Research Focus

Ain't misbehavin'? Genotype–environment interactions and the genetics of behavior

Trudy F.C. Mackay[1,3] and Robert R.H. Anholt[1,2,3]

[1] Department of Genetics, North Carolina State University, Raleigh, NC 27695, USA
[2] Department of Zoology, North Carolina State University, Raleigh, NC 27695, USA
[3] W.M. Keck Center for Behavioral Biology, North Carolina State University, Raleigh, NC 27695, USA

Although the multiple interacting genes affecting complex traits can readily be dissected, how much genotype–environment interactions contribute to variation in complex traits remains elusive. A recent study that quantified several behavioral phenotypes on the same mouse strains in different laboratories decades apart, shows that some behavioral differences between laboratories remain greatly replicable, whereas others are less robust over time. This report, together with studies from *Drosophila*, stresses the importance of understanding genotype–environment interactions.

Behavior in health and evolution

Behaviors are quantitative traits and their genetic underpinnings are complex. Phenotypic variation between individuals is attributable to the segregation of genetic polymorphisms at many interacting loci, with effects that are sensitive to the environment (Box 1). Understanding the genetic and environmental influences on the manifestation of behavior is important from the dual perspectives of human health and evolutionary biology. Psychiatric disorders such as depression and schizophrenia, in addition to addictive patterns of substance abuse, exact a heavy burden on health care in modern societies. Development of effective treatments depends on understanding the genetic mechanisms and environmental triggers that predispose to these disorders. From the evolutionary perspective, behaviors are components of fitness because they mediate survival through food localization and predator avoidance, and procreation through mate finding, reproduction and parental care. Thus, understanding the genetic basis of variation in behaviors can inform our understanding of adaptive evolution.

How reproducible are measures of behavior?

The behavioral literature is replete with examples of inconsistent results between different laboratories comparing the same genotypes of model organisms for ostensibly the same behavior [1], and of inability to replicate human association studies for behavioral traits. There are many possible reasons for such inconsistencies. First, the ability reliably to detect a difference in behavior between two genotypes (the genotypic effect) depends not on the absolute value of the effect in the units in which the behavior is measured, but the effect scaled by the standard deviation of the variation in environments experienced (the square root of V_E). In other words, large differences between genotypes are likely to be repeatable [2], but smaller differences might not be replicable if sample sizes are too small. Second, differences in genetic background could modify the effect of a mutation or natural polymorphism. Third, behavioral measurements could be affected by significant genotype–environment ($G \times E$) interactions.

Recently, Douglas Wahlsten and colleagues [3] performed a meta-analysis of behavioral data acquired by different laboratories on the same inbred mouse strains over a period of up to 50 years. Behavioral phenotypes included measures of locomotor activity, ethanol preference, and behavior on the elevated plus maze (an apparatus designed to assess anxiety-like behavior). This meta-analysis corroborated results from a previous study by Crabbe and colleagues [1], who subjected eight strains of mice – five commonly used inbred strains, an F2 population derived from two of the strains, and a mutant strain in which the 5-hydroxytryptamine 1B (5-HT-1B) receptor had been knocked out, and its control – to five behavioral tests in three laboratories at different locations. They observed significant differences between laboratories and significant laboratory by genotype interactions for measures of endogenous locomotor behavior, the increase in locomotor behavior following a cocaine injection, and performance in the elevated plus maze. Like the study by Crabbe *et al.* [1], Wahlsten and his colleagues [3] found that the effect of genotype was greatly significant for all behaviors. This is reassuring, and indicates that differences between genotypes are replicable among locations and over many decades. However, $G \times E$ interaction was pervasive in both studies [1,3]. Wahlsten and colleagues [3] showed that the $G \times E$ interaction effect for locomotor activity and ethanol preference averaged 17% and 18% of the main effects, respectively, and the average cross-environment genetic correlations were great (0.89 for both traits). By contrast, anxiety-like behavior assessed in the elevated plus maze varied markedly for the same genotypes tested in different environments. Thus, alcohol preference and locomotion showed greater phenotypic robustness than measurements of anxiety-like behavior. Because different genotypes were responsible for the interaction effects of different behaviors, these observations cannot be explained by one or two genotypes having extraordinary sensitivity to environmental

Corresponding author: Mackay, T.F.C. (trudy_mackay@ncsu.edu).
Available online 5 April 2007.

change. It is important to note that the presence of $G \times E$ interaction means that particular pairs of strains could present different results in two laboratories, even for traits with high cross-environment genetic correlations.

