The Company of **Biologists**

Identifying genes for neurobehavioural traits in rodents: progress and pitfalls

Amelie Baud[1],* and Jonathan Flint[2],*

ABSTRACT

Identifying genes and pathways that contribute to differences in neurobehavioural traits is a key goal in psychiatric research. Despite considerable success in identifying quantitative trait loci (QTLs) associated with behaviour in laboratory rodents, pinpointing the causal variants and genes is more challenging. For a long time, the main obstacle was the size of QTLs, which could encompass tens if not hundreds of genes. However, recent studies have exploited mouse and rat resources that allow mapping of phenotypes to narrow intervals, encompassing only a few genes. Here, we review these studies, showcase the rodent resources they have used and highlight the insights into neurobehavioural traits provided to date. We discuss what we see as the biggest challenge in the field – translating QTLs into biological knowledge by experimentally validating and functionally characterizing candidate genes – and propose that the CRISPR/Cas genome-editing system holds the key to overcoming this obstacle. Finally, we challenge traditional views on inbred versus outbred resources in the light of recent resource and technology developments.

KEY WORDS: Genetics of behaviour, Quantitative trait loci, Rodent resources

Introduction

According to a World Health Organization (WHO) survey of 14 countries in the Americas, Europe, the Middle East, Africa and Asia (Demyttenaere et al., 2004), psychiatric disorders are relatively common worldwide, although the overall prevalence varies widely by country. Anxiety disorders are the most common disorders, followed by mood, substance and impulse-control disorders. Importantly, psychiatric disorders are typically associated with impairment greater than that due to serious chronic physical diseases (Demyttenaere et al., 2004).

Despite the high prevalence of psychiatric disorders and the important individual and societal burden, relatively little is known about the underlying biological mechanisms. As a result, diagnostic categories such as those of the widely influential Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2000) and International Statistical Classification of Diseases and Related Health Problems (World Health Organization, 2004) are defined based on symptoms (e.g. guilty ruminations,

[1]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK. [2]Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA 90095-1761, USA.

*Authors for correspondence (abaud@ebi.ac.uk; JFlint@mednet.ucla.edu)

A.B., 0000-0003-2448-0283; J.F., 0000-0002-9427-4429

suicidal thoughts, low mood) and signs (e.g. impaired cognitive function, rapid speech).

Although such classification has enabled more reproducible diagnoses and better clinical management, it does not ensure that the diagnostic categories correspond to shared underlying causes and mechanisms (Owen, 2014). For example, major depressive disorder (MDD) is diagnosed if an individual presents with five out of nine symptoms, one of which must be depressed mood or loss of pleasure. Symptomatic heterogeneity suggests that psychiatric disorders are likely to stem from various causes and mechanisms (Casey et al., 2013). The same symptoms can also apply to different diagnostic categories (Allardyce et al., 2007) so that, in addition to heterogeneity within a disorder, common aetiologies between disorders are likely.

Research has provided evidence that the match between diagnostic categories and genetic causes is relatively poor. For example, genetic heterogeneity in MDD is evident from the poor genetic correlation between sexes [0.60, which is similar to the genetic correlation between MDD and bipolar disorder (BPD) at 0.64] (Flint and Kendler, 2014). Furthermore, several studies have provided evidence for a shared genetic component across diagnostic categories, including schizophrenia (SZ) and BPD (Lichtenstein et al., 2009; Purcell et al., 2009; Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013a), SZ and MDD, BPD and MDD, attention-deficit/hyperactivity disorder (ADHD) and MDD, and SZ and autism spectrum disorders (ASD) (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013a). Genome-wide association studies (GWAS; see Glossary, Box 1) have also identified individual genetic variants that contribute to multiple psychiatric disorders (Psychiatric GWAS Consortium and Bipolar Disorder Working Group, 2011; Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013b; Sullivan et al., 2012).

A further confounding factor is that healthy individuals may experience 'hallmark' symptoms of psychiatric disorders, such as delusions and hallucinations (Allardyce et al., 2007). Similarly, quantitative traits that are extreme in affected individuals vary substantially among healthy individuals (e.g. capacity for social communication is impaired in ASD but varies in the general population, Robinson et al., 2016). Importantly, variation of the relevant traits in healthy and affected individuals arises from similar genetic variants (Robinson et al., 2016; Lencz et al., 2014), suggesting that psychiatric disease is likely to be the extreme of a continuum rather than a discrete entity.

Therefore, there has been a push to reconsider the way in which research on psychiatric disorders is carried out and to stop focusing on current clinical diagnostic categories. Championing that vision, the American National Institute of Mental Health (NIMH) launched the Research Domain Criteria (RDoC) project (Insel et al., 2010), which encourages researchers to use a wider range of, preferably quantitative, measures related to mental health, including genetic, molecular, cellular, circuit-level and individual-level measures, as

Disease Models & Mechanisms

> ## Box 1. Glossary of key terms
>
> **Complex trait:** A phenotype that varies as a result of multiple genetic and environmental effects and their interplay.
>
> **X% confidence interval (CI):** Size of the genomic region that has X% chance to contain the variant responsible for the QTL.
>
> **CRISPR/Cas:** Clustered regularly interspaced short palindromic repeats/CRISPR-associated proteins – a genetic engineering system based on components of a prokaryotic immune system that can be used to change single nucleotides or larger genomic fragments in any genome.
>
> **Founder:** An inbred strain or outbred stock from which a resource is descended.
>
> **Genetic mapping:** The process by which regions of the genome associated with phenotypic variation are identified.
>
> **Genome-wide association study (GWAS):** The genome-wide mapping of single-nucleotide polymorphisms (SNPs) associated with a particular trait across many individuals.
>
> **Haplotype:** An unrecombined genomic segment inherited from one of the founders.
>
> **Hardy–Weinberg equilibrium:** A variant is said to be in Hardy–Weinberg equilibrium when the frequencies of the corresponding genotypes are constant across generations. Unless a variant is subject to strong evolutionary forces (such as selection or meiotic drive), genotype frequencies will be approximately constant across generations at the time scales we are considering here (years or decades).
>
> **Heritability:** The proportion of phenotypic variance in a population that is attributable to genetic effects.
>
> **Inbred strain:** An inbred strain corresponds to animals that are genetically identical (clones). It is derived by sister-brother mating for many (>20 in mice) generations. In an inbred strain, all genetic loci are homozygous.
>
> **LOD support interval:** The logarithm of the odds (LOD) ratio provides a measure of association between genotype and phenotype. The LOD will peak at a QTL and drop as distance to the QTL increases. In some circumstances the boundaries of a QTL can be defined based on the LOD profile and the distance it takes for it to drop by 1 unit, which gives 90% chance that thus-delimited QTL encompasses the causal variant. A 1.5 LOD drop interval corresponds to 95% confidence that the QTL encompasses the causal variant (Dupuis and Siegmund, 1999).
>
> **Outbred stock:** Animals that are genetically diverse and unique.
>
> **Quantitative complementation:** A method for testing the candidacy of a gene at a QTL. See Flint et al. (2005) for a description of this complicated test.
>
> **Quantitative trait locus (QTL):** A locus in the genome found to be associated with variation in a quantitative phenotype, such as height, weight or a measure of anxiety (for example how much an animal freezes in response to a frightening stimulus). When the trait mapped is expression level of a gene, QTLs are called eQTLs (for expression QTLs) and classified as cis (cis-eQTL) when they are close to the gene and trans (trans-eQTL) when they are distant to it.
>
> **Recombination event:** The exchange of genetic material between two homologous chromosomes during meiosis.

well as the family environment and social context, and to study the full spectrum of these variables in affected individuals and the general population (Insel et al., 2010; Casey et al., 2013). The goal of RDoC is that the classification of psychiatric disorders will, in the future, be informed by findings from neuroscience and genomics, and match aetiological processes. Furthermore, knowledge of the biological processes underlying psychiatric disorders would allow biological and behavioural tests to be used for diagnosis, prevention and treatment.

With increasing focus on specific aspects of psychiatric disorders rather than diagnostic categories, the importance of animal models in psychiatric research is evident (Markou et al., 2009). Many measures collected in humans to assess behavioural traits can be

readily collected in rodents (e.g. neuroimaging, startle test, biochemical measurements), and measures that can only be collected in animal models for practical or ethical reasons (e.g. gene expression in the brain, response to psychosocial stress) may also contribute to a better understanding of neurobehavioural processes.

Of course, the way phenotypic variation arises will be important when considering the relevance of animal models to psychiatric research. Because most common psychiatric disorders have been shown to arise from a large number of genetic variants and are considered to be complex traits (see Glossary, Box 1), populations of mice and rats that segregate a large number of naturally occurring variants and present continuous phenotypic variation will be most relevant.

In this Review, we highlight rodent resources that have shed light on the genetic basis of behaviour and associated neurophysiological traits. More specifically, we focus on the outcome that we believe is of greatest interest to geneticists with a strong focus on disease, namely the identification of genes associated with neurobehavioural variation. Once genes have been associated with a trait, the pathways, cell types and neural circuits that are involved can be inferred, and potential therapeutic targets identified.

Because of our focus on complex behaviours and traits, we will not discuss animal models created to assess the role of one gene in isolation; instead, we will review rodent resources in which a large number of genetic variants segregate and give rise to phenotypic variation (for a review of both strategies and how they complement each other, see Williams and Auwerx, 2015).

The same resources can be used for the study of behaviours as for other complex traits. Indeed, although behaviours are 'noisier' phenotypes than average i.e. their heritability (see Glossary, Box 1) is slightly lower than that of other complex traits both in mice (Valdar et al., 2006b; Nicod et al., 2016; Parker et al., 2016) and rats (Baud et al., 2014b; Rat Genome Sequencing and Mapping Consortium, 2013), the effect sizes of quantitative trait loci (QTLs; see Glossary, Box 1) are not significantly lower (Flint and Mackay, 2009; Flint and Mott, 2008) or only very slightly lower (Nicod et al., 2016; Parker et al., 2016) than the effect sizes of QTLs for other complex traits.

Mouse and rat resources available to finely dissect the genetic basis of complex traits have been extensively reviewed (Peters et al., 2007; Flint, 2011; Flint and Eskin, 2012; Mott and Flint, 2013; Gonzales and Palmer, 2014; Williams and Williams, 2016). These reviews compare breeding schemes and relate them to genetic characteristics such as the total number of variants segregating, allele frequencies, rate of decay of linkage disequilibrium (a measure of how wide unrecombined intervals are), regions with low levels of polymorphism and specific analytic requirements. They also provide a structured discussion of the advantages and disadvantages of each population. By contrast, here we review genetic studies of complex neurobehavioural traits that have used these resources. We limit ourselves to resources with moderate to high levels of recombination (see Glossary, Box 1) for the reasons explained in the first section below, and finish by discussing the increasing relevance of outbred resources (see Glossary, Box 1).

