal[94–96].

The inclusion of a genetic marker in both the fixed and random terms can deflate the test statistic for this marker, leading to a loss of power to detect

a QTL; this problem has been termed 'proximal contamination' (ref. 95). To avoid proximal contamination, we computed 19 different K matrices, each one excluding 1 of the 19 autosomes. To scan markers on a given chromosome, we used the version of K that did not include that chromosome. We have previously proposed this leave-one-chromosome-out (LOCO) approach as a simple solution for avoiding the problem of proximal contamination[23].

We used a permutation-based approach to calculate the genome-wide significance threshold for P values calculated in GEMMA. We estimated the distribution of P values under the null hypothesis by mapping QTLs in 1,000 randomly permuted data sets, then taking the threshold to be the $100(1-\alpha)$th percentile of this distribution, with $\alpha = 0.1$. Although this permutation test is technically only valid under the assumption that the samples are exchangeable[97], we have previously suggested that 'naive' permutations are generally sufficient[23]. Furthermore, given our observation that population structure is subtle, we expect that this simulation provides a good approximation to the null (**Supplementary Note**).

Heritability estimates. Instead of computing a point estimate for SNP heritability (h^2), which is the usual approach (for example, using the REML estimate[98]), we evaluated likelihood over a regular grid of values for h^2, which allowed us to directly quantify uncertainty in h^2 under the reasonable assumption of a uniform prior for the proportion of variance explained[96].

We estimated h^2 for our phenotypes[21]. Because the GBS SNPs did not completely tag all causal variants (and because we excluded the sex chromosomes), our estimates of h^2 underestimate a trait's true narrow-sense heritability. To estimate h^2, we assumed that all genetic markers made some small contribution to variation in the trait and that these contributions were normally distributed with the same variance[96,98,99]. Under this polygenic model, the covariance of the phenotype measurements was $\mathrm{Cov}(y_1, \ldots, y_n) = \sigma^2 H$, where $H = \left(I + \sigma_a^2 K\right)$, I is the $n \times n$ identity matrix, K is the $n \times n$ realized relatedness matrix, σ_a^2 is the variance of the additive genetic effects and σ^2 is the variance of the residuals. Under this formulation, σ_a^2 represents the relative contribution of the additive genetic variance, and we can use this parameter to provide an estimate for h^2

$$h^2 = \sigma_a^2 s_a / \left(\sigma_a^2 s_a + 1\right)$$

where s_a is the mean sample variance of all the available SNPs, or the mean of the diagonal entries of K assuming that the columns of X are centered so that each of the columns has a mean of zero. See the **Supplementary Note** for further details.

Expression QTL mapping. We used the RPKM measurements from RNA-seq and the GBS genotype data to map eQTLs. We performed an eQTL scan separately in each brain tissue (hippocampus, striatum and prefrontal cortex). First, we discarded genes with low levels of expression (RPKM <1) and genes that showed no variability in expression. For the remaining genes, we quantile normalized the expression data. To account for unknown confounders, we removed the linear effects of the first few principal components calculated from the $k \times n$ gene expression matrix, with k genes and n samples (20 principal components for hippocampus, 10 principal components for striatum and 20 principal components for prefrontal cortex)[41]. After removing the linear effects of the principal components, we again quantile normalized the expression data. We then used an LMM as implemented in GEMMA to scan for *cis*-eQTLs, as described above for the behavioral and physiological phenotypes. To define *cis*-eQTLs, we only considered SNPs within 1 Mb of the gene's transcribed region (preliminary analyses indicated that 1 Mb captured most of the significant signals; **Supplementary Figs. 19** and **20**). We used a permutation-based approach to calculate significance thresholds for P values in *cis*-eQTL mapping. We used 1,000 permutations of the expression values to compute a separate significance threshold for each gene, using only SNPs that were included in the *cis*-eQTL scan. In addition to *cis*-eQTL scans, we also performed genome-wide *trans*-eQTL scans for all genes. The genome-wide scans were performed using the same LMM that was used for *cis*-eQTL analyses, except that all SNPs outside a 2-Mb region flanking the gene were included in the *trans*-eQTL analysis. The significance threshold for *trans*-eQTLs was computed using permutations of 1,000 randomly selected genes in each tissue; this approach is permissible because all expression traits were quantile normalized (**Supplementary Note**).

Allele-specific expression. We performed an analysis of ASE to identify genes that had ASE QTLs. This analysis was conducted independently from the mapping of *cis*-eQTLs described above. We identified variants that had at least ten samples with high-confidence heterozygote genotype calls. For genes that contained at least one such variant, we compared the relative expression of the two alleles across these heterozygote samples. To account for overdispersion, we used a beta binomial model to fit the counts of the two alleles for each sample. We then used a likelihood-ratio test to test for significant deviation of the observed data from the expectation of equal counts for both alleles (**Supplementary Fig. 29** and **Supplementary Note**).

84. Majewski, J. & Pastinen, T. The study of eQTL variations by RNA-seq: from SNPs to phenotypes. *Trends Genet.* **27**, 72–79 (2011).
85. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63 (2009).
86. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
87. Roberts, A., Pimentel, H., Trapnell, C. & Pachter, L. Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics* **27**, 2325–2329 (2011).
88. Grabowski, P.P., Morris, G.P., Casler, M.D. & Borevitz, J.O. Population genomic variation reveals roles of history, adaptation and ploidy in switchgrass. *Mol. Ecol.* **23**, 4059–4073 (2014).
89. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
90. Van der Auwera, G.A. *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **11**, 11.10.1–11.10.33 (2013).
91. Keane, T.M. *et al.* Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477**, 289–294 (2011).
92. Pickrell, J.K. & Pritchard, J.K. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **8**, e1002967 (2012).
93. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**, 821–824 (2012).
94. Hayes, B.J., Visscher, P.M. & Goddard, M.E. Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res. (Camb.)* **91**, 47–60 (2009).
95. Listgarten, J. *et al.* Improved linear mixed models for genome-wide association studies. *Nat. Methods* **9**, 525–526 (2012).
96. Zhou, X., Carbonetto, P. & Stephens, M. Polygenic modeling with Bayesian sparse linear mixed models. *PLoS Genet.* **9**, e1003264 (2013).
97. Abney, M. Permutation testing in the presence of polygenic variation. *Genet. Epidemiol.* **39**, 249–258 (2015).
98. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
99. Speed, D., Hemani, G., Johnson, M.R. & Balding, D.J. Improved heritability estimation from genome-wide SNPs. *Am. J. Hum. Genet.* **91**, 1011–1021 (2012).

doi:10.1038/ng.3609