Genetic contributions to behavioural diversity at the gene–environment interface

Andres Bendesky and Cornelia I. Bargmann

Abstract | Recent work on behavioural variation within and between species has furthered our understanding of the genetic architecture of behavioural traits, the identities of relevant genes and the ways in which genetic variants affect neuronal circuits to modify behaviour. Here we review our understanding of the genetics of natural behavioural variation in non-human animals and highlight the implications of these findings for human genetics. We suggest that gene–environment interactions are central to natural genetic variation in behaviour and that genes affecting neuromodulatory pathways and sensory processing are preferred sites of naturally occurring mutations.

"All animals are equal but some animals are more equal than others." George Orwell

The word 'behaviour' refers to all observable actions of animals, including humans, ranging from simple reflex actions to complex behavioural sequences or patterns (BOX 1). A single behaviour represents a combined response to external stimuli and internal motivational states, as interpreted by the brain and modified by prior experience. Genetic systems that underlie sensory processing, emotion and motivation, neuronal development and plasticity are all essential for generating behaviours; these genetic programs vary between and within species, giving rise to behavioural diversity.

Many single-gene mutations that affect animal behaviours have been identified through classical genetic screens and knockout mutants, and single-gene mutations can affect human behaviours as well. A notable example of the universality and specificity of single behavioural genes is provided by the *period* gene and other genes that regulate circadian behaviours; these were first identified through induced mutations in *Drosophila melanogaster*[1] and were then found to be mutated in humans with advanced sleep phase syndrome (ASPS), a rare single-gene circadian disorder[2,3]. Other rare human mutations of large effect can give rise to specific syndromes, such as overeating and obesity (caused by mutations affecting the leptin receptor)[4] and narcolepsy with cataplexy (caused by mutations in hypocretin)[5]. In agreement with the idea that behavioural genes can have conserved functions in animals and humans, these genes were originally identified through mutations in mice and dogs[6–8]. Despite these striking examples, however, most common genetic variation in human and animal behaviour cannot be explained by known single-gene mutations.

In this Review, we present a current understanding of common genetic variation that gives rise to behavioural diversity. We describe conceptual insights and molecular discoveries from studies of genetic variation within species and between closely related species. Recent technological advances that are driven by genome sequencing allow the identification of genetic markers in any species, including humans; these methods, coupled with high-throughput genotyping, are facilitating advances in genetic mapping of behavioural traits. We first describe the evidence for a complex genetic architecture for most behavioural traits and then describe representative studies that move from a behaviour to specific molecules. Gene–environment interactions are an essential theme of behavioural genetics, and we use specific examples to show how the gene–environment interface illuminates the nature of behavioural variation. The next step for the field is understanding how the brain translates genetic changes into behaviour, and initial examples have taken steps in this direction. Finally, we argue that certain gene classes, including sensory genes and genes that affect neuromodulatory systems, are disproportionately associated with variation in behaviour because of their evolvability. We discuss these ideas in the context of balancing selection that shapes and maintains behavioural variation in nature.

Plasticity
The ability of nervous systems to change molecularly, physiologically or anatomically based on experience.

Evolvability
The ability of organisms to respond to selective pressures with adaptive genetic changes. By analogy, a gene's propensity to acquire adaptive mutations under selective pressures. Among other properties, evolvability is affected by the mutation rate, directed mutation, the degree of pleiotropy and epistasis and system robustness.

Howard Hughes Medical Institute and Laboratory for Neural Circuits and Behavior, The Rockefeller University, New York, New York 10065, USA.
Correspondence to C.I.B.
e-mail: cori@rockefeller.edu
doi:10.1038/nrg3065
Published online
8 November 2011

Box 1 | **The nature of behaviour and challenges in the study of behaviour**

	Simple component	Complex pattern or sequence
Feeding behaviour	Eating	Food selection and preference
Defensive behaviour	Escape	Behavioural suppression, hiding
Habitat selection	Chemotaxis or thermotaxis	Nest building, exploration
Reproductive behaviour	Mating	Courtship song or dance
Other social interaction	Aggregation	Territorial defence, migration

Behaviours can be divided into different categories based on their complexity or their purpose. A classical ethological approach might classify behaviours based on whether they are related to food acquisition, predator avoidance, habitat selection, reproduction or other social interactions (see the table for examples of each class). A neurobiological approach might classify behaviours based on the extent to which they are learned or based on the anatomical brain systems involved; this analysis would be orthogonal to the ethological one.

A typical behaviour involves a set of connected actions that take place over a period of time. One example would be the immediate withdrawal of a paw from a hot surface, a rapid retreat from the area and licking of the paw, followed by long-term avoidance of the area. A sequence of behavioural actions is rarely as stereotyped as suggested by this example, however, and even simple animals show a range of actions and sequences during a behaviour such as escape, mating or grooming. To capture this variability, over the past decade, the analysis of animal behaviour has become more sophisticated with the increasing use of high-throughput automated systems for behavioural monitoring combined with statistical analysis of behavioural events[116–118]. These tools are particularly useful for genetic analysis, for which quantitative data must be gathered for many individuals. Careful and accurate behavioural measurements are crucial to the genetic analysis of behaviour.

The special challenge in understanding behaviour is that external and internal variables affecting behaviour change over time, and a common response is only expected when all variables are held constant. As a result, gene–environment interactions are prominent features of behavioural variation. Variation is a property of many biological systems, but an animal's morphology is much more stable over time than its behaviour. Variation is essential to behaviour, not peripheral.

Although behavioural variation is conceptually interesting, it generates challenges both intellectually and technically. Natural behavioural variation exists on a continuum, and genetic variants change the probability of certain behaviours in a quantitative way, not a qualitative way. To complicate the issue, behavioural measurements can substantially differ between laboratories, even when care is taken to standardize every aspect of an experiment[119]. Even within a laboratory, anxiety measurements in mice are strongly influenced by the person handling the animals[120]. Moreover, even when all variables are held constant, a behavioural response is often probabilistic (which might have a selective advantage) rather than deterministic[121]. With these factors in mind, it is clear why it is an art to develop behavioural assays that are both specific and sensitive.

Balancing selection
Selection that maintains trait variation. Two alleles can be balanced if a heterozygote is more successful than either homozygote, if each of the two alleles is better-adapted to one of two alternative environments or if each allele promotes a different, but equally successful, survival strategy in the same environment.

Ethological approach
The study of animal behaviours motivated by their observation in nature.

Genetic architecture
The number, frequency, effect size, dominance relationship and interactions of genetic variants that affect a trait in populations of a species.

Linkage-based mapping
A genetic mapping technique that uses pedigree information and genetic markers to link a trait to a genomic location.

QTL mapping
QTL mapping of progeny from an intercross measures the correlation between trait values and DNA markers across the genome and infers the number of loci that affect the trait, their location and the contribution of each locus to the total trait variance.

Genetic association
A population-based mapping technique that measures the correlation between a DNA polymorphism and a trait.

Recombinant inbred lines
(RILs). Strains that are derived from crosses between two or more parental strains, followed by recombination of chromosomes and inbreeding to homozygosity. Typically, RILs are carefully genotyped at many loci. A panel of RILs can be a stable resource for QTL mapping.

A complex genetic basis for behavioural traits

A genetic contribution to behavioural variation has long been recognized both in animals and in humans. The strongest evidence for genetic effects on human behaviour comes from twin and adoption studies that demonstrate correlations between genetic relatedness and the risk for psychiatric disorders (BOX 2). To move from this general insight to a molecular basis of behavioural trait variation, one must first define the genetic architecture of a trait and then map the genes. For most laboratory organisms, linkage-based mapping techniques, such as QTL mapping, are used. In outbred populations, such as humans and wild animals, there is also the option of using genetic association approaches, provided that many individuals and genetic markers can be examined (for example, genome-wide association studies (GWASs)). In general, behavioural mapping methods are similar to those used for non-behavioural traits, but special care must be taken to minimize measurement noise (BOX 1).

QTL mapping of behaviours. QTL analysis between inbred strains (FIG. 1) has been used to map trait differences in: learning, fear, anxiety, circadian rhythm, responses to addictive drugs and activity levels in mice; fear, anxiety and responses to addictive drugs in rats; and olfactory behaviour, mating behaviour and locomotor reactivity in *D. melanogaster*[9–12]. The advantage of QTL mapping using defined crosses is that a sufficiently large collection of F2 generation animals or recombinant inbred lines (RILs) allows each locus to be tested rigorously and can detect genetic interactions (epistasis). The disadvantage is that discovery is limited to those alleles that vary between the two starting strains.

A modified QTL approach that is directed at capturing broader population variation starts from a pool of parental strains, not just two strains. In *D. melanogaster,* a collection of wild flies from the Raleigh Farmers' Market, North Carolina, United States, has been used to generate 192 inbred lines representing broader genetic variations from wild populations (see the *Drosophila* Genetic Reference Panel website). In mice, a Collaborative Cross has been structured to capture variation from eight different mouse strains in a pool of recombinant inbred lines[13]. In all QTL crosses, establishing stable inbred strains that can be genotyped once and then tested for many phenotypic traits provides an immense increase in experimental power; such carefully constructed strains exist for *Caenorhabditis elegans*[14], *D. melanogaster* (see the *Drosophila* Genetic Reference Panel website) and mice[13].

Box 2 | The genetics of human psychiatric disease: observations and mysteries

The relatives of psychiatric patients have an elevated risk of similar disorders, and the degree of risk correlates with the degree of genetic relatedness. The monozygotic twin of a schizophrenic patient has about a 50% risk of schizophrenia[122], which represents an enormous increase over the population risk of ~1%; however, the fact that the risk in monozygotic twins is less than 100% demonstrates the existence of non-genetic components. By the same formulation, an affected monozygotic twin with a disorder predicts a ~60% risk for autism[123,124] and a ~40% risk for bipolar disease[125–127], anxiety disorders[128] or depression[129,130].

A puzzling aspect of these studies appears when the risk to first-degree relatives is examined (dizygotic twins, other siblings, parents and children of affected individuals). For anxiety and depression, the risk to these individuals is about half that of the monozygotic twin[128,129,130]. This relationship matches theories of additive variation, in which multiple alleles act independently of one another to influence risk. In autism, schizophrenia and bipolar disorder, however, the risk to first-degree relatives is much lower than half that of the monozygotic twin[122–127]. For these disorders, the genetic risk may reflect new mutations and combinations of risk alleles with nonlinear interactions. In addition, dizygotic twins have a higher risk than other first-degree relatives, an observation that might reflect prenatal environment or other developmental risk factors. Genome-wide association studies indicate that common genetic variants contribute to the risk of schizophrenia and bipolar disorder but that such variants are less important than they are in other complex non-psychiatric diseases[131–134] (but also see REF. 135 for an alternative viewpoint).