Low levels of recombination in a population prevent gene identification

Gene identification typically starts with the mapping of QTLs and proceeds with integrative approaches, as reviewed recently by Moreno-Moral and Petretto (2016). A large number of QTLs for behaviour have been mapped in rodents [Mouse Genome Database,

Table 1. Rodent resources for complex traits genetics

Resource	Founders	No. founders	Breeding scheme	Inbred or outbred
Backcross (BC)	Classical inbred strains	2	2-way cross then backcross to founder strain	Outbred
Second generation intercross (F2)	Classical inbred strains	2	2-way cross then one intercross	Outbred
Congenic line	Classical inbred strains	2	2-way cross then multiple rounds of backcross and selection to obtain a line different from one of the founder strains at a single locus	Inbred
Advanced intercross line (AIL)	Classical inbred strains	2	2-way cross then multiple generations of intercross	Outbred
Heterogeneous stock (HS)	Classical inbred strains	> 2 (usually 8 or 4)	Rotational breeding over many generations	Outbred
Recombinant inbred line (RIL)	Classical inbred strains	2	2-way cross, intercross for one or more generations, then inbreeding	Inbred
Pre-collaborative cross (pre-CC)	Five classical and three wild-derived inbred strains	8	2-way, 4-way, 8-way cross then partial inbreeding	Partially inbred
Collaborative cross (CC)	Five classical and three wild-derived inbred strains	8	2-way, 4-way, 8-way cross then inbreeding	Inbred
Diversity outbred stock (DO)	Five classical and three wild-derived inbred strains	8	2-way, 4-way, 8-way cross, partial inbreeding then pseudorandom breeding	Outbred
Commercially available outbred stock	Unknown	'Swiss' stock	Pseudorandom mating	Outbred
Reduced complexity cross (RCC or coisogenic cross)	Substrains (i.e. very closely related strains)	2	2-way cross then one intercross	Outbred
Selection lines	An outbred population	–	Selection of animals with extreme phenotypes, used as breeders for the next generation, over many generations	Outbred

(Bult et al., 2016), www.informatics.jax.org; Rat Genome Database (Hayman et al., 2016; Shimoyama et al., 2014), http://rgd.mcw.edu/], but the genetic variants and genes that mediate the effects of these QTLs remain unclear (Flint et al., 2005; Flint and Mackay, 2009; Parker and Palmer, 2011). The main reason for this is that low levels of recombination exist in resources traditionally used for QTL mapping, namely F2 intercrosses and backcrosses (see Table 1 for a brief description of these populations).

Each chromosome in an F2 population results, on average, from one recombination event between the two founder genomes, meaning that large segments of DNA are unrecombined (Glossary, Box 1 and Fig. 1A). As all genetic variants within an unrecombined segment are perfectly correlated, if one variant is associated with a given trait then the other variants in the segment will also be associated with that trait (Fig. 1D). As a result, F2 QTLs typically encompass a very large number of genetic variants and tens if not hundreds of genes, all of which are candidate mediators of the QTL effect.

In order to narrow down the list of candidates, common strategies include integrating sequence information and/or expression data with QTLs. Sequence information available for the parental strains of the cross (Keane et al., 2011; Yalcin et al., 2011; Baud et al., 2014b; Hermsen et al., 2015) enables identification of all the variants that segregate in the cross. Those variants whose effect is predicted to be deleterious for protein structure or function (e.g. using the variant effect predictor, http://www.ensembl.org/info/docs/tools/vep/index.html) can be deemed likely to mediate the QTL effect.

Another way to prioritise among the variants and genes at a QTL is to use expression data, usually obtained from mRNA present in a relevant tissue. Expression can be mapped to identify which of the genes at the QTL are regulated by a local variant i.e. genes with a cis-eQTL (see Glossary, Box 1). Such genes are usually prioritized over genes without a cis-eQTL as they could mediate the QTL effects through changes in gene expression (e.g. gene A in Fig. 1). The strength of the correlation between gene expression and behaviour is often used to select candidate genes (Fig. 1D), but there

are caveats to this strategy (Doss et al., 2005; Mehrabian et al., 2005).

The strategies presented above have limited efficacy because typically, an F2 QTL encompasses a large number of deleterious variants (e.g. Parker et al., 2012) and about 10% of genes have a cis-eQTL in any tissue (Schadt et al., 2003). In addition, both strategies focus on protein-coding genes and make many assumptions, and therefore risk wrongly discarding the true genetic factors underlying the QTL. It is possible to alleviate these problems by mapping loci at higher resolution, thus reducing the number of candidate genes.

One way to obtain higher mapping resolution is to increase the sample size. However, obtaining adequate recombination levels with an F2 population would require phenotyping and genotyping of more than 10,000 animals (Mott and Flint, 2013). Breaking down one or a few intervals of interest into smaller unrecombined intervals by creating congenic lines (Table 1) is another option; however, such attempts have often seen the QTL disappear as it breaks down into two or more QTLs with smaller effects that cannot be detected (Flint and Mott, 2001; Legare et al., 2000). Finally, higher recombination levels can be achieved by crossing for multiple generations, as each chromosome will accumulate on average one meiotic recombination event each generation. Many breeding schemes exploit this process, and below, we discuss results obtained using the resulting populations.

Mapping neurobehavioural traits in moderately to highly recombinant populations

Advanced intercross lines (AILs) (Darvasi and Soller, 1995) are generated from two inbred founders by multiple generations of intercrossing (Table 1). In laboratory rats, a seventh-generation (F7) AIL descended from the experimental autoimmune encephalomyelitis (EAE, a model for multiple sclerosis)-susceptible DA and EAE-resistant PVG.1AV1 inbred strains was used to fine-map QTLs previously detected in an F2 cross. Of 1068 phenotyped rats, 152 affected and 162 unaffected animals were genotyped and used to map QTLs that were 1.3 Mb, 3 Mb and 5.5 Mb wide (1 LOD support intervals; see Glossary, Box 1).
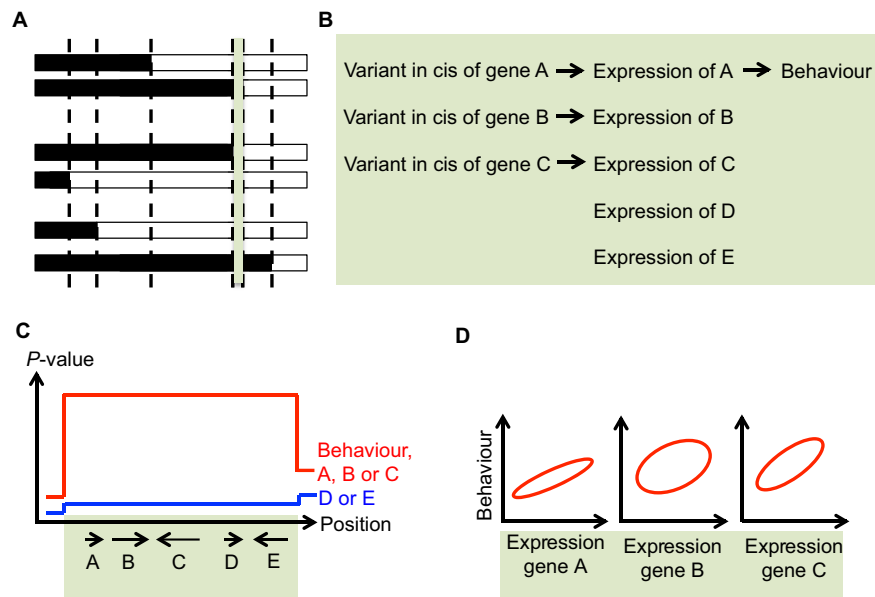
Fig. 1. Low levels of recombination preclude identification of genes associated with behavioural phenotypes. (A) Large unrecombined genomic segments in F2 crosses. Three pairs of homologous chromosomes, from three individuals of an F2 cross, are represented. As each chromosome presents on average only one recombination event, large segments of DNA are unrecombined in F2 crosses (even with hundreds of individuals). One such segment is highlighted in green. (B) Example scenario for behavioural variation. Variation in behaviour arises from differences in expression of gene A driven by a variant in cis of the gene. Gene A lies in the unrecombined segment shown in green in panel A with many other genes, four of which (gene B to gene E) are presented in this scenario. Expression of genes B and C is controlled by variants in cis, while expression of genes D and E is not. In reality the proportion of genes with a cis-eQTL in any given tissue is about 10%. (C) Primary analysis: QTL mapping of the behaviour and of the genes encompassed by the behavioural QTL. The position along the chromosome is shown on the x-axis and spans the green unrecombined interval. Genes A to E are represented by arrows. The y-axis shows the significance of the association between genotype and trait (behaviour or gene expression level). The red curve shows the shape of the association curve for behaviour and genes A, B and C, the blue curve that for genes D and E. The shape of the red curve denotes a QTL for the behaviour and a cis-eQTL for each of genes A, B, C. As genes A, B and C have a cis-eQTL, they will be considered candidate causal genes for the behavioural QTL effect. (D) Gene-behaviour correlation analysis to further prioritise candidate genes. The ovals represent the spread of the data points (not shown); hence the tighter the oval, the stronger the correlation between the variables on the x- and y-axes. A correlation between expression of gene A and behaviour is likely to exist as a result of the causal path shown in panel B. As genes B and C also have cis-eQTLs and the underlying genotypes are perfectly correlated with the genotypes of the cis-eQTL for gene A (no recombination in the green interval), a correlation will likely exist between expression of genes B and behaviour and expression of gene C and behaviour. The strength of these correlations will depend on many parameters, including non-local genetic and non-genetic effects. Hence, prioritising genes A, B, C based on the strength (or significance) of the correlations is not straightforward, and is subject to caveats.

Additional gene expression data, sequence data and prior knowledge were used to prioritise the genes at the QTL (Becanovic et al., 2006; Jagodic et al., 2004; Sheng et al., 2005).

A study of basal and methamphetamine-induced locomotor activity used 688 mice from a 34th-generation AIL descended from inbred strains SM/J and LG/J (Cheng et al., 2010). The authors report three genome-wide significant QTLs 0.5, 1.56 and 2.07 Mb wide (2 LOD support intervals), encompassing 1, 0 and more than 12 genes, respectively. *Csmd1* (CUB and Sushi multiple domains 1), which was identified in this study was knocked out for follow-up analyses but no effects on locomotor activity were detected (Distler et al., 2012). Another analysis of an eighth-generation AIL identified six QTLs ranging from 1.5 to 50 Mb in size, with a median of 15.6 Mb (1.8 LOD support intervals) (Parker et al., 2012). The studies highlighted here indicate that QTLs mapped in AILs can enable gene identification, provided enough recombination events have accumulated in the line over the generations.