Other traits, other organisms

$G \times E$ interaction is by no means confined to behavioral traits. Significant $G \times E$ interaction was also observed for mouse body weight [1] and brain size [3], with effects that were 17% and 10% of the main effects, respectively (considering only the two recent studies of brain weight). Notably, the relative contribution of $G \times E$ interaction to phenotypic variation for complex traits was minimized in these studies, because the animals were exposed to a limited range of environments typical of standard laboratory practice [1,3].

$G \times E$ interactions become more important when environmental conditions are deliberately altered. In *Drosophila*, a population of 98 recombinant inbred lines used to map quantitative trait loci (QTLs) by linkage to molecular markers was reared in several different temperature environments and culture conditions [4–8]. Between 47.3% [5] and 67.5% [8] of the QTLs affecting sensory bristle number exhibited $G \times E$, which accounted for ~14% of the total genetic variance for sensory bristle number [5,8]. The contribution of $G \times E$ is even more dramatic for fitness-related traits: 22 of 23 QTLs affecting longevity [6,7] and five of six QTLs affecting reproductive performance [4] showed environment-specific effects. Work in *Drosophila* revealed even more complex $G \times E$ effects when it became clear that epistatic interactions can also be environment-specific [8,9].

Studies mapping genes associated with complex traits in humans often suffer from lack of reproducibility. Whereas some studies are underpowered, difficulty in reproducing well-executed studies with good sample sizes in different populations might be caused by $G \times E$ or genetic background-dependent epistatic effects. For near-mendelian effects, this might not be a major problem, but it is a big hurdle for identifying disease susceptibility genes of subtle effects. Indeed, the promise of personalized medicine is predicated on not only the existence of substantial $G \times E$ but also an eventual understanding of its mechanistic basis. Although these studies are still in their infancy, several striking observations have already been made that hint at the generality of $G \times E$ interactions and their prevalence in determining human behavior and disease susceptibility. The low-activity allele of the monoamine oxidase A gene is associated with a tendency to develop violent behavior, but only if the individual was abused as a child [10]. Similarly, the serotonin transporter gene has functional long- and short-length polymorphisms in its promoter; individuals with one or two copies of the short allele exhibit more symptoms of depression if they had experienced stressful life events than individuals homozygous for the long allele [11]. Finally, several polymorphisms are significantly associated with the development of asthma-related phenotypes in children who grew up in sheltered environments, but are not associated with disease status in children who were previously exposed to allergenic environments [12].

Concluding remarks

Context-dependent effects of genes affecting variation in complex traits have been generally regarded as a nuisance, leading to difficulties in replicating results across studies, and in the worst case leading to rejection of important genetic effects that are significant in a particular environment, but not when averaged over the constellation of environments represented in a study. However, the discovery that context-dependent effects can be common motivates research to understand their origin. Can we determine whether particular categories of traits are more or less likely to exhibit $G \times E$, and what are the relevant environmental triggers of the interactions? What is the genetic basis of $G \times E$ interaction?

Box 1. Nature, nurture and their interaction

Like all complex traits, the behavior of an individual (its phenotypic value, P) is attributable to its unique genotype at all loci affecting the behavior (its genotypic value, G), the particular environmental circumstances to which the individual was exposed up to and during the measurement of the behavior (the environmental effect, E) and any interaction between genotype and environment ($G \times E$): $P = G + E + (G \times E)$ [14]. The $G \times E$ term occurs if a specific alteration in the environment has different effects on different genotypes. Thus, the variation in behavior among individuals in a population (V_P) has three major components, variation in genotypes (V_G), variation in environments experienced (V_E) and variation caused by the $G \times E$ interaction ($V_{G \times E}$): $V_P = V_G + V_E + V_{G \times E}$ [14].

When a quantitative measure of behavior can be obtained, it is possible to assess the relative contributions of nature (genetic variation) and nurture (environmental variation) to variation in individual behavior by statistical measures of the similarity of the behavior of related individuals (parents and offspring, siblings, twins and homozygous genotypes in model organisms). These measures are typically summarized as a ratio of genetic to total phenotypic variation, $V_G / V_P = V_G / (V_G + V_E)$, called the heritability [14,15]. Estimates of heritability of behavioral traits in humans and model organisms indicate the importance of both genetic and environmental sources of variation [16,17].