An unknown but growing fraction of schizophrenia and autism cases is associated with *de novo* mutations or rare transmitted mutations, often copy number variants (CNVs), which can cause a large increase in risk[136–142]. Each identified high-risk variant is present in at most 1% of patients, indicating that one disease can result from hundreds of different genetic causes. In other words, autism may not be a single, relatively common disorder; it may represent 100 rare disorders. Adding to this complexity, several SNPs and a rare single-gene risk factor (disrupted in schizophrenia 1 (*DISC1*)) increase risk for both schizophrenia and bipolar disorder, suggesting that shared genetic factors influence multiple disorders[131,132,143,144]. These results suggest that genetic causality and heterogeneity will challenge the existing classification categories for psychiatric disorders.

QTL approaches in rodents and flies unambiguously show that the genetic architecture of behavioural traits is complex. For example, a mouse QTL analysis of a single pair of inbred strains led to the identification of at least 16 distinct emotionality loci[10]. Moreover, a different set of QTLs was found for the same emotionality behaviour in different strains of mice[10]. Similar results from studies of locomotor activity and aggression in *D. melanogaster* support a similarly complex architecture[12,15]. Importantly, QTLs do not contribute equally to a phenotype. An analysis of over 200 behavioural QTLs affecting 20 different traits in mice and rats demonstrated that the effect size of QTLs is exponentially distributed: ~10% of the QTLs had large individual effects accounting for 10–20% of trait variance and a large number of loci contributed increasingly smaller effects[10,16]. The traits characterized in this analysis included motor activity, learning, emotionality traits and drug-related behaviours.

Linkage studies with introgression strains show large effects on behaviours. An approach that is complementary to QTL analysis is analysing introgression strains (FIG. 1; BOX 3). This approach is particularly powerful when the whole genome is covered in a panel of chromosome substitution strains (CSSs) or congenic strains; such panels have been developed in *C. elegans*[17,18], *D. melanogaster*[19], rats[20] and mice[21–23].

CSSs and congenic strains have been used to characterize fear and anxiety-related traits in mice and, in agreement with QTL crosses, they indicate that multiple loci contribute to these behavioural traits[24,25]. However, individual chromosomes can have large effects, sometimes accounting for half of the trait difference between the two parental strains[24] — a greater effect than is inferred for any single QTL in classical QTL mapping. Even more remarkably, many chromosomes from a single strain can have large effects, such that in combination they 'account' for more than 100% of the difference between the two starting strains[26]. The larger apparent effect of single chromosomes in a CSS is partly due to the statistical structure of the experiment. A QTL cross with multiple segregating loci is used to explain the segregating variance in a trait, which must add up to 100%. By contrast, studies of CSSs describe the effect of a genetic region on the mean trait value and more closely match the intuitive concept of effect size. In addition, QTL analysis detects the average effect of a variant across many different genetic backgrounds, including those in which epistatic interactions obscure the effects of the QTL, whereas chromosome substitution interrogates the variant in a single background. Many genetic effects are background-dependent; for example, the viability of certain gene knockouts in mice and yeast varies by strain[27,28]. Epistatic interactions among behavioural QTLs are well-recognized and have been detected in courtship, foraging, locomotion, learning and aggressive behaviours in insects[15,29–31], in exploratory behaviour in nematodes[32] and in fear and anxiety traits in mice[25].

Lessons for human behaviour. Based on animal studies, we expect any individual human behavioural trait to be affected by many different genetic variants. Across the entire population, individual variants are likely to have small to moderate effect sizes. However, the introgression strain results imply that in any single genetic background, a particular genetic variant may have a large effect that is lost when averaging over entire populations, because epistasis can decrease the impact of a variant in some individuals. At a practical level, this conclusion suggests that studies combining family-based designs with GWAS or sequencing approaches will help in finding causative genetic variants[33]. GWASs can detect many variants but, like QTL studies, the effect size of the variants will be diluted by genetic heterogeneity in the population. Human families are not as inbred as CSS strains, but their genetic backgrounds are considerably less heterogeneous than those of whole populations.

From behaviour to genes

Hundreds of QTLs that affect animal behaviour have been detected in genetic crosses, but the specific genes and gene variants that correspond to the QTL are just beginning to emerge from focused mapping approaches (described in BOX 3 and FIG. 2). The reduced cost of genome sequencing and the existence of well-characterized mapping strains are tools that should enable gene identification of many QTLs. The number of QTLs that

Emotionality
A set of fear- and anxiety-related behaviours, such as avoidance of exposed areas and inhibition of movement after foot shock.

Introgression strains
Strains into which defined DNA segments have been introduced from a different strain background through backcrossing. The introduced segments can be full chromosomes, as in chromosome substitution strains (CSSs) or smaller chromosomal intervals, as in congenics.

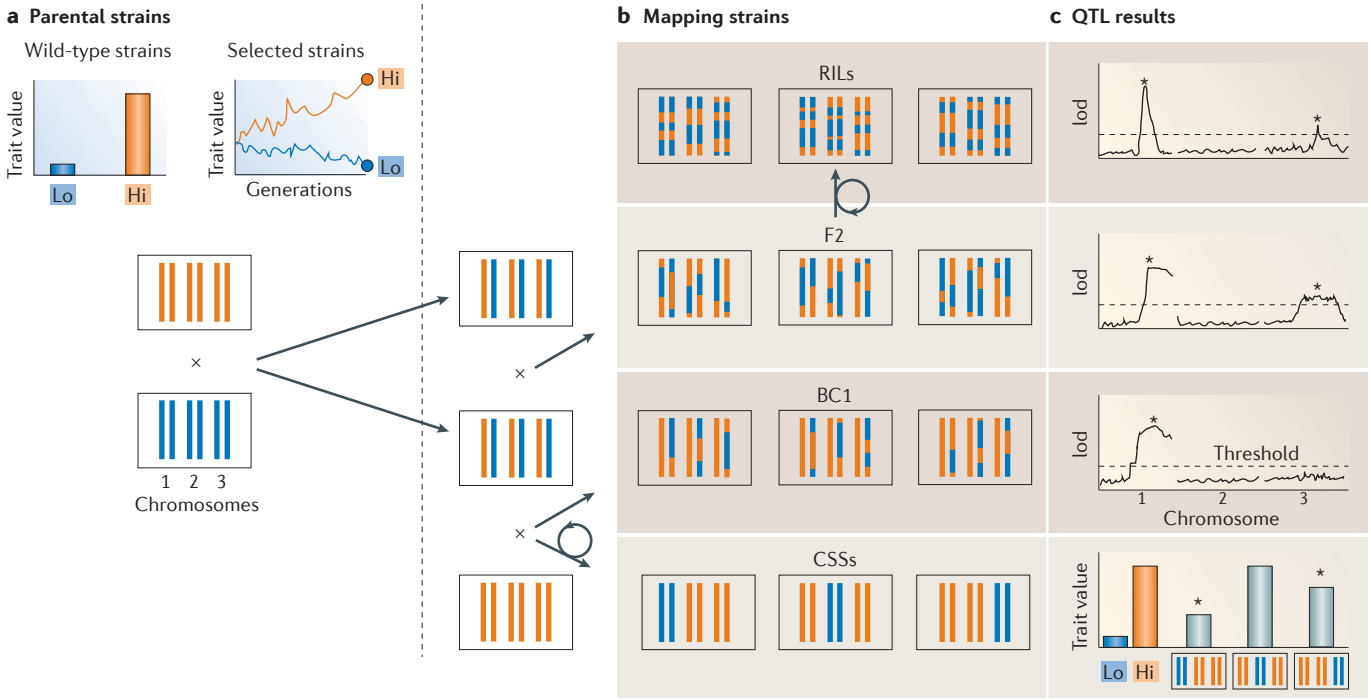


Figure 1 | **Strategies for linkage-based mapping.** The parental strains, called Hi and Lo in this example, can be wild-type strains that differ in the trait of interest (part **a**, top left) or strains that were selectively bred to differ in the trait, starting from a heterogeneous population (part **a**, top right). Three pairs of chromosomes (numbered 1–3) from the hypothetical Hi and Lo diploid individuals are shown in orange and blue, respectively. In this example, there are two additive QTLs, one of large effect on chromosome 1, with a dominant Lo allele, and another of small effect on chromosome 3, with a recessive Lo allele. Characterizing recombinant inbred lines (RILs) or the F2 generation (first and second rows, respectively, part **b**) identifies both QTLs (indicated by asterisks in all rows in part **c**). F2 gives information about the dominance of the QTLs, but RILs provide better resolution. Mapping with a backcross (BC1; third row, part **b**) can only identify QTLs that are recessive in the parental strain used for the backcross — Hi in this example. Only three strains for the RIL, F2 and BC1 are shown, but usually dozens to hundreds are used. Chromosome substitution strains (CSSs; fourth row, part **b**) identify both QTLs but give no indication of their location within the chromosome or their dominance. Arrows with circles indicate repeated crosses of the same type. lod, logarithm of the odds ratio.

have been definitely assigned to specific genes remains low, but encouraging initial results suggest that behavioural genes that have been identified through unbiased approaches can have conserved functions, leading to new insights into behaviour. Some illustrative examples are described below.

The first mapped behavioural gene,* foraging*, is conserved across species. *D. melanogaster* larvae fall into two groups based on their foraging strategy: rovers (which make up ~70% of the population) and sitters. Rovers travel greater distances in the presence of food, disperse more readily between food patches and pupate farther from the food supply than sitters do[34,35]. *D. melanogaster* larval foraging was the first naturally varying behaviour to be mapped to a specific gene of major effect, *foraging* (*for*)[36,37], which encodes the conserved cGMP-regulated protein kinase G (PKG). Rovers have more *for* mRNA in their brain than sitters and higher PKG activity[37], suggesting that the rover allele is a high-activity allele. Among the targets of PKG regulation are ion channels that regulate neuronal excitability[38].