Panels of recombinant inbred (RI) strains have traditionally been derived from inbreeding F2 animals (Table 1). In rats, two RI panels exist. The BXH/HXB panel was generated by reciprocal crossings of the spontaneously hypertensive rat (SHR/Ola) and the Brown Norway (BN-Lx/Cub) strains (Pravenec et al., 2004) and now contains 30 strains (Hubner et al., 2005). This panel has been used to map multiple behavioural traits, including startle, anxiety,

locomotion, conditioned taste aversion, alcohol consumption and learning (Printz et al., 2003; Conti et al., 2004; Bielavská et al., 2002; Tabakoff et al., 2009; Vanderlinden et al., 2014; Stuchlik et al., 2012). Although QTLs mapped in these studies were too large to allow identification of candidate genes, the panel was successfully included in a broader genetic study that identified a determinant of cardiac hypertrophy and mitochondrial function (McDermott-Roe et al., 2011). The second panel is derived from LE/Stm and F344/Stm and consists of 34 strains. Mapping of 109 traits, including neurobehavioural traits, was performed in this panel, but QTL intervals were again too large (>20 Mb) to allow gene identification (Voigt et al., 2008).

In mice, larger RI panels exist: the LXS panel is descended from inbred long sleep (ILS) and inbred short sleep (ISS) strains, which were selected for their ethanol sensitivity, and consists of 77 strains. Using 60 of those and more than 10 mice per strain, Bennett et al. (2015) mapped the genetic determinants of acute functional tolerance to the hypnotic effects of alcohol and identified a 23 Mb QTL (90% Bayesian credible interval) that included 716 genes. Similarly, large QTLs were mapped in a study of hearing loss in the same panel (Noben-Trauth et al., 2010). The BXD panel, descended from the inbred strain C57BL6/J, which is the mouse reference strain, and another strain, DBA/2J, contains around 120 lines that are almost fully inbred and are available from the Jackson Laboratory, and another set of 30-40 that are being inbred by

Williams, Lu and colleagues at the University of Tennessee Health Science Center (UTHSC) (Pandey and Williams, 2014). In this panel, 42 of the strains are derived from an F2 cross and the remaining are derived from F9 to F14 AILs, which improved the mapping resolution achievable with the panel (Peirce et al., 2004). It should be pointed out that very few studies make use of the full panel: many studies use only 30-40 strains and multiple replicates within each strain, resulting in very large QTLs (Putman et al., 2016; Harenza et al., 2014; Parsons et al., 2012; Ye et al., 2014; Young et al., 2016; Dickson et al., 2016).

Using 62 BXD strains, Carhuatanta et al. (2014) mapped five QTLs for fear and anxiety under chronic stress conditions. The size of the QTLs (1 LOD support intervals) ranged from 2.5 to 30.2 Mb with a median of 12.1 Mb. The narrowest interval encompassed only 15 genes, but no strong candidate gene was identified. Using 72 strains, Cook et al. (2015) mapped two QTLs 8.5 and 10 Mb in size (1.5 LOD support intervals) for anxiety-related traits following ethanol injection. Integrative approaches led the authors to prioritise four genes. Pandey and Williams (2014) used RNA expression data to show that the distance between the top cis-eQTL marker and the cognate gene could be as small as 0.6 Mb using 69 BXD strains. This demonstrates that the full BXD panel should enable causal genetic variants to be mapped very precisely. Using more strains would also improve the experimental power to detect QTLs. It is unclear, however, what the optimal number of strains and replicates is for a fixed number of mice. A very important advantage of the BXD panel is that a large amount of freely available phenotypic and molecular data that have been accumulated over the years on BXD strains and a web-based statistical analysis suite exists that can integrate these data to prioritise genes at QTLs (www.GeneNetwork.org). In this respect, the BXD panel is a unique resource.

AILs and RI panels exist that are descended not from two but four or eight inbred founder strains. They are respectively referred to as heterogeneous stocks (HS) (McClearn et al., 1970; Demarest et al., 2001; Hansen and Spuhler, 1984; Hitzemann et al., 2009; Iancu et al., 2010) and the collaborative cross (CC) (Churchill et al., 2004) (Table 1). HS are descended from the founder strains through usually more than 50 generations of outbreeding, which contributes to increasing mapping resolution. Mapping in HS and other populations descended from more than two inbred founders typically proceeds by reconstructing the chromosomes of the HS animals as mosaics of the founder genomes (Fig. 2) and mapping using haplotypes (Glossary,

Box 1) rather than genotypes. QTL intervals obtained are typically 4 Mb wide (Rat Genome Sequencing and Mapping Consortium et al., 2013; Valdar et al., 2006a).

Because the complete sequence of the strains from which HS are derived is now available both in mice (Keane et al., 2011; Yalcin et al., 2011) and rats (Baud et al., 2014b; Hermsen et al., 2015), 'merge analysis' can be used to refine the mapping and prioritise variants and the genes affected for further study (Mott et al., 2000; Rat Genome Sequencing and Mapping Consortium et al., 2013; Valdar et al., 2006a). Briefly, merge analysis proceeds by identifying variants whose strain distribution pattern amongst the founder strains of the cross is consistent with the QTL effect. Importantly, such prioritisation is based solely on association statistics and therefore makes no assumption as to the mechanisms underlying the QTL (e.g. changes in gene expression, altered protein sequence). Anxiety and EAE-related traits have been investigated in the rat NIH HS (Hansen and Spuhler, 1984) and QTLs were identified where a single candidate gene was identified by merge analysis (Rat Genome Sequencing and Mapping Consortium et al., 2013; Baud et al., 2014a): *Ctnnd2* (catenin delta-2) is associated with conditioned anxiety and *Fam198b* is associated with weight loss as a result of EAE. A role for *Ctnnd2* in anxiety in mice is supported by altered behaviour of *Ctnnd2*-knockout mice in contextual fear conditioning paradigms (Israely et al., 2004) and an association between *CTNND2* and anxiety has been found in humans too (Nivard et al., 2014). *Ctnnd2* therefore seems to contribute to anxiety in multiple species (rats, mice, humans). At other QTLs mapped in the rat NIH-HS, additional information on the variants segregating at the QTL, namely their predicted effect on protein function, was used to identify candidate genes: MHC class II genes *RT1-Da* and *RT1-Bb* were identified as potential contributors to EAE-associated weight loss. In mice, independent mapping of emotionality in the Boulder and the Northport HS (Talbot et al., 1999; Valdar et al., 2006a) identified the same 2 Mb region around *Rgs2*, a gene that was proven to underlie a QTL for emotionality in commercially available outbred mice (described below, Yalcin et al., 2004). Fine-mapping of anxiety-related traits in the Northport HS together with analysis of functional gene annotations and testing of a conditional mutant mice provided evidence that actin filament depolymerisation and expression of Cofilin-1 (*Cfl1*) in the ventral hippocampus may mediate anxiety (Goodson et al., 2012).
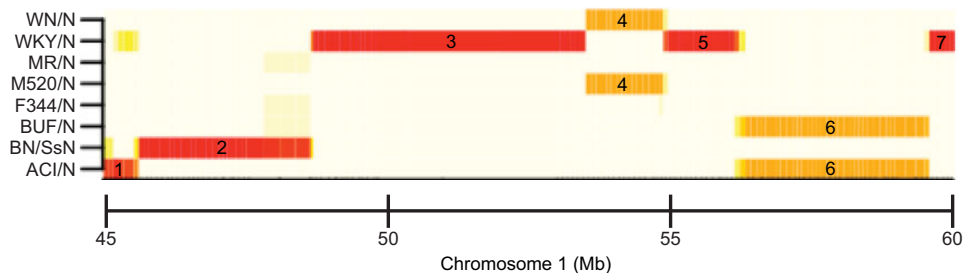

Fig. 2. Reconstruction of the chromosomes of an HS rat as a probabilistic mosaic of the eight founder genomes. A 16 Mb region of chromosome 1 is represented (*x*-axis). The figure describes the probability that the pair of homologous chromosomes is descended from each of the eight founders (*y*-axis) using a colour code (white: 0; red: 2; yellow ~1). Seven sub-segments (labelled 1-7) can be identified based on breaks in the colour pattern. The first sub-segment (45-45.5 Mb) has a probability of 2 (maximum probability) to be descended from the founder ACI/N; this means that we can be confident both homologous chromosomes are descended from that founder. Similarly, both chromosomes are descended from BN/SsN in the second segment and from WKY/N in the third segment. The fourth segment shows equal probability (probability of 1) to be descended from WN/N and M520/N; this can be interpreted either as one chromosome is descended from WN/N and the other from M520/N or those two founders are similar in that region and so we are unsure whether both chromosomes are descended from one or the other founder. Both chromosomes are descended from WKY/N in segments 5 and 7, and they are descended from BUF/N and/or ACI/N in segment 6. This fragmented pattern of genomic inheritance is referred to as 'mosaic'. The probabilities can be used as input to the statistical models used to map QTLs.

The rat NIH-HS was used to derive replicated high- and low-alcohol-drinking lines through bi-directional selection (Foroud et al., 2000) (Table 1). Subsequent mapping of alcohol preference identified QTLs that were too large for gene identification but, combined with an analysis of signatures of selection, the selection lines pointed to specific genes (*Grin2a*, *Cyp4f18*, *Myo9b*, *Pgls* and *Fam129c*) (Lo et al., 2016), illustrating the potential of this combined approach.

The CC is a panel of recombinant inbred strains descended from eight inbred founders including three wild-derived inbred strains, which ensure a very high level of genetic diversity and variation in almost any trait of interest (Churchill et al., 2004). Interestingly, the wild-derived strains contribute genetic variants that cause anomalous behaviour in CC mice (in comparison with classical laboratory mice) with practical consequences, thus care is required when interpreting behavioural data (Chesler, 2014). Published studies have so far mostly used pre-CC mice, i.e. mice from partially (~75%) inbred lines in the CC breeding colonies (Table 1). Using pre-CC mice, high mapping resolution was achieved across a range of phenotypes and molecular traits. QTLs for reproductive, behavioural, physiological and morphological traits were 4 Mb wide on average (1.5 LOD support intervals) (Philip et al., 2011). Philip et al. (2011) reported that the QTL interval detected for 'activity after sleep deprivation' harbours only three positional candidates: a microRNA and two genes: *Ntm* (neurotrimin), a neural cell-adhesion molecule that plays a role in brain development and *Snx19* (sorting nexin 19), which possesses a regulator of G protein-coupled receptor signalling (RGS) domain. The same authors identified a QTL interval for thermal nociception comprising just six genes: *Slit2*, *Pacrgl*, *Gpr125*, *Dhx15*, *Sod3* and *Kcnip4*.