However, these analyses cannot estimate the relative contribution of $G \times E$ to variation in behavior. This requires replicate measures of the behavior of the same group of genotypes in at least two environments. In humans, this can be done if we have defined alleles of genes affecting a behavior, whereas in model organisms the genotypes are typically mutant alleles or lines that have been inbred to homozygosity. Factorial analysis of variance (ANOVA [18]) is then used to assess the significance of the G, E and $G \times E$ terms (Figure I). $G \times E$ can occur if a particular difference in the environment affects some genotypes more than others, and if the rank order of the genotypes changes in the different environments. This is most easily visualized in plots of reaction norms (Figure I). For the simple case of two environments, the reaction norm of each genotype is the line joining the mean performance of the genotype in the two environments. If the reaction norms are parallel, there is no $G \times E$; if the reaction norms are not parallel and instead fan out or cross, there is $G \times E$. Alternatively, one can consider the behavior of a given genotype in two environments as two genetically correlated traits; $G \times E$ is evident if the cross-environment genetic correlation is less than one [19]. This can be visualized by a bivariate plot of the mean performance of the genotypes in the two environments, where $G \times E$ appears as a deviation from a perfect correlation (Figure I).
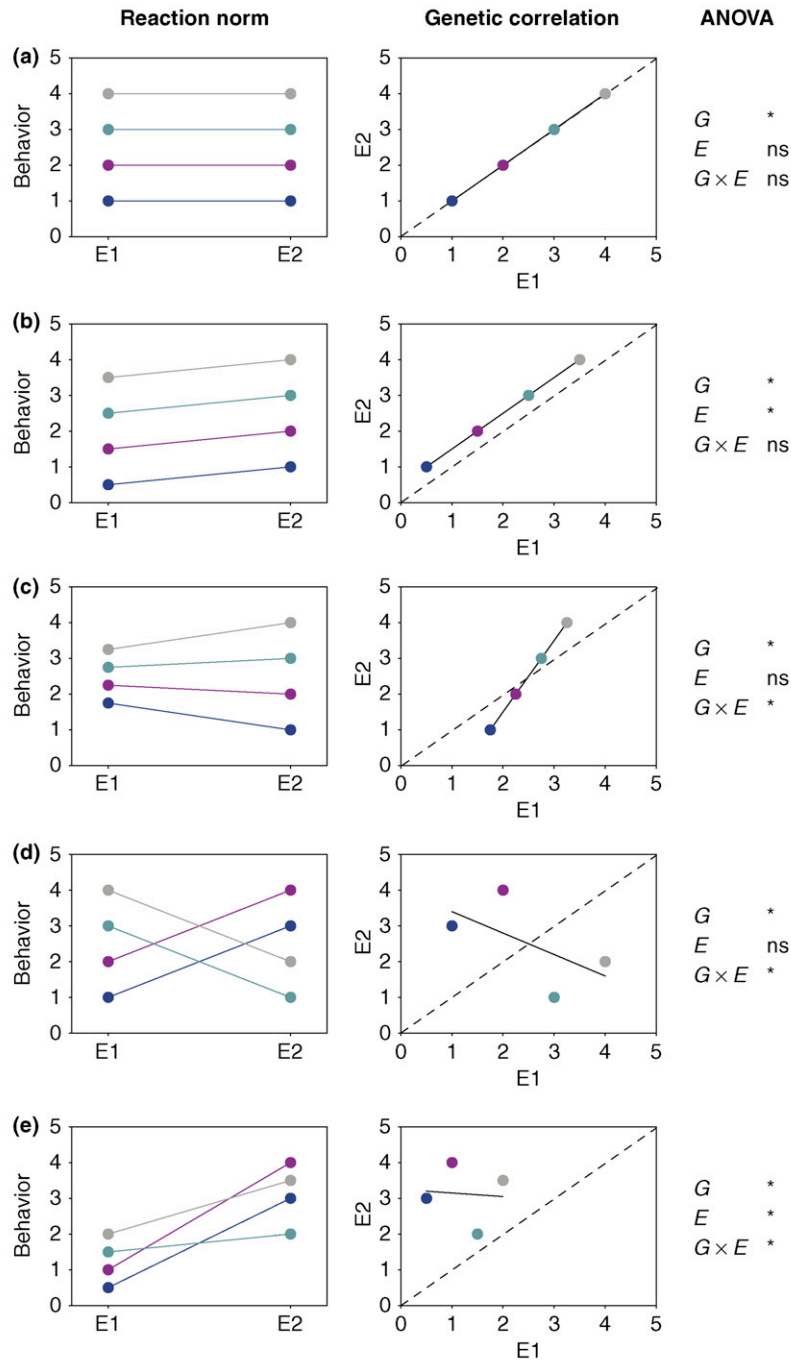
Figure I. Relationships between reaction norms, cross-environment genetic correlations and results of two-way ANOVA depicting significance of the main effects of genotype (G), environment (E) and the genotype–environment interaction ($G \times E$). E1 and E2 represent two environments, and four genotypes are color-coded. Asterisks indicate statistical significance; ns indicates a nonsignificant result. **(a)** Significant differences between genotypes, but the effects are identical in both environments. **(b)** Significant differences between genotypes and environments, but the rank order and relative magnitude of the genotype effects are constant across environments. This represents a situation in which there is phenotypic plasticity, but no genetic variation in phenotypic plasticity ($G \times E = 0$). **(c)** Significant differences between genotypes, but not environments. Here the variation among genotypes is greater in E2 than E1, although the rank order is preserved. The difference in genetic variance between environments gives rise to significant $G \times E$. **(d)** Significant differences between genotypes, but not environments. In this situation the magnitude of genetic variation is the same in both environments, but the rank order of the genotypes changes, leading to significant $G \times E$. **(e)** All of the ANOVA terms are significant, indicating significant genetic variation, phenotypic plasticity and genetic variation in phenotypic plasticity. $G \times E$ is attributable to changes in both variance and rank order of genotypes in the two environments.