The discovery of *for* in flies led to the realization that PKG is a prominent regulator of invertebrate behaviour. For example, young adult honeybees engage in nursing activities in the hive, and older honeybees become foragers that leave the hive to retrieve food. Transcript levels and activity of the bee *for* orthologue are higher in foragers than in nurses, and stimulation of PKG activity with cGMP analogues can induce premature foraging behaviours in young bees[39]. These results suggest that developmental regulation of PKG in an individual honeybee modifies its behaviour. Another insect species in which different behavioural forms vary in PKG activity is the red harvester ant, in which the *for* orthologue is expressed at lower levels in foragers than in nest workers (the opposite pattern from honeybees)[40]. The influence of PKG on behavioural variation extends to a nematode worm, *Pristionchus pacificus,* which is attracted to insect hosts by their pheromones. Natural variation in *P. pacificus* responses to insect pheromones maps to *egg laying defective 4* (*egl-4*), a homologue of the *for* gene[41]. These results show that PKG is a hotspot for different kinds of variation: genetic variation in PKG distinguishes different flies and nematodes, and stage- or caste-specific PKG modulation affects behaviour in individual honeybees and ants. Evidence for other behavioural 'hotspot' genes is addressed below.

lod
Logarithm of the odds ratio. A term that indicates the likelihood that a genomic region is linked to the trait being measured. A genome-wide lod threshold is set to correct for multiple comparisons.

Box 3 | **Functional validation of quantitative trait genes and nucleotides**

From QTLs to genes

The strategy for moving from a QTL to a gene depends on the organism and on the complexity of the trait. One strategy is introgression, in which progressively smaller regions spanning a QTL are crossed into a recipient genetic background until a single gene or variant can be defined (gene *a* in the example, shown with a dominant Lo allele; gene *b* is included for reference). Panels **Aa** and **Ab** show small genomic regions from the Lo background strain (blue) introgressed into the Hi background strain (orange) in genome-tagged strains (GTS1–4, panel **Aa**) and backcross (BC1–4) lines (panel **Ab**). A second strategy is linkage disequilibrium mapping of QTLs in outbred populations (OB1–4), which relies upon co-segregation of a QTL with a DNA marker through many recombination events over many generations (panel **Ac**). The plot in panel **Ac** shows gene *a* falling above a probability threshold for being associated with a trait. Linkage disequilibrium mapping is also the basis of genome-wide association studies (GWASs) in humans.

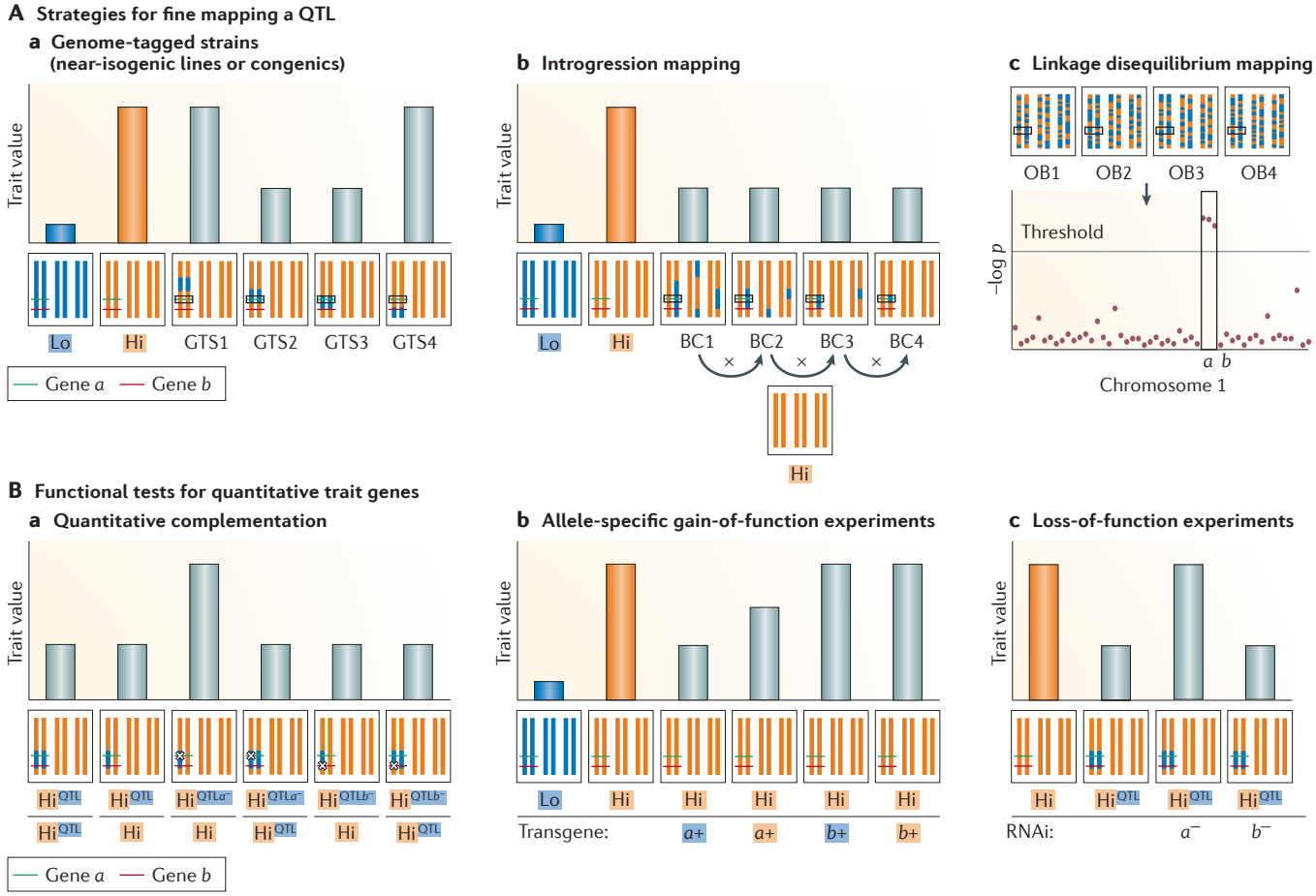
An alternative to fine mapping is searching for mutations in candidate genes. Many candidate behavioural genes that appeared to be promising in initial studies have failed to be replicated, however, and the consensus in the field is that unbiased approaches are necessary. A discovery-based approach for finding promising genes is genome-wide analysis of gene expression patterns, which can be combined with linkage approaches and mutation identification (FIG. 2).

From genes to quantitative trait nucleotides

The ultimate goal of mapping is the identification of the quantitative trait nucleotide (QTN) that affects the trait, which can be coding changes that affect a protein or proximal or distant non-coding changes that regulate gene expression. Specific genetic tests can define the relevance of QTNs and the affected quantitative trait gene (QTG; gene *a* in this example, which has a dominant Lo allele).

Quantitative complementation. The quantitative complementation test, first described in *Drosophila melanogaster*[145] (explained in detail in REF. 146), measures the extent to which null mutations complement the QTL allele with reduced activity (usually the recessive QTL allele). Strains that carry deletions or null mutations in genes within the QTL are crossed to both parental strains used for QTL mapping. An interaction between the QTL and the null mutation suggests that differential activity of the tested gene gives rise to behavioural variation. A null mutant can fail to complement because it is allelic to the QTG or because a mutation in another gene interacts with the QTG. To minimize the effect of multigenic interactions, advanced quantitative complementation is performed between strains with near-identical genetic backgrounds[15,32,88] (panel **Ba**): for example, by introgressing both the QTL and the null mutation into one parental strain.

Gain- and loss-of-function tests. Gain-of-function transgenes are usually introduced into the strain that carries the recessive allele of the QTL, as the recessive allele typically has reduced activity compared to the dominant QTL[37] (panel **Bb**). Transgenes can be made with DNA from both parental strains[32,61,88]; the expectation is that DNA from the strain with the dominant QTL will be more potent at affecting the behaviour than DNA from the recessive strain. If DNA from both strains rescues equally, the QTG may have been cloned, but the QTN may not be present in the tested transgene. Loss-of-function experiments in the dominant strain should provide the reciprocal answer to gain-of-function transgenes in the recessive strain (panel **Bc**). For example, RNAi of QTGs should usually transform the behaviour of the dominant strain in the direction of the recessive strain[32].



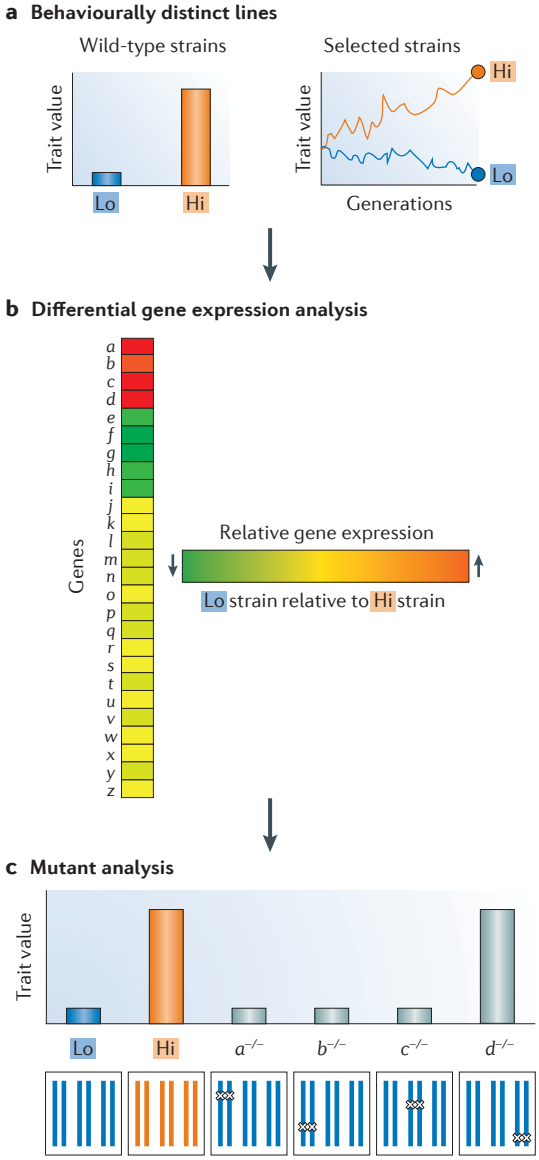


Figure 2 | **Differential gene expression analysis to identify genetic pathways mediating variation in behaviour.** Transcript abundance can be compared between two strains that either differ naturally in a behaviour of interest (**a**, left panel) or that are selected to differ at a trait (**a**, right panel). The role of genes that vary in expression between strains (**b**) can be further explored by testing mutants in those genes for altered behaviour (**c**). In this example, although genes *a*, *b*, *c* and *d* were expressed at higher levels in the Lo strain (**b**), only a null mutation of *d* affected the behaviour of the Lo strain (**c**). The technique can be applied to behaviourally selected strains alone or in combination with QTL analysis. This technique identifies genes with a biological role in behaviour, but does not identify QTLs per se. The genes that are differentially expressed in microarrays may be functional QTLs, or they may be indirect targets of the underlying genetic processes; linkage or identification of a mutation would be necessary to show that natural variation segregates at these loci. This strategy has been used to define molecular signatures and candidate genes associated with geotaxis, locomotion and aggression behaviours in *Drosophila melanogaster*[83,147–149].