Large pre-CC studies typically used many lines and one or few mice per line, suggesting that more lines and fewer replicates per line may be the best design for genetic mapping. Breeding of pre-CC mice is difficult and many lines have gone extinct (Threadgill et al., 2011; Chesler et al., 2008). A subset of 75 (as of time of writing) extant lines have reached a sufficient degree of inbreeding and meet criteria to be distributed to all investigators [available lines are updated at www.csbio.unc.edu/CCstatus/index.py?run= AvailableLines; see also Morgan and Welsh (2015) for informatics resources available to analyse CC data]. As this number is much lower than the number of lines used in pre-CC studies (~250 in pre-CC studies) and because CC lines have different genetic characteristics compared with pre-CC lines, the potential of CC mice needs evaluating. CC lines that are available to all investigators have been used in a very limited number of published studies to date and these studies used only a subset of all currently available lines (e.g. 16 in Mao et al., 2015). Thus, the potential of currently available CC lines to allow identification of genes associated with behaviour and other complex traits is unclear.

The CC panel – and indeed any recombinant inbred panel – can further be used by intercrossing CC lines (Graham et al., 2015; Rasmussen et al., 2014) or crossing CC lines to an inbred strain with a genetic variant of interest to identify modifiers of the variant (Chitsazan et al., 2016).

An outbred population of mice called diversity outbred (DO) stock (Svenson et al., 2012; Churchill et al., 2012) was also derived from pre-CC lines (Table 1). A website has been set up to host DO data and support their analysis (do.jax.org; see also Morgan and Welsh, 2015; Gatti et al., 2014). A total of 38 behavioural measures related to activity, anxiety and response to novelty were mapped in 283 DO mice from generations 4 and 5. Five significant QTLs were identified, with 1.5 LOD support intervals ranging from 1 to 7 Mb

(Logan et al., 2013). A QTL for climbing frequency during the tail suspension test (a measure described by the authors as one of wildness rather than depressed mood, which the test is traditionally used to measure) encompassed only five genes. Fine-mapping by identifying those variants at the QTL that have a strain distribution pattern that is consistent with the QTL effect in the founder strains (a similar approach to merge analysis) failed for this and some of the other QTLs, because variants with the appropriate strain distribution pattern were found to populate the entire QTL region (Logan et al., 2013). This situation is expected to be common with DO QTLs as a result of wild-derived founder strains being very different from each other and from the five classical inbred founder strains, and because wild-derived alleles drive most of the QTLs. As DO and CC share the same founders, this situation is also common in CC QTLs (Durrant et al., 2011). As mentioned and illustrated in Logan et al. (2013), wild-derived genetic variants also cause DO mice (like CC mice) to behave 'inappropriately' on certain tests (e.g. climbing in the tail suspension test), thus warranting care in the analysis and interpretation of the corresponding data.

Although both CC and DO were designed to present as much genetic variation as possible, an opposite strategy led to the development of reduced complexity crosses (RCC, also called coisogenic crosses). RCC are crosses of two closely related strains, typically two strains that are descended from a single inbred strain through independent evolution in different laboratories (Table 1). Kumar et al. (2013) observed differences in cocaine response between C57BL/6J, a C57BL/6 strain established and maintained at the Jackson Laboratory since 1948, and C57BL/6N, a C57BL/6 strain established at NIH in 1951. Subsequent mapping in an F2 cross identified a QTL that accounted for 61% of the genetic variance (i.e. the phenotypic variation of genetic origin). The QTL interval was 22 Mb wide, which in any other cross would have precluded identification of candidate genes. However, because there is very little genetic variation between the two founder strains, a single genetic variant (not gene) affecting the protein sequence within the QTL interval was suspected to cause the QTL effect. This variant encodes a serine-to-phenylalanine missense mutation in *Cyfip2* (cytoplasmic FMRP-interacting protein 2), a gene implicated in Fragile-X mental retardation.

Kumar et al. (2013) followed up on this candidate gene by measuring the stability of the corresponding protein in the founders C57BL/6J and C57BL/6N, and by behavioural, histological and electrophysiological profiling of a *Cyfip2* mouse knockout. All analyses supported a role for *Cyfip2* in the response to cocaine exposure. This study illustrates how crosses where only a few genetic variants segregate can solve the recurrent problem of translating QTLs into information about causal genes and variants. Importantly, this study also shows that in RCC-based studies, the size of QTL intervals is not a good indication of the potential to identify causal variants and genes (Williams and Williams, 2016), as it is limited by genotyping density rather than recombination levels.

Additional phenotypes including behaviours have been shown to vary across C57BL/6 substrains (Simon et al., 2013; Khisti et al., 2006; Mulligan et al., 2008; Kirkpatrick and Bryant, 2014) and are amenable to genetic mapping. Substrains also exist in the laboratory rat and the genetic variants that segregate between some of them have been catalogued (Atanur et al., 2013; Hermsen et al., 2015), allowing genetic study of phenotypic differences (Zhang-James et al., 2013).

The last type of resource we will review is commercially available outbred rats and mice, which refers to animals that have been

bred by commercial vendors for decades, primarily for use in pharmacological studies (Table 1). These populations are typically maintained using an outbreeding regime in order to preserve genetic diversity. For this reason, they should never be called 'strains' as a strain is the result of many generations (usually 20) of inbreeding. Moreover, a review of commercially available mice by Yalcin et al. (2010) highlighted important genetic differences between colonies of the same stock, where colony refers to 'a population of mice maintained as a mating population at a single location', and stock refers to 'a collection of colonies that are given the same stock designation by the breeders'. Thus, colony is the unit of interest for genetic studies using commercially available outbred mice, and the same is likely to be true for rats. Yalcin et al. (2010) investigated genetic characteristics that are important for genetic mapping (in particular, rate of linkage disequilibrium decay and minor allele frequencies) of 66 available colonies, providing valuable information to those wishing to choose a colony for genetic mapping. Prior to this study, 729 mice from the HsdOla:MF1-UK colony (nomenclature according to Yalcin and Flint, 2012) had been used to fine-map a chromosome-1 QTL for anxiety in mice (Yalcin et al., 2004). The study led to the identification of a regulator of G-protein signalling, *Rgs2*, as a strong positional candidate, and confirmed that it accounted for some of the QTL effect by quantitative complementation (Glossary, Box 1).

In 2016, two studies were published that investigated the genetic basis of a variety of complex traits, including many neurobehavioural traits, in commercially available outbred mice from the Crl:CFW(SW)-US_P08 (CFW) colony (Nicod et al., 2016; Parker et al., 2016). The size of the 95% CI for the 255 QTLs mapped by Nicod et al. (2016) using ~1800 mice ranged from 0.01 to 7.33 Mb, with a mean at 1.50 Mb and 43% of intervals smaller than 1 Mb (the size of CIs was not reported by Parker et al., 2016). It constitutes the highest mapping resolution achieved genome-wide in any rodent resource thus far. In the study by Nicod et al. (2016), 7 QTLs for neurobehavioural traits encompassed a single gene, thus identifying 7 very strong positional candidates: (1) Glutamate receptor metabotropic 7 (*Grm7*) was associated with total distance travelled in elevated plus maze (see Ellenbroek and Youn, 2016 for a description of this test) and a role for *Grm7* in locomotor activity in a new environment is supported by evidence from a mouse knockout (Cryan et al., 2003). (2) Unc-13 homolog C (*Unc13c*) was associated with the number of long (>1 min) sleep episodes, and there is also evidence for differential expression of the human orthologue in individuals with poor sleep quality (Reddy et al., 2014). (3) *Met* and (4) *Rtkn2* were associated with startle pulse reactivity; (5) *Ppargc1a* with number of long (>1 min) sleep episodes; (6) *Adarb2* with basal home cage activity and (7) *Pcdh17* with total distance travelled in elevated plus maze, although there was no independent evidence for a role of these genes in these phenotypes.

In addition to collecting multiple behavioural phenotypes related to conditioned fear, anxiety-like behaviour, methamphetamine sensitivity and prepulse inhibition on ~900 mice, Parker et al. (2016) measured gene expression in a subset of the mice in the hippocampus (*n*=79), striatum (*n*=55) and prefrontal cortex (*n*=54). Integrating QTLs for behavioural and gene expression traits, they prioritised those genes at QTLs that had a cis-eQTL in a relevant brain region and whose expression was correlated with the behaviour. *Azi2* (5-azacytidine-induced gene 2) was the best candidate gene at a QTL for methamphetamine sensitivity, and *Zmynd11* (zinc finger, MYND domain-containing 11) at a QTL for anxiety-like behaviour. *Zmynd11* has independently been suggested to be involved in various psychiatric disorders (Coe et al., 2014).

Commercially available outbred rats also exist, but the genetic characteristics of the different colonies have not yet been assessed, nor have such rats been used in mapping studies.

We recapitulate in Table S1 the 54 significant QTLs reported in the studies reviewed above and that have led to the identification of 1 to 5 positional candidate genes. This table demonstrates that the genetic determinants of behaviour can now be mapped with high precision both in rats and mice. Importantly, we have not reported in Table S1 genes cloned using congenic strains. Those would only strengthen our point, however, which is that the time is over when high numbers of positional candidate genes at QTLs were the limiting factor for the usefulness of rodent models in genetic studies. What seems to be the limiting factor is our ability to follow up on narrow QTLs with experimental validation and functional characterization, as discussed below.

Validation of QTLs: challenges and opportunities

Candidate genes have been followed up for only three out of 54 QTLs (for a fourth QTL, there was prior evidence confirming a role of the candidate gene *Ctnnd2*). The first case is that of *Rgs2*, a gene associated with anxiety in the mouse Northport heterogeneous stock. Its contribution to the QTL was confirmed by quantitative complementation (Yalcin et al., 2004). The second case is *Csmd1*, a gene associated with locomotor activity in a 34th-generation AIL (Cheng et al., 2010). In this case, no effect of the gene on locomotor activity was observed in the corresponding knockout model (Distler et al., 2012). The third example is *Cyfip2*, a gene associated with cocaine response in a RCC between the C57BL/6J and C57BL/6N substrains (Kumar et al., 2013). Follow-up included demonstrating differential stability of CYFIP2 between the founder strains, and histological, electrophysiological and behavioural testing of a *Cyfip2* knockout.