Quantitative complementation test
A variant of the classical genetic complementation test that measures the interaction between genetic mutations and two naturally occurring alleles to determine whether the natural alleles are allelic to the mutants.

Mouse emotionality is affected by G protein modulation. Of the 450 behavioural QTLs that have been detected in mice (see the Mouse Genome Informatics website), the first to be compellingly mapped to a specific gene was regulator of G protein signalling 2 (*Rgs2*), which emerged from the studies of emotionality described above[42]. Fine mapping of the complex emotionality trait led to the fragmentation of one strong QTL into three neighbouring QTLs within the original region, a phenomenon that is commonly observed in quantitative genetics[15,32,43–46]. Genetic proof that *Rgs2* was a relevant locus for emotionality came from quantitative complementation tests[42] (BOX 3).

A knockout mutation of *Rgs2* shows high anxiety, supporting the idea that this gene regulates emotionality[42]. *Rgs2* and other RGS proteins shorten the duration of G-protein-coupled receptor (GPCR) signalling by stimulating GTP hydrolysis to inactivate heterotrimeric G proteins[47]. The effect of the *Rgs2* knockout on anxiety-related behaviours suggests that *Rgs2* limits signalling of GPCR pathways that produce anxiety. The nature of the relevant GPCR might be understood by studying other phenotypes of *Rgs2* knockouts, which include hypertension with evidence of increased sympathetic function and disruption of angiotensin and vasopressin GPCR signalling[48,49]. Both blood pressure and anxiety are strongly stress-responsive, suggesting that a common system regulated by *Rgs2* could limit stress responses in the brain and in peripheral tissues.

Emotionality in mice is a trait with similarities to human emotional traits, such as anxiety, neuroticism or emotional stability. This suggests that further analysis of the mouse trait could shed light on related human traits[50,51]. In support of this idea, the brain regions involved in mouse emotionality include the amygdala, which is implicated in human fear-related behaviours[52,53].

Animal models of human disorders: modelling endophenotypes. A number of animal behavioural genetic studies attempt to model features of human psychiatric disorders called endophenotypes, which are simpler markers correlated with the disorder[54]. For example, prepulse inhibition — the suppression of an acoustic startle response by a prior stimulus — is a behaviour that is often diminished in schizophrenic patients[55], and it can be studied in animals to model aspects of schizophrenia[56]. A QTL cross between two mouse strains identified six QTLs that affect prepulse inhibition, as well as a number of loci that are associated with auditory sensitivity and other general behavioural traits[57]. One of the prepulse inhibition QTLs is linked to fatty acid binding protein 7 (*Fabp7*), and a *Fabp7* mouse knockout recapitulates the behavioural effect of the QTL. Although more needs to be done to strengthen the connection, the *Fabp7*-targeted mouse has reduced neurogenesis in the hippocampus, a developmental defect that is potentially consistent with neurodevelopmental defects of schizophrenic patients.

Endophenotypes are also used as animal models of drug addiction and related behaviours. In humans, acute tolerance to the intoxicating effects of alcohol

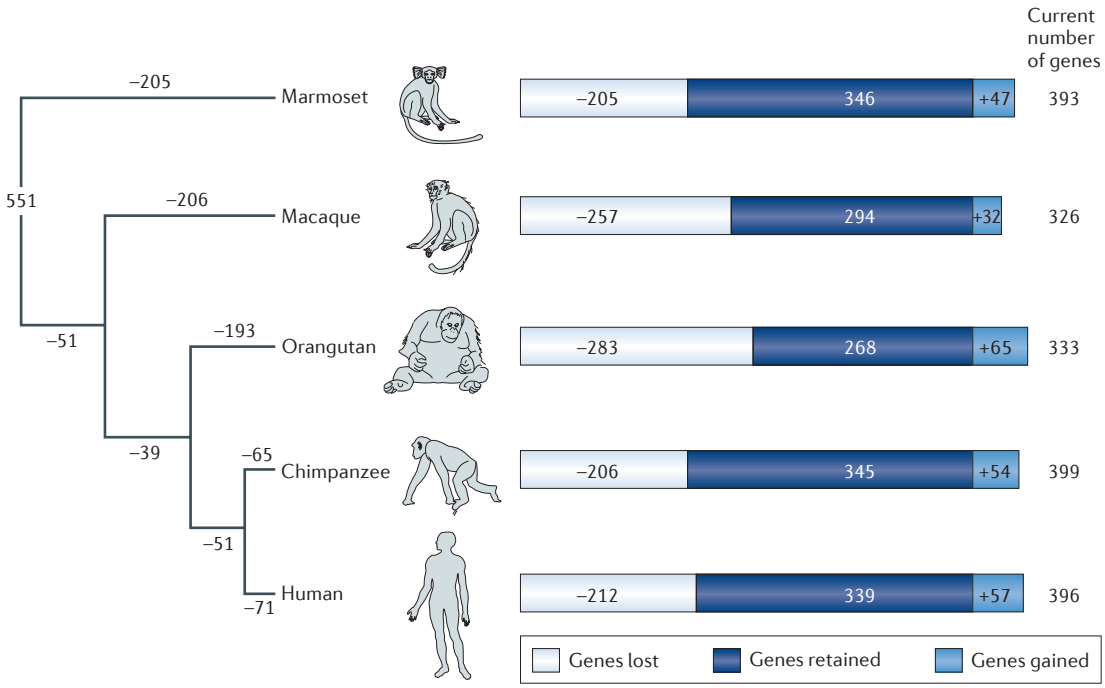


Figure 3 | **Olfactory receptor gene evolution in primates.** Changes in the olfactory receptor gene repertoire in five primate species, based on cross-genomic comparisons. The common ancestor of primates had at least 551 olfactory receptor genes (left). Many olfactory receptor genes have been lost along the five branches to modern primates (numbers of gene losses are shown on each branch). On the right, the number of functional olfactory receptor genes in modern primate species is shown. Humans have lost 212 of the original 551 olfactory receptor genes (white bars), retained 339 olfactory receptor genes (dark blue bars) and gained 57 olfactory receptor genes through gene duplication and divergence (light blue bars), giving a current total of 396 genes. The rate of change in human olfactory receptor genes is similar to the rate in other primate lineages. This figure is modified, with permission, from REF. 150 © (2010) Oxford Univ. Press.

partly predicts alcohol addiction, and this endophenotype of addiction is easily modelled in animals. Many QTLs for alcohol tolerance, dependence and withdrawal have been mapped in rodents, including one tolerance QTL in rats that is tentatively assigned to the neuropeptide Y gene[58] and a withdrawal QTL in mice that probably corresponds to the multiple PDZ (*MPDZ*) gene[59]. Behavioural endophenotypes have the advantage of being traits that are easy to score. One limitation of endophenotypes is that they can be as genetically complex as a disease but are only partly correlated with the disease state.

Genes and the environment

Many behaviours are triggered by sensory cues, and most are regulated by environmental context. The dichotomy between genetic regulation (nature) and environmental regulation (nurture) of behaviour is false: many genes that affect behaviour do so by affecting an animal's detection of, response to or interaction with environmental cues. Examples of each category appear below.

Sensory genes and the environment: variation in chemoreceptor genes. Animals interact with the environment through different sensory modalities, and modifications of these systems appear to be a site of frequent behavioural adaptations.

Receptors for smell and taste belong to large gene families and represent the fastest-evolving neuronal genes in animal genomes, including the human genome (FIG. 3). Modification of smell and taste receptors can lead to rapid changes in behaviour, as can be illustrated by studies of the laboratory adaptation of the nematode *C. elegans*. High-density growth of *C. elegans* in the laboratory resulted in the deletion of two pheromone receptor genes that regulate development based on population density[60]. Both pheromone receptor genes were deleted independently in two strains grown at a high density in different locations. Moreover, a similar pheromone receptor gene was deleted following high-density growth of a different nematode species, *Caenorhabditis briggsae*[60]. Thus, a shift in the environment (in this case, an artificial shift in density) can cause a repeatable change in the repertoire of chemoreceptor genes. Another *C. elegans* chemosensory gene, *globin 5* (*glb-5*), has mutated in association with growth in a high-oxygen laboratory environment[61]. Reduced *glb-5* activity in the laboratory strain decreases the animal's sensitivity to oxygen[61,62] and affects other oxygen-regulated behaviours, such as their tendency to aggregate with other animals[61].

Changes in human chemoreceptor genes are also associated with specific sensory changes. Human polymorphism at a taste receptor (*TAS2R*) leads to

differences in perception of the bitter substance phenylthiocarbamide[63], and polymorphism at the olfactory receptor *OR7D4* leads to differences in perception of androstenone odours[64]. These variations in sensory perception alter the ingestion of bitter food[65] and modify physiological responses to odorants in human sweat, respectively[64,66].

Because there are so many chemoreceptor genes, each tuned to different chemicals, genetic changes to these molecules provide a simple path for modifying specific behaviours without deleterious effects. Variation in chemoreceptor genes between species that occupy different environmental niches carries this principle to the next evolutionary level. Substantial shifts in the olfactory and gustatory receptor repertoire are observed in specialist *Drosophila sechellia* fruit flies that feed exclusively on the *Morinda citrifolia* fruit[67], and a sweet taste receptor gene has been lost in carnivorous cats[68].

Sensory genes and the environment: variation in other sensory systems. Visual systems also show evidence of rapid adaptation within and between animal species. Changes in the sequence and number of visual opsin genes, which encode cone photoreceptor proteins, have occurred repeatedly in vertebrate evolution[69]. Humans and old world monkeys have a recent opsin gene duplication that allows red–green colour discrimination (reviewed in REF. 70). New world monkeys lack this duplication but show evidence of intraspecies variation in colour discrimination: two alleles of a long-to-middle (L–M) wavelength-sensitive opsin gene on the X chromosome are maintained by balancing selection, and females that are heterozygous at this locus are able to discriminate more colours than hemizygous males or homozygous females[71].

At a higher level, sensory systems can rapidly remodel the design or number of sensory organs to change behaviour. For example, the tetra fish *Astyanax mexicanus* exists in two forms, one sighted and surface-dwelling, the other blind and cave-dwelling. The loss of sight in cave-dwelling populations has been accompanied by expanded cell numbers in the mechanosensory organs of the lateral line, an adaptation that increases sensitivity to vibrations from food falling on the water surface[72]. In humans, genes that are required in the auditory system show signatures of positive selection, and it has been speculated that these changes are related to a sophisticated human ability: the use of language[73]. The observation that sensory perception genes are among the genes that are under the strongest positive selection in the human genome[74,75] supports the conclusion that sensory genes are a preferred target of behavioural adaptation.