Why are QTLs so difficult to follow up on? Firstly, there is ever increasing evidence, mostly from human genetics studies, that the mechanisms of genetic control are not straightforward. In line with this, 43% of trait-associated SNPs in humans lie in intergenic regions and 45% lie in introns (Hindorff et al., 2009); intronic variants have been shown to regulate distant genes in some instances (Rask-Andersen et al., 2015); finally, there is also mounting evidence that QTLs often arise from multiple variants and possibly multiple genes (Rat Genome Sequencing and Mapping Consortium et al., 2013; Allen et al., 2010). As a result, investigators may have limited confidence in the variants or genes they identify as candidates at QTLs.

A second difficulty lies in creating a genetically engineered knockout model in order to validate a candidate gene. Until recently, knockout models were available for only a few genes, and creating a knockout model was very time- and resource consuming. However, things are changing: first, more and more knockouts are becoming available through a very large international consortium – the International Mouse Phenotyping Consortium (http://www.mousephenotype.org/); secondly, the recent advent of the CRISPR/Cas technology for genome editing (Glossary, Box 1) has made it possible for individual laboratories to easily create their own, personalised knockout model (Barrangou, 2014). We refer the reader to three reviews on CRISPR/Cas technology, which include a comparison with ZFN and TALEN technologies (Gaj et al., 2013; Kim and Kim, 2014; Sander and Joung, 2014). In brief, the simplicity and high efficiency of the CRISPR/Cas system makes it a very attractive alternative to traditional knockout procedures.

In the future, it is possible that knockouts will not be the go-to models for following up on QTLs. Indeed, variants underlying

QTLs have subtle effects that loss-of-function models poorly recapitulate, and variants associated with complex traits have been shown to often sit outside the protein-coding regions. Thus, it may be more appropriate to validate variants rather than genes. CRISPR/Cas technology permits replacement of specific single nucleotides and addition or deletion of specific sequences. For example, Yoshimi et al. (2014) showed that they could revert each of three mutations of F344 rats (albino, non-agouti and hooded). Thus, CRISPR/Cas opens new avenues for follow up of QTLs through variant editing. Unfortunately, even the highest mapping resolution achieved to this day leaves tens if not hundreds of variants segregating at most QTLs. To narrow down the list of candidate variants to a tractable number (10, 20 maybe for CRISPR), one can use additional information regarding the position of the variants (e.g. lying in a promoter or at a splice site), their predicted effect on protein structure (e.g. predicted to affect binding with DNA), functional evidence from the mouse ENCODE project (www.mouseencode.org/), or from analysis of evolutionary constraints, for example.

Finally, an added difficulty of validating a QTL using a genetically engineered model is the genetic background on which the mutant is created. The same mutation may have observable effects in one background but not in another (Holmes et al., 2003) and it may even have opposing effects in different backgrounds (Sittig et al., 2016) Therefore, when the goal is to validate a QTL, a mutant should ideally be created on the background in which the QTL was detected. Unfortunately, in all mapping populations except for congenic lines, genetic variants exist not only in the QTL interval but everywhere in the genome; this means that each animal or strain of a mapping population represents a unique genetic background and there is not a single background on which the QTL should be placed. One solution is to place QTL alleles on multiple genetic backgrounds representative of the original population. For example, the same mutation could be introduced in multiple DO mice and mutant DO mice compared with non-modified DO mice. This expensive validation strategy will maximize the investigators' chances to replicate the QTL effect, and will have the added benefit of increasing the generalisability of the results to multiple genetic backgrounds – an important first step towards applying the information to humans (Sittig et al., 2016). The efficiency of the CRISPR/Cas technology in seemingly any background is yet another major advantage of this technology, which truly seems to hold the key to the future of animal model-based investigation of complex traits.

Inbred versus outbred resources

In light of the success of outbred resources in precisely mapping phenotypic variation, we wish to discuss what has been seen as unique advantages of inbred resources. Inbred resources include RI panels derived from two or more progenitors and combinations of inbred and RI strains [e.g. the hybrid mouse diversity panel, HMDP (Bennett et al., 2010)]; outbred resources include AIL, HS, DO and commercially available outbred stocks. Traditionally, the main selling points of inbred resources were threefold. First, most inbred strains have already been genotyped and need not be genotyped again, whereas outbred animals are genetically unique and thus need to be genotyped in each experiment. Recently, next-generation sequencing (NGS) has been used to genotype commercially available outbred mice, using two different approaches (Davies et al., 2016; Parker et al., 2016), showing that NGS is an attractive alternative to microarrays for genotyping outbred populations.

Second, multiple measurements can be made on the same genotypes (using multiple animals of the same inbred strain),

thus reducing non-genetic variation and facilitating detection of genetic associations. However, when inbred panels are large enough (e.g. BXD RI panel), using all strains and only one animal per strain might be more powerful than using fewer strains and multiple animals per strain (Fig. 1B in Andreux et al., 2012). Using one animal per strain effectively means using an outbred sample.

The third advantage is that the same genotypes can be used in multiple experiments. This permits accumulation of phenotypic data over the years, which facilitates systems genetics approaches. The best example of this is the accumulation of more than 5000 organismal phenotypes, mRNA expression in ~33 tissues, microRNA and protein expression data, as well as metabolomics data in the BXD RI panel (www.genenetwork.org). The availability of multiple animals with the same genotypes has also facilitated studies of gene by environment (G×E) interactions and sex-specific genetic effects (G×S) (Peirce et al., 1998). Last but not least, this panel has been used to evaluate the robustness of positive findings through replication studies.

However, it is possible to tackle these goals using outbred populations, as illustrated by the following studies. Krohn et al. (2014) identified G×S effects in outbred HS mice, showing that such effects were widespread but relatively small. French et al. (2015) replicated benzene-induced genotoxicity in one study of two cohorts of DO mice, illustrating the concept of replication for different sets of genotypes. They did not go as far as mapping QTLs in the two cohorts however, but instead combined them for QTL mapping (presumably because of sample size considerations). In theory, genotype to phenotype associations involving variants in Hardy–Weinberg equilibrium (see Glossary, Box 1) could be investigated in multiple samples collected in different generations of an outbred population, as allelic frequencies are expected to remain constant. However, there is a caveat in that other variants that are not in Hardy–Weinberg equilibrium might cause differences in genetic background and interact with the evaluated variant, modifying its effect. Thus, the potential of outbred populations for replication studies and to investigate G×S and G×E interactions remains to be fully exploited.

Conclusions

Multiple highly recombinant rat and mouse resources now allow precise mapping of phenotypic variation. However, experimental validation and functional characterisation of candidate genes remain major obstacles in the way of identifying causal genes and pathways. Both steps are crucial to turn QTLs into biological understanding of the mechanisms underlying behaviour. The advent of the CRISPR/Cas technology holds great promises in this regard, yet important experimental parameters such as the genetic background on which mutants are generated need to be considered carefully.

It is important to stress that many of the resources highlighted here are not fixed – they are evolving, be it by addition of new RI strains (e.g. BXD panel), further inbreeding of current lines (e.g. CC) or further outbreeding of current population (e.g. DO, HS). Furthermore, new resources can easily be created from existing ones (e.g. RIX), and resources can be combined (Bennett et al., 2010; Furlotte et al., 2012). Thus, rodent resources can only improve in the future.

Disease Models & Mechanisms

References
Allardyce, J., Gaebel, W., Zielasek, J. and van Os, J. (2007). Deconstructing
Psychosis conference February 2006: the validity of schizophrenia and alternative
approaches to the classification of psychosis. *Schizophr. Bull.* **33**, 863-867.
**Allen, H. L., Estrada, K., Lettre, G., Berndt, S. I., Weedon, M. N., Rivadeneira, F.,
Willer, C. J., Jackson, A. U., Vedantam, S., Raychaudhuri, S. et al.** (2010).
Hundreds of variants clustered in genomic loci and biological pathways affect
human height. *Nature* **467**, 832-838.
American Psychiatric Association (2000). *Diagnostic and Statistical Manual of
Mental Disorders, Fourth Edition: DSM-IV-TR®*. Washington, DC: American
Psychiatric Association.
**Andreux, P. A., Williams, E. G., Koutnikova, H., Houtkooper, R. H., Champy, M.-
F., Henry, H., Schoonjans, K., Williams, R. W. and Auwerx, J.** (2012). Systems
genetics of metabolism: the use of the BXD murine reference panel for multiscalar
integration of traits. *Cell* **150**, 1287-1299.
**Atanur, S. S., Diaz, A. G., Maratou, K., Sarkis, A., Rotival, M., Game, L.,
Tschannen, M. R., Kaisaki, P. J., Otto, G. W., Ma, M. C. J. et al.** (2013). Genome
sequencing reveals loci under artificial selection that underlie disease phenotypes
in the laboratory rat. *Cell* **154**, 691-703.
Barrangou, R. (2014). Cas9 targeting and the CRISPR revolution. *Science* **344**,
707-708.
Baud, A., Flint, J. and Fernandez-Teruel, A. and **Rat Genome Sequencing
Mapping Consortium** (2014a). Identification of genetic variants underlying anxiety
and multiple sclerosis in heterogeneous stock rats. *World J. Neurosci.* **4**, 216-224.
**Baud, A., Guryev, V., Hummel, O., Johannesson, M., Rat Genome Sequencing
and Mapping Consortium and Flint, J.** (2014b). Genomes and phenomes of a
population of outbred rats and its progenitors. *Sci. Data* **1**, 140011.
**Becanovic, K., Jagodic, M., Sheng, J. R., Dahlman, I., Aboul-Enein, F.,
Wallstrom, E., Olofsson, P., Holmdahl, R., Lassmann, H. and Olsson, T.**
(2006). Advanced intercross line mapping of Eae5 reveals Ncf-1 and CLDN4 as
candidate genes for experimental autoimmune encephalomyelitis. *J. Immunol.*
176, 6055-6064.
**Bennett, B. J., Farber, C. R., Orozco, L., Kang, H. M., Ghazalpour, A., Siemers,
N., Neubauer, M., Neuhaus, I., Yordanova, R., Guan, B. et al.** (2010). A high-
resolution association mapping panel for the dissection of complex traits in mice.
Genome Res. **20**, 281-290.
**Bennett, B., Larson, C., Richmond, P. A., Odell, A. T., Saba, L. M., Tabakoff, B.,
Dowell, R. and Radcliffe, R. A.** (2015). Quantitative trait locus mapping of acute
functional tolerance in the LXS recombinant inbred strains. *Alcoholism* **39**, 611-620.
Bielavská, E., Kren, V., Musilová, A., Zídek, V. and Pravenec, M. (2002). Genome
scanning of the HXB/BXH sets of recombinant inbred strains of the rat for
quantitative trait loci associated with conditioned taste aversion. *Behav. Genet.*
32, 51-56.
Bult, C. J., Eppig, J. T., Blake, J. A., Kadin, J. A. and Richardson, J. E. and **the
Mouse Genome Database Group** (2016). Mouse genome database 2016.
Nucleic Acids Res. **44**, D840-D847.
Carhuatanta, K. A. K., Shea, C. J. A., Herman, J. P. and Jankord, R. (2014).
Unique genetic loci identified for emotional behavior in control and chronic stress
conditions. *Front. Behav. Neurosci.* **8**, 341.
**Casey, B. J., Craddock, N., Cuthbert, B. N., Hyman, S. E., Lee, F. S. and Ressler,
K. J.** (2013). DSM-5 and RDoC: progress in psychiatry research? *Nat. Rev.
Neurosci.* **14**, 810-814.
**Cheng, R., Lim, J. E., Samocha, K. E., Sokoloff, G., Abney, M., Skol, A. D. and
Palmer, A. A.** (2010). Genome-wide association studies and the problem of
relatedness among advanced intercross lines and other highly recombinant
populations. *Genetics* **185**, 1033-1044.
Chesler, E. J. (2014). Out of the bottleneck: the Diversity Outcross and
Collaborative Cross mouse populations in behavioral genetics research. *Mamm.
Genome* **25**, 3-11.
**Chesler, E. J., Miller, D. R., Branstetter, L. R., Galloway, L. D., Jackson, B. L.,
Philip, V. M., Voy, B. H., Culiat, C. T., Threadgill, D. W., Williams, R. W. et al.**
(2008). The Collaborative Cross at Oak Ridge National Laboratory: developing a
powerful resource for systems genetics. *Mamm. Genome* **19**, 382-389.
**Chitsazan, A., Ferguson, B., Ram, R., Mukhopadhyay, P., Handoko, H. Y.,
Gabrielli, B., Soyer, P. H., Morahan, G. and Walker, G. J.** (2016). A mutation in
the Cdon gene potentiates congenital nevus development mediated by
NRASQ61K. *Pigment Cell Melanoma Res.* **29**, 459-464.