Genes and the response to environmental cues. An animal's sensitivity to environmental cues is determined not only by its sensory receptors but also by its shifting internal states. For example, hungry animals are more sensitive to food-related cues and less sensitive to aversive cues than well-fed animals. This example describes behavioural variability within one individual, but the interface between internal and external cues also represents a site for behavioural variation between individuals. An example of natural variation at this interface is an animal's choice of whether to abandon a depleting food supply, which is known in behavioural ecology as the exploration–exploitation decision[76]. In *C. elegans*, as in other animals, abandoning a food supply is strongly modulated by environmental cues, including food quality, food quantity and animal density[77,78]. It is also modulated by genetic variation and has been studied using recombinant inbred lines from two *C. elegans* strains that differ in their tendency to leave food[14,32]. One QTL for the exploration–exploitation behaviour corresponds to the G-protein-coupled *tyramine receptor 3* (*tyra-3*), which is related to vertebrate adrenergic receptors. Non-coding polymorphisms in *tyra-3* alter its expression levels in sensory neurons that detect food cues[32] and apparently modulate sensitivity to those food cues. The effect of *tyra-3* is only observed at intermediate food levels — all animals, regardless of the *tyra-3* allele, remain on abundant food and abandon food that is present at low amounts[32]. These results show how genetic variation interacts with the environment to regulate behaviour: internal arousal states, signalled through adrenergic receptors, can couple strongly or weakly (depending on the adrenergic receptor allele) with a sensory input that modulates behaviour.

The ligand for *tyra-3* is tyramine[79], one of several invertebrate neurotransmitters that are related to vertebrate adrenaline and noradrenaline[80]. This class of transmitters is linked to arousal states in invertebrates and vertebrates[80]. Exploration versus exploitation decisions in primates are also regulated by noradrenaline release[81], suggesting that these transmitters have broad roles in decision making.

Genetic variation and the environment act on common substrates. Another connection between genetic and environmental regulation of behaviour is illustrated by studies of *D. melanogaster* aggression. Fruit flies defend food resources or potential mates with attack behaviours such as lunging and boxing[82]. The genetic underpinnings of these behaviours have been analysed by QTL approaches[15] and by selective breeding strategies in which the most aggressive flies in genetically heterogeneous populations were selectively mated for more than 20 generations[83,84]. Both approaches found multigenic effects on aggression.

As a way of identifying candidate loci that regulate aggression, microarray analysis was used to search for genes that were differentially expressed between highly aggressive and less aggressive flies derived from selective breeding or from random inbreeding of wild-derived populations[83–85] (FIG. 2). The overall transcriptional differences were substantial in the three separate studies, but the identified gene sets were largely non-overlapping. A possible explanation for this finding is that outbred populations are so diverse that multiple independent combinations of alleles can lead to highly aggressive or non-aggressive behaviour. Nonetheless, 25 of the differentially expressed genes had effects on aggression when tested using knockout alleles, supporting the validity of the candidate transcripts as regulators of behaviour[83–85].

One informative gene that emerged from the artificially selected aggressive lines was *Cyp6a20*, which encodes one of multiple cytochrome P450 enzymes in *D. melanogaster*. Expression of this gene is low in aggressive strains, and reduced expression is sufficient for behavioural differences in aggression[83]. *Cyp6a20* is expressed in olfactory sensory organs, where it may regulate the responses to pheromones that influence aggression[86]. Interestingly, mRNA levels of *Cyp6a20* are reduced in flies that are reared in isolation, which are more aggressive than socially experienced flies[86]. This observation suggests that social experience modulates aggression by changing *Cyp6a20* levels. Thus, *Cyp6a20* sits at the intersection of genetic and environmental influences on behaviour: either a genetic change or an environmental change that decreases *Cyp6a20* expression leads to increased aggression.

Gene–environment interactions are strongly supported in humans as well and may form a framework for understanding many psychiatric disorders and risk factors. In depression, for example, genetic susceptibility (having an identical twin who is depressed) interacts with risk from environmental insults, such as divorce or death in the family[87].

From genes to circuits that affect behaviour

A genetic change that affects behaviour acts in the context of the neural processes that generate the behaviour. In some cases, such as changes in sensory receptor genes, this relationship is straightforward. In other cases, studying the behavioural gene can provide new insights into brain circuits, as illustrated by studies of animal social behaviour.

From a social gene to a social circuit. Most *C. elegans* strains are social feeders that aggregate in the presence of food. The laboratory strain N2, however, is a solitary feeder: its low levels of aggregation are associated with a high-activity allele of a neuropeptide Y receptor homologue, *npr-1*, which differs from a low-activity (high-aggregation) allele at a single amino acid residue[88]. The high-activity solitary allele arose during laboratory cultivation[61] and increases fitness in the laboratory environment[89,90].

Expression of *npr-1* is present in ~10% of *C. elegans* neurons, but its influence on aggregation is dominated by its effects on a single pair of integrating neurons called RMG neurons[91]. RMG neurons form electrical synapses with many sensory neurons that stimulate aggregation, suggesting that RMG couples multiple sensory inputs that drive a common behaviour. The high-activity *npr-1* variant partially uncouples this circuit to diminish aggregation behaviour without disrupting other important roles of the sensory cues. The discovery of this circuit element via the *npr-1* variant shows how genetic approaches can advance neurobiological studies.

RMG neurons
Caenorhabditis elegans neurons that are essential for aggregation and are linked by electrical synapses to multiple classes of sensory neurons that detect oxygen, pheromones, noxious cues and nutrients.

Differential gene expression pattern leads to changes in behaviour: the case of neuropeptide receptors. Variation in social behaviour is commonly observed within and between mammalian species, and here too genetic studies have provided fresh insights into the neurobiology of social behaviour. Two related neuropeptides, oxytocin and arginine vasopressin (AVP), are important regulators of mammalian social and reproductive behaviour[92], and genetic variation in AVP signalling has been linked to the different social behaviour of monogamous prairie voles and polygamous montane voles[93–95]. Both vole species have functional AVP genes and functional vasopressin receptor genes, but they differ in their expression of the vasopressin V1a receptor (V1aR). A brain region that is involved in the neurobiology of reward called the ventral pallidum only expresses V1aR in monogamous voles[96] and, remarkably, affiliative behaviour of polygamous montane voles is substantially increased by virus-mediated introduction of arginine vasopressin receptor 1A (*Avpr1a*; the gene encoding V1aR) into the ventral pallidum[97]. These results implicate differential expression of the V1aR neuropeptide receptor in the differential organization of social behaviours in the two vole species. They also point to the ventral pallidum as a site that can encode rewarding features of social cues. Little is known about the human circuits for social behaviour, but the rodent pathways provide a starting point for further investigation.

Emerging themes in the genetics of behaviour

Is it possible to derive general principles about natural variation and the evolution of behaviour by analogy with common principles that have been uncovered in evolutionary developmental biology[98]? Early signs are promising. Natural variation in the gene encoding PKG affects foraging behaviour in both insects and nematodes, and variation in the *period* gene affects circadian rhythm in flies and humans. Other indications of common themes are not so strongly tied to a single gene, but may be tied to classes of genes, such as the sensory receptor genes described above. We suggest that highly evolvable behavioural genes will be characterized by diversity, as exemplified by multigene families, and by modular flexibility, which is the ability to easily form new behavioural connections.

Adaptable neuromodulatory pathways. Several behavioural trait genes are associated with G-protein-coupled neurotransmitter receptors or their regulators: *Rgs2* in mice, *Avpr1a* in voles, and *tyra-3* and *npr-1* in nematodes. In each case, the GPCR system is associated with internal motivational states — anxiety, affiliation, arousal or hunger — that set thresholds for behavioural responses to external stimuli.

We suggest that GPCR pathways are amenable to natural variation because of their diversity and modular flexibility. All animal genomes encode dozens of GPCRs for neuropeptides and for modulatory bioamines that modify neuronal excitability and synaptic strength. These modulators are typically not essential for core neurotransmission and this, coupled with their variety, leaves room for evolvability. Moreover, neuromodulators can act at a distance and not just at local synapses; this allows them to broadcast internal motivational or arousal states. Action at a distance enables the creation

Stomatogastric ganglion
A part of the crustacean nervous system that coordinates digestive tract movements. It consists of 30 defined neurons and has been extensively used to study neural circuit dynamics, connections and modulation.

of new behavioural links between distant brain areas by simply modifying the site of receptor expression without requiring growth of new anatomical connections. In agreement with the hypothesis that neuromodulators are substrates for behavioural diversity, neuropeptides and neuropeptide receptor expression patterns evolve rapidly. Expression of oxytocin and vasopressin receptors is highly variable among rodent species[99,100], and cross-species comparisons of the stomatogastric ganglion of crustaceans show a near-invariant set of neurons but a divergence in neuropeptide expression[101,102].

Genetic variation in GCPRs and other regulators of neurotransmission has also been suggestively associated with human behavioural and psychiatric traits. A cautious view of these results is warranted, as promising results have often failed to maintain significance upon meta-analysis[103–107]. With that concern in mind, recent studies have implicated GPCR variants in the risk for psychiatric disorders: rare microduplications in the vasoactive intestinal peptide receptor 1 (*VIPR1*) gene in schizophrenia[108], a polymorphism in the 5-hydroxytryptamine (serotonin) receptor 2B (*HTR2B)* gene in Finns with severe impulsivity[109] and common polymorphisms at the pituitary adenylate cyclase-activating polypeptide receptor gene *PAC1* in post-traumatic stress disorder[110].

Balancing selection for behavioural traits. Taking a step back from the specific genes that affect behaviour, why does behavioural variation persist within a species? Genetic variation is generated by mutation and maintained through drift, population-specific selection or balancing selection[111], and either population-specific or balancing selection can support variety in behaviours[112]. Emotionality traits in mice are potentially subject to balancing selection in different environments: a predator-rich environment may favour animals that are highly responsive to potentially dangerous stimuli, whereas a predator-poor environment may relax that selective pressure and favour bolder animals[113]. Foraging activity is another behavioural axis that is subject to balancing selection: depending on resource distribution in the environment, high activity levels that promote exploration may be more or less advantageous than low activity levels that conserve energy resources. The rover and sitter alleles of *D. melanogaster* larvae are maintained in wild populations by balancing selection of an interesting kind: different alleles are favoured depending on the density of larvae and the relative frequency of rover and sitter alleles in the population[114,115]. Density-dependent selection and frequency-dependent selection are special cases of balancing selection that are relevant to social behaviours as well as foraging.