**Churchill, G. A., Airey, D. C., Allayee, H., Angel, J. M., Attie, A. D., Beatty, J.,
Beavis, W. D., Belknap, J. K., Bennett, B., Berrettini, W. et al.** (2004). The
Collaborative Cross, a community resource for the genetic analysis of complex
traits. *Nat. Genet.* **36**, 1133-1137.
Churchill, G. A., Gatti, D. M., Munger, S. C. and Svenson, K. L. (2012). The
Diversity Outbred mouse population. *Mamm. Genome* **23**, 713-718.
**Coe, B. P., Witherspoon, K., Rosenfeld, J. A., Van Bon, B. W. M., Vulto-van
Silfhout, A. T., Bosco, P., Friend, K. L., Baker, C., Buono, S., Vissers, L. E. L.
M. et al.** (2014). Refining analyses of copy number variation identifies specific
genes associated with developmental delay. *Nat. Genet.* **46**, 1063-1071.
Conti, L. H., Jirout, M., Breen, L., Vanella, J. J., Schork, N. J. and Printz, M. P.
(2004). Identification of quantitative trait loci for anxiety and locomotion
phenotypes in rat recombinant inbred strains. *Behav. Genet.* **34**, 93-103.
Cook, M. N., Baker, J. A., Heldt, S. A., Williams, R. W., Hamre, K. M. and Lu, L.
(2015). Identification of candidate genes that underlie the QTL on chromosome 1
that mediates genetic differences in stress-ethanol interactions. *Physiol.
Genomics* **47**, 308-317.
Cross-Disorder Group of the Psychiatric Genomics Consortium (2013a).
Genetic relationship between five psychiatric disorders estimated from genome-
wide SNPs. *Nat. Genet.* **45**, 984-994.
Cross-Disorder Group of the Psychiatric Genomics Consortium (2013b).
Identification of risk loci with shared effects on five major psychiatric disorders: a
genome-wide analysis. *Lancet* **381**, 1371-1379.
Cryan, J. F. et al. (2003). Antidepressant and anxiolytic-like effects in mice lacking
the group III metabotropic glutamate receptor mGluR7. *Eur. J. Neurosci.* **17**,
2409-2417.
Darvasi, A. and Soller, M. (1995). Advanced intercross lines, an experimental
population for fine genetic mapping. *Genetics* **141**, 1199-1207.
Davies, R. W., Flint, J., Myers, S. and Mott, R. (2016). Rapid genotype imputation
from sequence without reference panels. *Nat. Genet.* **48**, 965-969.
Demarest, K., Koyner, J., McCaughran, J., Jr, Cipp, L. and Hitzemann, R.
(2001). Further characterization and high-resolution mapping of quantitative trait
loci for ethanol-induced locomotor activity. *Behav. Genet.* **31**, 79-91.
**Demyttenaere, K., Bruffaerts, R., Posada-Villa, J., Gasquet, I., Kovess, V.,
Lepine, J. P., Angermeyer, M. C., Bernert, S., De Girolamo, G., Morosini, P.
et al.** (2004). Prevalence, severity, and unmet need for treatment of mental
disorders in the World Health Organization World Mental Health Surveys. *JAMA*
291, 2581-2590.
**Dickson, P. E., Miller, M. M., Calton, M. A., Bubier, J. A., Cook, M. N., Goldowitz,
D., Chesler, E. J. and Mittleman, G.** (2016). Systems genetics of intravenous
cocaine self-administration in the BXD recombinant inbred mouse panel.
Psychopharmacology **233**, 701-714.
Distler, M. G., Opal, M. D., Dulawa, S. C. and Palmer, A. A. (2012). Assessment of
behaviors modeling aspects of schizophrenia in Csmd1 mutant mice. *PLoS ONE*
7, e51235.
Doss, S., Schadt, E. E., Drake, T. A. and Lusis, A. J. (2005). Cis-acting expression
quantitative trait loci in mice. *Genome Res.* **15**, 681-691.
Dupuis, J. and Siegmund, D. (1999). Statistical methods for mapping quantitative
trait loci from a dense set of markers. *Genetics* **151**, 373-386.
**Durrant, C., Tayem, H., Yalcin, B., Cleak, J., Goodstadt, L., de Villena, F. P.-M.,
Mott, R. and Iraqi, F. A.** (2011). Collaborative Cross mice and their power to map
host susceptibility to Aspergillus fumigatus infection. *Genome Res.* **21**,
1239-1248.
Ellenbroek, B. and Youn, J. (2016). Rodent models in neuroscience research: is it
a rat race? *Dis. Model. Mech.* **9**, 1079-1087.
Flint, J. (2011). Mapping quantitative traits and strategies to find quantitative trait
genes. *Methods* **53**, 163-174.
Flint, J. and Eskin, E. (2012). Genome-wide association studies in mice. *Nat. Rev.
Genet.* **13**, 807-817.
Flint, J. and Kendler, K. S. (2014). The genetics of major depression. *Neuron* **81**,
484-503.
Flint, J. and Mackay, T. F. C. (2009). Genetic architecture of quantitative traits in
mice, flies, and humans. *Genome Res.* **19**, 723-733.
Flint, J. and Mott, R. (2001). Finding the molecular basis of quantitative traits:
successes and pitfalls. *Nat. Rev. Genet.* **2**, 437-445.
Flint, J. and Mott, R. (2008). Applying mouse complex-trait resources to
behavioural genetics. *Nature* **456**, 724-727.
Flint, J., Valdar, W., Shifman, S. and Mott, R. (2005). Strategies for mapping and
cloning quantitative trait genes in rodents. *Nat. Rev. Genet.* **6**, 271-286.
**Foroud, T., Bice, P., Castelluccio, P., Bo, R., Miller, L., Ritchotte, A., Lumeng, L.,
Li, T.-K. and Carr, L. G.** (2000). Identification of quantitative trait loci influencing
alcohol consumption in the high alcohol drinking and low alcohol drinking rat lines.
Behav. Genet. **30**, 131-140.
**French, J. E., Gatti, D. M., Morgan, D. L., Kissling, G. E., Shockley, K. R.,
Knudsen, G. A., Shepard, K. G., Price, H. C., King, D. and Witt, K. L.** (2015).
Diversity outbred mice identify population-based exposure thresholds and genetic
factors that influence benzene-induced genotoxicity. *Environ. Health Perspect.*
123, 237.
Furlotte, N. A., Kang, E. Y., Van Nas, A., Farber, C. R., Lusis, A. J. and Eskin, E.
(2012). Increasing association mapping power and resolution in mouse genetic

studies through the use of meta-analysis for structured populations. *Genetics* **191**, 959-967.

Gaj, T., Gersbach, C. A. and Barbas, C. F. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31**, 397-405.

Gatti, D. M., Svenson, K. L., Shabalin, A., Wu, L.-Y., Valdar, W., Simecek, P., Goodwin, N., Cheng, R., Pomp, D., Palmer, A. et al. (2014). Quantitative trait locus mapping methods for diversity outbred mice. *G3* **4**, 1623-1633.

Gonzales, N. M. and Palmer, A. A. (2014). Fine-mapping QTLs in advanced intercross lines and other outbred populations. *Mamm. Genome* **25**, 271-292.

Goodson, M., Rust, M. B., Witke, W., Bannerman, D., Mott, R., Ponting, C. P. and Flint, J. (2012). Cofilin-1: a modulator of anxiety in mice. *PLoS Genet.* **8**, e1002970.

Graham, J. B., Thomas, S., Swarts, J., McMillan, A. A., Ferris, M. T., Suthar, M. S., Treuting, P. M., Ireton, R., Gale, M. and Lund, J. M. (2015). Genetic diversity in the collaborative cross model recapitulates human west nile virus disease outcomes. *MBio* **6**, e00493-15.

Hansen, C. and Spuhler, K. (1984). Development of the National Institutes of Health genetically heterogeneous rat stock. *Alcoholism* **8**, 477-479.

Harenza, J. L., Muldoon, P. P., De Biasi, M., Damaj, M. I. and Miles, M. F. (2014). Genetic variation within the Chrna7 gene modulates nicotine reward-like phenotypes in mice. *Genes Brain Behav.* **13**, 213-225.

Hayman, G. T., Laulederkind, S. J. F., Smith, J. R., Wang, S.-J., Petri, V., Nigam, R., Tutaj, M., De Pons, J., Dwinell, M. R. and Shimoyama, M. (2016). The Disease Portals, disease–gene annotation and the RGD disease ontology at the Rat Genome Database. *Database* **2016**, baw034.

Hermsen, R., de Ligt, J., Spee, W., Blokzijl, F., Schäfer, S., Adami, E., Boymans, S., Flink, S., van Boxtel, R., van der Weide, R. H. et al. (2015). Genomic landscape of rat strain and substrain variation. *BMC Genomics* **16**, 357.

Hindorff, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S. and Manolio, T. A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. USA* **106**, 9362-9367.

Hitzemann, R., Edmunds, S., Wu, W., Malmanger, B., Walter, N., Belknap, J., Darajian, P. and McWeeney, S. (2009). Detection of reciprocal quantitative trait loci for acute ethanol withdrawal and ethanol consumption in heterogeneous stock mice. *Psychopharmacology* **203**, 713-722.

Holmes, A., Li, Q., Murphy, D. L., Gold, E. and Crawley, J. N. (2003). Abnormal anxiety-related behavior in serotonin transporter null mutant mice: the influence of genetic background. *Genes Brain Behav.* **2**, 365-380.

Hubner, N., Wallace, C. A., Zimdahl, H., Petretto, E., Schulz, H., Maciver, F., Mueller, M., Hummel, O., Monti, J., Zidek, V. et al. (2005). Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. *Nat. Genet.* **37**, 243-253.

Iancu, O. D., Darajian, P., Walter, N. A. R., Malmanger, B., Oberbeck, D., Belknap, J., McWeeney, S. and Hitzemann, R. (2010). Genetic diversity and striatal gene networks: focus on the heterogeneous stock-collaborative cross (HS-CC) mouse. *BMC Genomics* **11**, 585.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C. and Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am. J. Psychiatry* **167**, 748-751.

Israely, I., Costa, R. M., Xie, C. W., Silva, A. J., Kosik, K. S. and Liu, X. (2004). Deletion of the neuron-specific protein delta-catenin leads to severe cognitive and synaptic dysfunction. *Curr. Biol.* **14**, 1657-1663.

Jagodic, M., Becanovic, K., Sheng, J. R., Wu, X., Bäckdahl, L., Lorentzen, J. C., Wallström, E. and Olsson, T. (2004). An advanced intercross line resolves Eae18 into two narrow quantitative trait loci syntenic to multiple sclerosis candidate loci. *J. Immunol.* **173**, 1366-1373.

Keane, T. M., Goodstadt, L., Danecek, P., White, M. A., Wong, K., Yalcin, B., Heger, A., Agam, A., Slater, G., Goodson, M. et al. (2011). Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477**, 289-294.

Khisti, R. T., Wolstenholme, J., Shelton, K. L. and Miles, M. F. (2006). Characterization of the ethanol-deprivation effect in substrains of C57BL/6 mice. *Alcohol* **40**, 119-126.

Kim, H. and Kim, J.-S. (2014). A guide to genome engineering with programmable nucleases. *Nat. Rev. Genet.* **15**, 321-334.

Kirkpatrick, S. L. and Bryant, C. D. (2014). Behavioral architecture of opioid reward and aversion in C57BL/6 substrains. *Front. Behav. Neurosci.* **8**, doi:10.3389/fnbeh.2014.00450.

Krohn, J., Speed, D., Palme, R., Touma, C., Mott, R. and Flint, J. (2014). Genetic interactions with sex make a relatively small contribution to the heritability of complex traits in mice. *PLoS ONE* **9**, e96450.

Kumar, V., Kim, K., Joseph, C., Kourrich, S., Yoo, S.-H., Huang, H. C., Vitaterna, M. H., de Villena, F. P.-M., Churchill, G., Bonci, A. et al. (2013). C57BL/6N mutation in cytoplasmic FMRP interacting protein 2 regulates cocaine response. *Science* **342**, 1508-1512.

Legare, M. E., Bartlett, F. S. and Frankel, W. N. (2000). A major effect QTL determined by multiple genes in epileptic EL mice. *Genome Res.* **10**, 42-48.

Lencz, T., Knowles, E., Davies, G., Guha, S., Liewald, D. C., Starr, J. M., Djurovic, S., Melle, I., Sundet, K., Christoforou, A. et al. (2014). Molecular

genetic evidence for overlap between general cognitive ability and risk for schizophrenia: a report from the Cognitive Genomics consorTium (COGENT). *Mol. Psychiatry* **19**, 168-174.

Lichtenstein, P., Yip, B. H., Björk, C., Pawitan, Y., Cannon, T. D., Sullivan, P. F. and Hultman, C. M. (2009). Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* **373**, 234-239.

Lo, C.-L., Lossie, A. C., Liang, T., Liu, Y., Xuei, X., Lumeng, L., Zhou, F. C. and Muir, W. M. (2016). High resolution genomic scans reveal genetic architecture controlling alcohol preference in bidirectionally selected rat model. *PLoS Genet.* **12**, e1006178.

Logan, R. W., Robledo, R. F., Recla, J. M., Philip, V. M., Bubier, J. A., Jay, J. J., Harwood, C., Wilcox, T., Gatti, D. M., Bult, C. J. et al. (2013). High-precision genetic mapping of behavioral traits in the diversity outbred mouse population. *Genes Brain Behav.* **12**, 424-437.

Mao, J.-H., Langley, S. A., Huang, Y., Hang, M., Bouchard, K. E., Celniker, S. E., Brown, J. B., Jansson, J. K., Karpen, G. H. and Snijders, A. M. (2015). Identification of genetic factors that modify motor performance and body weight using Collaborative Cross mice. *Sci. Rep.* **5**, 16247.

Markou, A., Chiamulera, C., Geyer, M. A., Tricklebank, M. and Steckler, T. (2009). Removing obstacles in neuroscience drug discovery: the future path for animal models. *Neuropsychopharmacology* **34**, 74-89.

McClearn, G. E., Wilson, J. R. and Meredith, W. (1970). The use of isogenic and heterogenic mouse stocks in behavioral research. In *Contributions to behavior-genetic analysis: The mouse as a prototype*, pp. 3-22. New York: Appleton-Century-Crofts.

McDermott-Roe, C., Ye, J., Ahmed, R., Sun, X.-M., Serafin, A., Ware, J., Bottolo, L., Muckett, P., Cañas, X., Zhang, J. et al. (2011). Endonuclease G is a novel determinant of cardiac hypertrophy and mitochondrial function. *Nature* **478**, 114-118.

Mehrabian, M., Allayee, H., Stockton, J., Lum, P. Y., Drake, T. A., Castellani, L. W., Suh, M., Armour, C., Edwards, S., Lamb, J. et al. (2005). Integrating genotypic and expression data in a segregating mouse population to identify 5-lipoxygenase as a susceptibility gene for obesity and bone traits. *Nat. Genet.* **37**, 1224-1233.

Moreno-Moral, A. and Petretto, E. (2016). From integrative genomics to systems genetics in the rat to link genotypes to phenotypes. *Dis. Model. Mech.* **9**, 1097-1110.

Morgan, A. P. and Welsh, C. E. (2015). Informatics resources for the Collaborative Cross and related mouse populations. *Mamm. Genome* **26**, 521-539.

Mott, R. and Flint, J. (2013). Dissecting quantitative traits in mice. *Annu. Rev. Genomics Hum. Genet.* **14**, 421-439.

Mott, R., Talbot, C. J., Turri, M. G., Collins, A. C. and Flint, J. (2000). A method for fine mapping quantitative trait loci in outbred animal stocks. *Proc. Natl. Acad. Sci. USA* **97**, 12649-12654.

Mulligan, M. K., Ponomarev, I., Boehm, S. L., Owen, J. A., Levin, P. S., Berman, A. E., Blednov, Y. A., Crabbe, J. C., Williams, R. W., Miles, M. F. et al. (2008). Alcohol trait and transcriptional genomic analysis of C57BL/6 substrains. *Genes Brain Behav.* **7**, 677-689.

Nicod, J., Davies, R. W., Cai, N., Hassett, C., Goodstadt, L., Cosgrove, C., Yee, B. K., Lionikaite, V., McIntyre, R. E., Remme, C. A. et al. (2016). Genome-wide association of multiple complex traits in outbred mice by ultra-low-coverage sequencing. *Nat. Genet.* **48**, 912-918.

Nivard, M. G., Mbarek, H., Hottenga, J. J., Smit, J. H., Jansen, R., Penninx, B. W., Middeldorp, C. M. and Boomsma, D. I. (2014). Further confirmation of the association between anxiety and CTNND2: replication in humans. *Genes Brain Behav.* **13**, 195-201.

Noben-Trauth, K., Latoche, J. R., Neely, H. R. and Bennett, B. (2010). Phenotype and genetics of progressive sensorineural hearing loss (Snhl1) in the LXS set of recombinant inbred strains of mice. *PLoS ONE* **5**, e11459.

Owen, M. J. (2014). New approaches to psychiatric diagnostic classification. *Neuron* **84**, 564-571.

Pandey, A. K. and Williams, R. W. (2014). Genetics of gene expression in CNS. *Int. Rev. Neurobiol.* **116**, 195.

Parker, C. C. and Palmer, A. A. (2011). Dark matter: are mice the solution to missing heritability?. *Front. Genet.* **2**, 32.

Parker, C. C., Cheng, R., Sokoloff, G. and Palmer, A. A. (2012). Genome-wide association for methamphetamine sensitivity in an advanced intercross mouse line. *Genes Brain Behav.* **11**, 52-61.

Parker, C. C., Gopalakrishnan, S., Carbonetto, P., Gonzales, N. M., Leung, E., Park, Y. J., Aryee, E., Davis, J., Blizard, D. A., Ackert-Bicknell, C. L. et al. (2016). Genome-wide association study of behavioral, physiological and gene expression traits in outbred CFW mice. *Nat. Genet.* **48**, 919-926.

Parsons, M. J., Grimm, C., Paya-Cano, J. L., Fernandes, C., Liu, L., Philip, V. M., Chesler, E. J., Nietfeld, W., Lehrach, H. and Schalkwyk, L. C. (2012). Genetic variation in hippocampal microRNA expression differences in C57BL/6 J X DBA/2 J (BXD) recombinant inbred mouse strains. *BMC Genomics* **13**, 476.

Peirce, J. L., Derr, R., Shendure, J., Kolata, T. and Silver, L. M. (1998). A major influence of sex-specific loci on alcohol preference in C57Bl/6 and DBA/2 inbred mice. *Mamm. Genome* **9**, 942-948.