In this analysis, the reproducibility of an animal's environment should have predictable effects on its genetic variability. If an important environmental cue is entirely reliable, such as the circadian cycle in the tropics, information about that cue should be encoded by the genome. If an environmental cue is entirely unreliable, animals may learn about it from individual experience. In the intermediate domain, information about food supply, predators, weather patterns or population density may be variable or constant, and balancing selection may encode different degrees of genetic versus experience-dependent behaviour, accordingly.

Human populations may also be subject to balancing selection for behavioural adaptations. In that context, it may be fruitful to consider traits that are frequently under balancing selection in animals, such as foraging strategies, activity levels and sensitivity to threat.

Perspective

The analysis of natural variation in behaviour has convincingly shown a complex genetic basis and pervasive interactions between genetic variants and the environment. Current excitement focuses on identifying more of the molecules that are involved in behavioural variation and translating these genetic discoveries into neurobiological and evolutionary insight.

1. Konopka, R. J. & Benzer, S. Clock mutants of *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **68**, 2112–2116 (1971).
2. Toh, K. L. *et al.* An hPer2 phosphorylation site mutation in familial advanced sleep phase syndrome. *Science* **291**, 1040–1043 (2001).
3. Xu, Y. *et al.* Functional consequences of a *CKIδ* mutation causing familial advanced sleep phase syndrome. *Nature* **434**, 640–644 (2005).
4. Clement, K. *et al.* A mutation in the human leptin receptor gene causes obesity and pituitary dysfunction. *Nature* **392**, 398–401 (1998).
5. Peyron, C. *et al.* A mutation in a case of early onset narcolepsy and a generalized absence of hypocretin peptides in human narcoleptic brains. *Nature Med.* **6**, 991–997 (2000).
6. Lee, G. H. *et al.* Abnormal splicing of the leptin receptor in diabetic mice. *Nature* **379**, 632–635 (1996).
7. Chemelli, R. M. *et al.* Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. *Cell* **98**, 437–451 (1999).
8. Lin, L. *et al.* The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* **98**, 365–376 (1999).
9. Fanara, J. J., Robinson, K. O., Rollmann, S. M., Anholt, R. R. & Mackay, T. F. *Vanaso* is a candidate quantitative trait gene for *Drosophila* olfactory behavior. *Genetics* **162**, 1321–1328 (2002).
10. Flint, J. Analysis of quantitative trait loci that influence animal behavior. *J. Neurobiol.* **54**, 46–77 (2003). **This paper contains a systematic description of behavioural QTLs identified in rodent studies.**
11. Gleason, J. M. & Ritchie, M. G. Do quantitative trait loci (QTL) for a courtship song difference between *Drosophila simulans* and *D. sechellia* coincide with candidate genes and intraspecific QTL? *Genetics* **166**, 1303–1311 (2004).
12. Jordan, K. W., Morgan, T. J. & Mackay, T. F. Quantitative trait loci for locomotor behavior in *Drosophila melanogaster*. *Genetics* **174**, 271–284 (2006).
13. Churchill, G. A. *et al.* The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nature Genet.* **36**, 1133–1137 (2004).
14. Rockman, M. V. & Kruglyak, L. Recombinational landscape and population genomics of *Caenorhabditis elegans*. *PLoS Genet.* **5**, e1000419 (2009). **This paper reports the development of recombinant inbred advanced-intercross lines for *C. elegans* trait mapping.**
15. Edwards, A. C. & Mackay, T. F. Quantitative trait loci for aggressive behavior in *Drosophila melanogaster*. *Genetics* **182**, 889–897 (2009).
16. Flint, J., Valdar, W., Shifman, S. & Mott, R. Strategies for mapping and cloning quantitative trait genes in rodents. *Nature Rev. Genet.* **6**, 271–286 (2005). **This paper includes an analysis of effect sizes of hundreds of behavioural and physiological QTLs.**
17. Chen, W. C. *Construction and use of* Caenorhabditis elegans *chromosome substitution strains to map a novel p38 component involved in innate immunity*. Thesis, Stanford Univ. (2008).
18. Doroszuk, A., Snoek, L. B., Fradin, E., Riksen, J. & Kammenga, J. A genome-wide library of CB4856/N2 introgression lines of *Caenorhabditis elegans*. *Nucleic Acids Res.* **37**, e110 (2009).
19. Hollocher, H., Ting, C. T., Wu, M. L. & Wu, C. I. Incipient speciation by sexual isolation in *Drosophila melanogaster*: extensive genetic divergence without reinforcement. *Genetics* **147**, 1191–1201 (1997).
20. Mattson, D. L. *et al.* Chromosome substitution reveals the genetic basis of Dahl salt-sensitive hypertension and renal disease. *Am. J. Physiol. Renal Physiol.* **295**, F837–F842 (2008).
21. Nadeau, J. H., Singer, J. B., Matin, A. & Lander, E. S. Analysing complex genetic traits with chromosome substitution strains. *Nature Genet.* **24**, 221–225 (2000).
22. Iakoubova, O. A. *et al.* Genome-tagged mice (GTM): two sets of genome-wide congenic strains. *Genomics* **74**, 89–104 (2001).
23. Singer, J. B. *et al.* Genetic dissection of complex traits with chromosome substitution strains of mice. *Science* **304**, 445–448 (2004).
24. Singer, J. B., Hill, A. E., Nadeau, J. H. & Lander, E. S. Mapping quantitative trait loci for anxiety in chromosome substitution strains of mice. *Genetics* **169**, 855–862 (2005).