Peirce, J. L., Lu, L., Gu, J., Silver, L. M. and Williams, R. W. (2004). A new set of BXD recombinant inbred lines from advanced intercross populations in mice. *BMC Genet.* **5**, 7.

Peters, L. L., Robledo, R. F., Bult, C. J., Churchill, G. A., Paigen, B. J. and Svenson, K. L. (2007). The mouse as a model for human biology: a resource guide for complex trait analysis. *Nat. Rev. Genet.* **8**, 58-69.

Philip, V. M., Sokoloff, G., Ackert-Bicknell, C. L., Striz, M., Branstetter, L., Beckmann, M. A., Spence, J. S., Jackson, B. L., Galloway, L. D., Barker, P. et al. (2011). Genetic analysis in the Collaborative Cross breeding population. *Genome Res.* **21**, 1223-1238.

Pravenec, M., Zidek, V., Landa, V., Simakova, M., Mlejnek, P., Kazdova, L., Bila, V., Krenova, D. and Kren, V. (2004). Genetic analysis of "metabolic syndrome" in the spontaneously hypertensive rat. *Physiol. Res.* **53**, S15-S22.

Printz, M. P., Jirout, M., Jaworski, R., Alemayehu, A. and Kren, V. (2003). Invited Review: HXB/BXH rat recombinant inbred strain platform: a newly enhanced tool for cardiovascular, behavioral, and developmental genetics and genomics. *J. Appl. Physiol.* **94**, 2510-2522.

Psychiatric GWAS Consortium and **Bipolar Disorder Working Group** (2011). Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat. Genet.* **43**, 977-983.

Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., Sullivan, P. F., Sklar, P., Ruderfer, D. M., McQuillin, A., Morris, D. W. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748-752.

Putman, A. H., Wolen, A. R., Harenza, J. L., Yordanova, R. K., Webb, B. T., Chesler, E. J. and Miles, M. F. (2016). Identification of quantitative trait loci and candidate genes for an anxiolytic-like response to ethanol in BXD recombinant inbred strains. *Genes Brain Behav.* **15**, 367-381.

Rask-Andersen, M., Almén, M. S. and Schiöth, H. B. (2015). Scrutinizing the FTO locus: compelling evidence for a complex, long-range regulatory context. *Hum. Genet.* **134**, 1183-1193.

Rasmussen, A. L., Okumura, A., Ferris, M. T., Green, R., Feldmann, F., Kelly, S. M., Scott, D. P., Safronetz, D., Haddock, E., LaCasse, R. et al. (2014). Host genetic diversity enables Ebola hemorrhagic fever pathogenesis and resistance. *Science* **346**, 987-991.

Rat Genome Sequencing and Mapping Consortium (2013). Combined sequence-based and genetic mapping analysis of complex traits in outbred rats. *Nat. Genet.* **45**, 767-775.

Reddy, S. Y., Rasmussen, N. A., Fourie, N. H., Berger, R. S., Martino, A. C., Gill, J., Longchamps, R., Wang, X. M., Heitkemper, M. M. and Henderson, W. A. (2014). Sleep quality, BDNF genotype and gene expression in individuals with chronic abdominal pain. *BMC Med. Genomics* **7**, 61.

Robinson, E. B., St Pourcain, B., Anttila, V., Kosmicki, J. A., Bulik-Sullivan, B., Grove, J., Maller, J., Samocha, K. E., Sanders, S. J., Ripke, S. et al. (2016). Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nat. Genet.* **48**, 552-555.

Sander, J. D. and Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* **32**, 347-355.

Schadt, E. E., Monks, S. A., Drake, T. A., Lusis, A. J., Che, N., Colinayo, V., Ruff, T. G., Milligan, S. B., Lamb, J. R., Cavet, G. et al. (2003). Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422**, 297-302.

Sheng, J. R., Jagodic, M., Dahlman, I., Becanovic, K., Nohra, R., Marta, M., Iacobaeus, E., Olsson, T. and Wallström, E. (2005). Eae19, a new locus on rat chromosome 15 regulating experimental autoimmune encephalomyelitis. *Genetics* **170**, 283-289.

Shimoyama, M., De Pons, J., Hayman, G. T., Laulederkind, S. J., Liu, W., Nigam, R., Petri, V., Smith, J. R., Tutaj, M., Wang, S.-J. et al. (2014). The Rat Genome Database 2015: genomic, phenotypic and environmental variations and disease. *Nucleic Acids Res.* **43**, D743-D750.

Simon, M. M., Greenaway, S., White, J. K., Fuchs, H., Gailus-Durner, V., Wells, S., Sorg, T., Wong, K., Bedu, E., Cartwright, E. J. et al. (2013). A comparative phenotypic and genomic analysis of C57BL/6J and C57BL/6N mouse strains. *Genome Biol.* **14**, R82.

Sittig, L. J., Carbonetto, P., Engel, K. A., Krauss, K. S., Barrios-Camacho, C. M. and Palmer, A. A. (2016). Genetic background limits generalizability of genotype-phenotype relationships. *Neuron* **91**, 1253-1259.

Stuchlik, A., Hatalová, H., Grzyb, A. N., Petrasek, T., Prokopova, I., Overall, R., Silhavy, J., Zidek, V., Kempermann, G. and Vales, K. (2012). P-381-Genetic

and correlation analysis of spatial learning of the rat hxb/bxh recombinant inbred strains in the carousel maze, a spatial avoidance paradigm. *Eur. Psychiatry* **27**, 1.

Sullivan, P. F., Daly, M. J. and O'Donovan, M. (2012). Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat. Rev. Genet.* **13**, 537-551.

Svenson, K. L., Gatti, D. M., Valdar, W., Welsh, C. E., Cheng, R., Chesler, E. J., Palmer, A. A., McMillan, L. and Churchill, G. A. (2012). High-resolution genetic mapping using the Mouse Diversity outbred population. *Genetics* **190**, 437-447.

Tabakoff, B., Saba, L., Printz, M., Flodman, P., Hodgkinson, C., Goldman, D., Koob, G., Richardson, H. N., Kechris, K., Bell, R. L. et al. (2009). Genetical genomic determinants of alcohol consumption in rats and humans. *BMC Biol.* **7**, 70.

Talbot, C. J., Nicod, A., Cherny, S. S., Fulker, D. W., Collins, A. C. and Flint, J. (1999). High-resolution mapping of quantitative trait loci in outbred mice. *Nat. Genet.* **21**, 305-308.

Threadgill, D. W., Miller, D. R., Churchill, G. A. and de Villena, F. P.-M. (2011). The collaborative cross: a recombinant inbred mouse population for the systems genetic era. *ILAR J.* **52**, 24-31.

Valdar, W., Solberg, L. C., Gauguier, D., Burnett, S., Klenerman, P., Cookson, W. O., Taylor, M. S., Rawlins, J. N. P., Mott, R. and Flint, J. (2006a). Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nat. Genet.* **38**, 879-887.

Valdar, W., Solberg, L. C., Gauguier, D., Cookson, W. O., Rawlins, J. N. P., Mott, R. and Flint, J. (2006b). Genetic and environmental effects on complex traits in mice. *Genetics* **174**, 959-984.

Vanderlinden, L. A., Saba, L. M., Printz, M. P., Flodman, P., Koob, G., Richardson, H. N., Hoffman, P. L. and Tabakoff, B. (2014). Is the alcohol deprivation effect genetically mediated? Studies with HXB/BXH recombinant inbred rat strains. *Alcoholism* **38**, 2148-2157.

Voigt, B., Kuramoto, T., Mashimo, T., Tsurumi, T., Sasaki, Y., Hokao, R. and Serikawa, T. (2008). Evaluation of LEXF/FXLE rat recombinant inbred strains for genetic dissection of complex traits. *Physiol. Genomics* **32**, 335-342.

Williams, E. G. and Auwerx, J. (2015). The convergence of systems and reductionist approaches in complex trait analysis. *Cell* **162**, 23-32.

Williams, R. and Williams, E. (2016). Resources for systems genetics. In *Systems Genetics: Methods and Protocols. Methods in Molecular Biology 1488* (Eds. K. Schughart and R. W. Williams). New York: Humana.

World Health Organization (2004). *International Statistical Classification of Diseases and Related Health Problems*. Geneva: World Health Organization.

Yalcin, B. and Flint, J. (2012). Association studies in outbred mice in a new era of full-genome sequencing. *Mamm. Genome* **23**, 719-726.

Yalcin, B., Willis-Owen, S. A. G., Fullerton, J., Meesaq, A., Deacon, R. M., Rawlins, J. N. P., Copley, R. R., Morris, A. P., Flint, J. and Mott, R. (2004). Genetic dissection of a behavioral quantitative trait locus shows that Rgs2 modulates anxiety in mice. *Nat. Genet.* **36**, 1197-1202.

Yalcin, B., Nicod, J., Bhomra, A., Davidson, S., Cleak, J., Farinelli, L., Østerås, M., Whitley, A., Yuan, W., Gan, X. et al. (2010). Commercially available outbred mice for genome-wide association studies. *PLoS Genet.* **6**, e1001085.

Yalcin, B., Wong, K., Agam, A., Goodson, M., Keane, T. M., Gan, X., Nellåker, C., Goodstadt, L., Nicod, J., Bhomra, A. et al. (2011). Sequence-based characterization of structural variation in the mouse genome. *Nature* **477**, 326-329.

Ye, R., Carneiro, A. M. D., Airey, D., Sanders-Bush, E., Williams, R. W., Lu, L., Wang, J., Zhang, B. and Blakely, R. D. (2014). Evaluation of heritable determinants of blood and brain serotonin homeostasis using recombinant inbred mice. *Genes Brain Behav.* **13**, 247-260.

Yoshimi, K., Kaneko, T., Voigt, B. and Mashimo, T. (2014). Allele-specific genome editing and correction of disease-associated phenotypes in rats using the CRISPR–Cas platform. *Nat. Commun.* **5**, 4240.

Young, E. E., Bryant, C. D., Lee, S. E., Peng, X., Cook, B., Nair, H. K., Dreher, K. J., Zhang, X., Palmer, A. A., Chung, J. M. et al. (2016). Systems genetic and pharmacological analysis identifies candidate genes underlying mechanosensation in the von Frey test. *Genes Brain Behav.* **15**, 604-615.

Zhang-James, Y., Middleton, F. A. and Faraone, S. V. (2013). Genetic architecture of Wistar-Kyoto rat and spontaneously hypertensive rat substrains from different sources. *Physiol. Genomics* **45**, 528-538.

Disease Models & Mechanisms

Supplementary Table S1. Tractable QTLs for neurobehavioral traits mapped in rodent resources.

Click here to Download Table S1