25. Gale, G. D. *et al.* A genome-wide panel of congenic mice reveals widespread epistasis of behavior quantitative trait loci. *Mol. Psychiatry* **14**, 631–645 (2009).
26. Shao, H. *et al.* Genetic architecture of complex traits: large phenotypic effects and pervasive epistasis. *Proc. Natl Acad. Sci. USA* **105**, 19910–19914 (2008).
References 24 and 26 describe a systematic analysis of mouse chromosome-substitution strains that argues for large effects of individual QTLs when they are examined in homogeneous rather than heterogeneous strain backgrounds.
27. Threadgill, D. W. *et al.* Targeted disruption of mouse EGF receptor: effect of genetic background on mutant phenotype. *Science* **269**, 230–234 (1995).
28. Dowell, R. D. *et al.* Genotype to phenotype: a complex problem. *Science* **328**, 469 (2010).
Knockouts of yeast genes are examined in the genetic backgrounds of two strains, showing that 5% of all 'essential' genes are essential in one background but not in the other.
29. Meffert, L. M., Hicks, S. K. & Regan, J. L. Nonadditive genetic effects in animal behavior. *Am. Nat.* **160**, S198–S213 (2002).
30. Ruppell, O., Pankiw, T. & Page, R. E. Jr. Pleiotropy, epistasis and new QTL: the genetic architecture of honey bee foraging behavior. *J. Hered.* **95**, 481–491 (2004).
31. Arizmendi, C., Zuleta, V., Ruiz-Dubreuil, G. & Godoy-Herrera, R. Genetics analysis of larval foraging behavior in *Drosophila funebris*. *Behav. Genet.* **38**, 525–530 (2008).
32. Bendesky, A., Tsunozaki, M., Rockman, M. V., Kruglyak, L. & Bargmann, C. I. Catecholamine receptor polymorphisms affect decision-making in *C. elegans*. *Nature* **472**, 313–318 (2011).
A *C. elegans* receptor related to mammalian adrenergic receptors is reported as regulating an exploration–exploitation decision in this paper.
33. Ott, J., Kamatani, Y. & Lathrop, M. Family-based designs for genome-wide association studies. *Nature Rev. Genet.* **12**, 465–474 (2011).
34. Sokolowski, M. B. Foraging strategies of *Drosophila melanogaster:* a chromosomal analysis. *Behav. Genet.* **10**, 291–302 (1980).
35. Sokolowski, M. B. *et al.* Ecological genetics and behaviour of *Drosophila melanogaster* larvae in nature. *Anim. Behav.* **34**, 403–408 (1986).
36. de Belle, J. S. & Sokolowski, M. B. Heredity of rover/sitter: alternative foraging strategies of *Drosophila melanogaster*. *Heredity* **59**, 73–83 (1987).
37. Osborne, K. A. *et al.* Natural behavior polymorphism due to a cGMP-dependent protein kinase of *Drosophila*. *Science* **277**, 834–836 (1997).
This paper discusses the molecular identification of the *D. melanogaster for* gene as a cGMP-dependent protein kinase.
38. Renger, J. J., Yao, W. D., Sokolowski, M. B. & Wu, C. F. Neuronal polymorphism among natural alleles of a cGMP-dependent kinase gene, *foraging*, in *Drosophila*. *J. Neurosci.* 19, RC28 (1999).
39. Ben-Shahar, Y., Robichon, A., Sokolowski, M. B. & Robinson, G. E. Influence of gene action across different time scales on behavior. *Science* **296**, 741–744 (2002).
40. Ingram, K. K., Oefner, P. & Gordon, D. M. Task-specific expression of the *foraging* gene in harvester ants. *Mol. Ecol.* **14**, 813–818 (2005).
41. Hong, R. L., Witte, H. & Sommer, R. J. Natural variation in *Pristionchus pacificus* insect pheromone attraction involves the protein kinase EGL-4. *Proc. Natl Acad. Sci. USA* **105**, 7779–7784 (2008).
42. Yalcin, B. *et al.* Genetic dissection of a behavioral quantitative trait locus shows that *Rgs2* modulates anxiety in mice. *Nature Genet.* **36**, 1197–1202 (2004).
This paper describes the identification of *Rgs2* as the causative gene for an emotionality trait in mouse.
43. Stam, L. F. & Laurie, C. C. Molecular dissection of a major gene effect on a quantitative trait: the level of alcohol dehydrogenase expression in *Drosophila melanogaster*. *Genetics* **144**, 1559–1564 (1996).
44. Legare, M. E., Bartlett, F. S., 2nd & Frankel, W. N. A major effect QTL determined by multiple genes in epileptic EL mice. *Genome Res.* **10**, 42–48 (2000).
45. Steinmetz, L. M. *et al.* Dissecting the architecture of a quantitative trait locus in yeast. *Nature* **416**, 326–330 (2002).
46. Thomson, M. J., Edwards, J. D., Septiningsih, E. M., Harrington, S. E. & McCouch, S. R. Substitution mapping of dth1.1, a flowering-time quantitative trait locus (QTL) associated with transgressive variation in rice, reveals multiple sub-QTL. *Genetics* **172**, 2501–2514 (2006).
47. Grafstein-Dunn, E., Young, K. H., Cockett, M. I. & Khawaja, X. Z. Regional distribution of regulators of G-protein signaling (RGS) 1, 2, 13, 14, 16, and GAIP messenger ribonucleic acids by *in situ* hybridization in rat brain. *Mol. Brain Res.* **88**, 113–123 (2001).
48. Heximer, S. P. *et al.* Hypertension and prolonged vasoconstrictor signaling in RGS2-deficient mice. *J. Clin. Invest.* **111**, 1259 (2003).
49. Sun, X., Kaltenbronn, K. M., Steinberg, T. H. & Blumer, K. J. RGS2 is a mediator of nitric oxide action on blood pressure and vasoconstrictor signaling. *Mol. Pharmacol.* **67**, 631–639 (2005).
50. Flint, J. The genetic basis of neuroticism. *Neurosci. Biobehav. Rev.* **28**, 307–316 (2004).
51. Willis-Owen, S. A. & Flint, J. Identifying the genetic determinants of emotionality in humans; insights from rodents. *Neurosci. Biobehav. Rev.* **31**, 115–124 (2007).
52. Morris, J. S. *et al.* A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature* **383**, 812–815 (1996).
53. LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E. & Phelps, E. A. Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron* **20**, 937–945 (1998).
54. Kendler, K. S. & Neale, M. C. Endophenotype: a conceptual analysis. *Mol. Psychiatry* **15**, 789–797 (2010).
55. Braff, D. L., Geyer, M. A. & Swerdlow, N. R. Human studies of prepulse inhibition of startle: normal subjects, patient groups, and pharmacological studies. *Psychopharmacology* **156**, 234–258 (2001).
56. Geyer, M. A., Krebs-Thomson, K., Braff, D. L. & Swerdlow, N. R. Pharmacological studies of prepulse inhibition models of sensorimotor gating deficits in schizophrenia: a decade in review. *Psychopharmacology* **156**, 117–154 (2001).
57. Watanabe, A. *et al.* Fabp7 maps to a quantitative trait locus for a schizophrenia endophenotype. *PLoS Biol.* **5**, e297 (2007).
58. Carr, L. G. *et al.* A quantitative trait locus for alcohol consumption in selectively bred rat lines. *Alcohol Clin. Exp. Res.* **22**, 884–887 (1998).
59. Shirley, R. L., Walter, N. A., Reilly, M. T., Fehr, C. & Buck, K. J. *Mpdz* is a quantitative trait gene for drug withdrawal seizures. *Nature Neurosci.* **7**, 699–700 (2004).
60. McGrath, P. T. *et al.* Parallel evolution of domesticated *Caenorhabditis* species targets pheromone receptor genes. *Nature* **477**, 321–325 (2011).
61. McGrath, P. T. *et al.* Quantitative mapping of a digenic behavioral trait implicates globin variation in *C. elegans* sensory behaviors. *Neuron* **61**, 692–699 (2009).
62. Persson, A. *et al.* Natural variation in a neural globin tunes oxygen sensing in wild *Caenorhabditis elegans*. *Nature* **458**, 1030–1033 (2009).
63. Kim, U. K. *et al.* Positional cloning of the human quantitative trait locus underlying taste sensitivity to phenylthiocarbamide. *Science* **299**, 1221–1225 (2003).
This paper reports the mapping of a polymorphic sensation of bitter taste in humans, familiar to many from 'taste paper' experiments in science class, to a taste receptor.
64. Keller, A., Zhuang, H., Chi, Q., Vosshall, L. B. & Matsunami, H. Genetic variation in a human odorant receptor alters odour perception. *Nature* **449**, 468–472 (2007).
This study shows that differential olfactory sensitivity to androstenone, an androgen-derived odour, is associated with polymorphism in a specific olfactory receptor in humans.
65. Hayes, J. E. *et al.* Allelic variation in *TAS2R* bitter receptor genes associates with variation in sensations from and ingestive behaviors toward common bitter beverages in adults. *Chem. Senses* **36**, 311–319 (2011).
66. Wyart, C. *et al.* Smelling a single component of male sweat alters levels of cortisol in women. *J. Neurosci.* **27**, 1261–1265 (2007).
67. McBride, C. S. Rapid evolution of smell and taste receptor genes during host specialization in *Drosophila sechellia*. *Proc. Natl Acad. Sci. USA* **104**, 4996–5001 (2007).
68. Li, X. *et al.* Pseudogenization of a sweet-receptor gene accounts for cats' indifference toward sugar. *PLoS Genet.* **1**, 27–35 (2005).
69. Collin, S. P. & Trezise, A. E. The origins of colour vision in vertebrates. *Clin. Exp. Optom* **87**, 217–223 (2004).
70. Jacobs, G. H. Evolution of colour vision in mammals. *Phil. Trans. R. Soc. B* **364**, 2957–2967 (2009).
71. Hiwatashi, T. *et al.* An explicit signature of balancing selection for color-vision variation in new world monkeys. *Mol. Biol. Evol.* **27**, 453–464 (2010).
72. Yoshizawa, M., Goricki, S., Soares, D. & Jeffery, W. R. Evolution of a behavioral shift mediated by superficial neuromasts helps cavefish find food in darkness. *Curr. Biol.* **20**, 1631–1636 (2010).
73. Akey, J. M. Constructing genomic maps of positive selection in humans: where do we go from here? *Genome Res.* **19**, 711–722 (2009).
74. Clark, A. G. *et al.* Inferring nonneutral evolution from human–chimp–mouse orthologous gene trios. *Science* **302**, 1960–1963 (2003).
75. Nielsen, R. *et al.* A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biol.* **3**, e170 (2005).
76. Stephens, D. W. & Kerbs, J. R. *Foraging Theory* (Princeton Univ. Press, 1987).
77. Goubault, M. N., Outreman, Y., Poinsot, D. & Cortesero, A. M. Patch exploitation strategies of parasitic wasps under intraspecific competition. *Behav. Ecol.* **16**, 693–701 (2005).
78. Shtonda, B. B. & Avery, L. Dietary choice behavior in *Caenorhabditis elegans*. *J. Exp. Biol.* **209**, 89–102 (2006).
79. Wragg, R. T. *et al.* Tyramine and octopamine independently inhibit serotonin-stimulated aversive behaviors in *Caenorhabditis elegans* through two novel amine receptors. *J. Neurosci.* **27**, 13402–13412 (2007).
80. Roeder, T. Tyramine and octopamine: ruling behavior and metabolism. *Annu. Rev. Entomol.* **50**, 447–477 (2005).
81. Aston-Jones, G. & Cohen, J. D. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* **28**, 403–450 (2005).
82. Chen, S., Lee, A. Y., Bowens, N. M., Huber, R. & Kravitz, E. A. Fighting fruit flies: a model system for the study of aggression. *Proc. Natl Acad. Sci. USA* **99**, 5664–5668 (2002).
83. Dierick, H. A. & Greenspan, R. J. Molecular analysis of flies selected for aggressive behavior. *Nature Genet.* **38**, 1023–1031 (2006).
This paper reports the identification of the *Cyp6a20* gene as a regulator of *D. melanogaster* aggression through transcriptional profiling of highly aggressive strains.
84. Edwards, A. C., Rollmann, S. M., Morgan, T. J. & Mackay, T. F. Quantitative genomics of aggressive behavior in *Drosophila melanogaster*. *PLoS Genet.* **2**, e154 (2006).
85. Edwards, A. C. *et al.* A transcriptional network associated with natural variation in *Drosophila* aggressive behavior. *Genome Biol.* **10**, R76 (2009).
86. Wang, L., Dankert, H., Perona, P. & Anderson, D. J. A common genetic target for environmental and heritable influences on aggressiveness in *Drosophila*. *Proc. Natl Acad. Sci. USA* **105**, 5657–5663 (2008).
This study demonstrates that genetic variation and social experience converge on *Cyp6a20* to regulate aggression in male *D. melanogaster*.
87. Kendler, K. S. *et al.* Stressful life events, genetic liability, and onset of an episode of major depression in women. *Am. J. Psychiatry* **152**, 833–842 (1995).
This paper discusses the exemplary use of human twin studies to understand genotype–environment interactions in major depression.
88. de Bono, M. & Bargmann, C. I. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. *Cell* **94**, 679–689 (1998).
89. Reddy, K. C., Andersen, E. C., Kruglyak, L. & Kim, D. H. A polymorphism in *npr-1* is a behavioral determinant of pathogen susceptibility in *C. elegans*. *Science* **323**, 382–384 (2009).
90. Weber, K. P. *et al.* Whole genome sequencing highlights genetic changes associated with laboratory domestication of *C. elegans*. *PLoS ONE* **5**, e13922 (2010).

91. Macosko, E. Z. *et al.* A hub-and-spoke circuit drives pheromone attraction and social behaviour in *C. elegans*. *Nature* **458**, 1171–1175 (2009).
This study demonstrates that variation in the *C. elegans* neuropeptide receptor gene *npr-1* alters sensory processing in a gap junction circuit to regulate aggregation.
92. Lim, M. M. & Young, L. J. Neuropeptidergic regulation of affiliative behavior and social bonding in animals. *Horm. Behav.* **50**, 506–517 (2006).
93. Shapiro, L. E. & Dewsbury, D. A. Differences in affiliative behavior, pair bonding, and vaginal cytology in two species of vole (*Microtus ochrogaster* and *M. montanus*). *J. Comp. Psychol.* **104**, 268–274 (1990).
94. Winslow, J. T., Hastings, N., Carter, C. S., Harbaugh, C. R. & Insel, T. R. A role for central vasopressin in pair bonding in monogamous prairie voles. *Nature* **365**, 545–548 (1993).
95. Young, L. J., Nilsen, R., Waymire, K. G., MacGregor, G. R. & Insel, T. R. Increased affiliative response to vasopressin in mice expressing the V1a receptor from a monogamous vole. *Nature* **400**, 766–768 (1999).
96. Insel, T. R., Wang, Z. X. & Ferris, C. F. Patterns of brain vasopressin receptor distribution associated with social organization in microtine rodents. *J. Neurosci.* **14**, 5381–5392 (1994).
97. Lim, M. M. *et al.* Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene. *Nature* **429**, 754–757 (2004).
98. Stern, D. L. & Orgogozo, V. Is genetic evolution predictable? *Science* **323**, 746–751 (2009).
99. Insel, T. R., Gelhard, R. & Shapiro, L. E. The comparative distribution of forebrain receptors for neurohypophyseal peptides in monogamous and polygamous mice. *Neuroscience* **43**, 623–630 (1991).
100. Insel, T. R. & Shapiro, L. E. Oxytocin receptor distribution reflects social organization in monogamous and polygamous voles. *Proc. Natl Acad. Sci. USA* **89**, 5981–5985 (1992).
101. Marder, E., Calabrese, R. L., Nusbaum, M. P. & Trimmer, B. Distribution and partial characterization of FMRFamide-like peptides in the stomatogastric nervous systems of the rock crab, *Cancer borealis*, and the spiny lobster, *Panulirus interruptus*. *J. Comp. Neurol.* **259**, 150–163 (1987).
102. Verley, D. R., Doan, V., Trieu, Q., Messinger, D. I. & Birmingham, J. T. Characteristic differences in modulation of stomatogastric musculature by a neuropeptide in three species of *Cancer* crabs. *J. Comp. Physiol. A* **194**, 879–886 (2008).
103. Clarke, H., Flint, J., Attwood, A. S. & Munafo, M. R. Association of the *5- HTTLPR* genotype and unipolar depression: a meta-analysis. *Psychol. Med.* **40**, 1767–1778 (2010).
104. Kohli, M. A. *et al.* The neuronal transporter gene *SLC6A15* confers risk to major depression. *Neuron* **70**, 252–265 (2011).
105. Munafo, M. R., Yalcin, B., Willis-Owen, S. A. & Flint, J. Association of the dopamine D4 receptor (*DRD4*) gene and approach-related personality traits: meta-analysis and new data. *Biol. Psychiatry* **63**, 197–206 (2008).
106. Barnett, J. H., Scoriels, L. & Munafo, M. R. Meta-analysis of the cognitive effects of the catechol-*O*-methyltransferase gene Val158/108Met polymorphism. *Biol. Psychiatry* **64**, 137–144 (2008).
107. Green, A. E. *et al.* Using genetic data in cognitive neuroscience: from growing pains to genuine insights. *Nature Rev. Neurosci.* **9**, 710–720 (2008).
108. Vacic, V. *et al.* Duplications of the neuropeptide receptor gene *VIPR2* confer significant risk for schizophrenia. *Nature* **471**, 499–503 (2011).
109. Bevilacqua, L. *et al.* A population-specific *HTR2B* stop codon predisposes to severe impulsivity. *Nature* **468**, 1061–1066 (2010).
110. Ressler, K. J. *et al.* Post-traumatic stress disorder is associated with PACAP and the PAC1 receptor. *Nature* **470**, 492–497 (2011).
111. Charlesworth, B. & Charlesworth, D. *Elements of Evolutionary Genetics* (Roberts and Company Publishers, 2010).
112. Wolf, M., van Doorn, G. S., Leimar, O. & Weissing, F. J. Life-history trade-offs favour the evolution of animal personalities. *Nature* **447**, 581–584 (2007).
113. Giles, N. & Huntingford, F. A. Predation risk and inter-population variation in antipredator behaviour in the three-spined stickleback, *Gasterosteus aculeatus* L. *Anim. Behav.* **32**, 264–275 (1984).
114. Sokolowski, M. B., Pereira, H. S. & Hughes, K. Evolution of foraging behavior in *Drosophila* by density-dependent selection. *Proc. Natl Acad. Sci. USA* **94**, 7373–7377 (1997).
115. Fitzpatrick, M. J., Feder, E., Rowe, L. & Sokolowski, M. B. Maintaining a behaviour polymorphism by frequency-dependent selection on a single gene. *Nature* **447**, 210–212 (2007).
This study shows that two alleles of the *for* gene are maintained in *D. melanogaster* larvae by balancing selection that is based on allele frequency in the population.
116. Goulding, E. H. *et al.* A robust automated system elucidates mouse home cage behavioral structure. *Proc. Natl Acad. Sci. USA* **105**, 20575–20582 (2008).
117. Branson, K., Robie, A. A., Bender, J., Perona, P. & Dickinson, M. H. High-throughput ethomics in large groups of *Drosophila*. *Nature Methods* **6**, 451–457 (2009).
118. Albrecht, D. R. & Bargmann, C. I. High-content behavioral analysis of *Caenorhabditis elegans* in precise spatiotemporal chemical environments. *Nature Methods* **8**, 599–605 (2011).
119. Crabbe, J. C., Wahlsten, D. & Dudek, B. C. Genetics of mouse behavior: interactions with laboratory environment. *Science* **284**, 1670–1672 (1999).
This paper reports cautionary results demonstrating the sensitivity of behavioural analysis to small differences in testing conditions.
120. Flint, J. Mapping quantitative traits and strategies to find quantitative trait genes. *Methods* **53**, 163–174 (2011).
121. Maye, A., Hsieh, C. H., Sugihara, G. & Brembs, B. Order in spontaneous behavior. *PLoS ONE* **2**, e443 (2007).
122. Tandon, R., Keshavan, M. S. & Nasrallah, H. A. Schizophrenia, "just the facts" what we know in 2008. 2. Epidemiology and etiology. *Schizophr. Res.* **102**, 1–18 (2008).
123. Bailey, A. *et al.* Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol. Med.* **25**, 63–77 (1995).
124. Hallmayer, J. *et al.* Genetic heritability and shared environmental factors among twin pairs with autism. *Arch. Gen. Psychiatry* 4 Jul 2011 (doi:10.1001/archgenpsychiatry.2011.76).
125. Kendler, K. S., Pedersen, N. L., Neale, M. C. & Mathe, A. A. A pilot Swedish twin study of affective illness including hospital- and population-ascertained subsamples: results of model fitting. *Behav. Genet.* **25**, 217–232 (1995).
126. McGuffin, P. *et al.* The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Arch. Gen. Psychiatry* **60**, 497–502 (2003).
127. Kieseppa, T., Partonen, T., Haukka, J., Kaprio, J. & Lonnqvist, J. High concordance of bipolar I disorder in a nationwide sample of twins. *Am. J. Psychiatry* **161**, 1814–1821 (2004).
128. Hettema, J. M., Neale, M. C. & Kendler, K. S. A review and meta-analysis of the genetic epidemiology of anxiety disorders. *Am. J. Psychiatry* **158**, 1568–1578 (2001).
129. Sullivan, P. F., Neale, M. C. & Kendler, K. S. Genetic epidemiology of major depression: review and meta-analysis. *Am. J. Psychiatry* **157**, 1552–1562 (2000).
130. Kendler, K. S., Gatz, M., Gardner, C. O. & Pedersen, N. L. A Swedish national twin study of lifetime major depression. *Am. J. Psychiatry* **163**, 109–114 (2006).
131. O'Donovan, M. C. *et al.* Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nature Genet.* **40**, 1053–1055 (2008).
132. Purcell, S. M. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752 (2009).
133. Shi, J. *et al.* Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* **460**, 753–757 (2009).
134. Stefansson, H. *et al.* Common variants conferring risk of schizophrenia. *Nature* **460**, 744–747 (2009).
135. Wray, N. R. & Visscher, P. M. Narrowing the boundaries of the genetic architecture of schizophrenia. *Schizophr. Bull.* **36**, 14–23 (2010).
136. Karayiorgou, M. *et al.* Schizophrenia susceptibility associated with interstitial deletions of chromosome 22q11. *Proc. Natl Acad. Sci. USA* **92**, 7612–7616 (1995).
This foundational study shows the importance of a rare *de novo* copy number variant (CNV) in conferring susceptibility to schizophrenia.
137. Sebat, J. *et al.* Strong association of *de novo* copy number mutations with autism. *Science* **316**, 445–449 (2007).
138. Stefansson, H. *et al.* Large recurrent microdeletions associated with schizophrenia. *Nature* **455**, 232–236 (2008).
139. Xu, B. *et al.* Strong association of *de novo* copy number mutations with sporadic schizophrenia. *Nature Genet.* **40**, 880–885 (2008).
140. Levy, D. *et al.* Rare *de novo* and transmitted copy-number variation in autistic spectrum disorders. *Neuron* **70**, 886–897 (2011).
141. O'Roak, B. J. *et al.* Exome sequencing in sporadic autism spectrum disorders identifies severe *de novo* mutations. *Nature Genet.* **43**, 585–589 (2011).
142. Sanders, S. J. *et al.* Multiple recurrent *de novo* CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* **70**, 863–885 (2011).
143. Chubb, J. E., Bradshaw, N. J., Soares, D. C., Porteous, D. J. & Millar, J. K. The DISC locus in psychiatric illness. *Mol. Psychiatry* **13**, 36–64 (2008).
This paper reports the identification of a single gene in a Scottish family that increases risk for both schizophrenia and bipolar disorder.
144. Williams, H. J. *et al.* Most genome-wide significant susceptibility loci for schizophrenia and bipolar disorder reported to date cross-traditional diagnostic boundaries. *Hum. Mol. Genet.* **20**, 387–391 (2011).
145. Long, A. D., Mullaney, S. L., Mackay, T. F. & Langley, C. H. Genetic interactions between naturally occurring alleles at quantitative trait loci and mutant alleles at candidate loci affecting bristle number in *Drosophila melanogaster*. *Genetics* **144**, 1497–1510 (1996).
This study describes the development of the quantitative complementation test as a tool for QTL validation.
146. Mackay, T. F. Quantitative trait loci in *Drosophila*. *Nature Rev. Genet.* **2**, 11–20 (2001).
147. Toma, D. P., White, K. P., Hirsch, J. & Greenspan, R. J. Identification of genes involved in *Drosophila melanogaster* geotaxis, a complex behavioral trait. *Nature Genet.* **31**, 349–353 (2002).
148. Sambandan, D., Carbone, M. A., Anholt, R. R. & Mackay, T. F. Phenotypic plasticity and genotype by environment interaction for olfactory behavior in *Drosophila melanogaster*. *Genetics* **179**, 1079–1088 (2008).
149. Ayroles, J. F. *et al.* Systems genetics of complex traits in *Drosophila melanogaster*. *Nature Genet.* **41**, 299–307 (2009).
150. Matsui, A., Go, Y. & Niimura, Y. Degeneration of olfactory receptor gene repertories in primates: no direct link to full trichromatic vision. *Mol. Biol. Evol.* **27**, 1192–1200 (2010).

Acknowledgements
We thank P. McGrath, S. Flavell and E. Glater for discussions.

Competing interests statement
The authors declare no competing financial interests.

FURTHER INFORMATION
The *Drosophila* Genetic Reference Panel (DGRP): http://mackay.gnets.ncsu.edu/MackaySite/DGRP.html
Mouse Genome Informatics: http://www.informatics.jax.org

ALL LINKS ARE ACTIVE IN THE ONLINE PDF