There are at least two possible models that are not mutually exclusive [13]. First, alleles could have varying effects on the phenotype in different environments (the 'allelic sensitivity' hypothesis). In this case, the genes affecting trait sensitivity to environmental variation will colocalize with the genes affecting variation in the trait. Second, regulatory genes could respond to specific environmental cues to turn on or modulate the expression of structural genes that directly influence the phenotype (the 'gene regulation' hypothesis). In this case, the genes

affecting trait sensitivity to environmental variation will be distinct from those that affect variation in the trait in a specific environment. In either case, we must ask whether $G \times E$ interactions are specific for each trait, or whether there are pleiotropic effects on $G \times E$ for multiple traits. How do $G \times E$ interactions affect epistatic interactions? What are the relative contributions of changes in genetic variance and changes in rank order of genotype performance across environments?

Understanding the genetic basis of $G \times E$ interactions is the next frontier in the analysis of complex traits. The study by Wahlsten *et al.* [3] urges us to rethink the context of reproducibility of genetic experiments on complex traits. One should not *a priori* dismiss well-executed studies that cannot be repeated precisely in a different population, but rather try to understand the mechanistic basis for the observed irreproducibility in the context of $G \times E$ and interactions between loci (epistasis). In this sense, the mice studied by Crabbe *et al.* [1] and Wahlsten *et al.* [3] were not necessarily misbehaving.

References

1 Crabbe, J.C. *et al.* (1999) Genetics of mouse behavior: interactions with laboratory environment. *Science* 284, 1670–1672
2 Pfaff, D. (2001) Precision in mouse behavior genetics. *Proc. Natl. Acad. Sci. U. S. A.* 98, 5957–5960
3 Wahlsten, D. *et al.* (2006) Stability of inbred mouse strain differences in behavior and brain size between laboratories and across decades. *Proc. Natl. Acad. Sci. U. S. A.* 103, 16364–16369
4 Fry, J.D. *et al.* (1998) QTL mapping of genotype-environment interaction for fitness in *Drosophila melanogaster*. *Genet. Res.* 71, 133–141
5 Gurganus, M.C. *et al.* (1998) Genotype-environment interaction for quantitative trait loci affecting sensory bristle number in *Drosophila melanogaster*. *Genetics* 149, 1883–1898
6 Vieira, C. *et al.* (2000) Genotype-environment interaction for quantitative trait loci affecting lifespan in *Drosophila melanogaster*. *Genetics* 154, 213–227
7 Leips, J. and Mackay, T.F.C. (2000) Quantitative trait loci for lifespan in *Drosophila melanogaster*: Interactions with genetic background and larval density. *Genetics* 155, 1773–1788
8 Dilda, C.L. and Mackay, T.F.C. (2002) The genetic architecture of *Drosophila* sensory bristle number. *Genetics* 162, 1655–1674
9 Sambandan, D. *et al.* (2006) Dynamic genetic interactions determine odor-guided behavior in *Drosophila melanogaster*. *Genetics* 174, 1349–1363
10 Caspi, A. *et al.* (2002) Role of genotype in the cycle of violence in maltreated children. *Science* 297, 851–854
11 Caspi, A. *et al.* (2003) Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 301, 386–389
12 Ober, C. and Thompson, E.E. (2005) Rethinking genetic models of asthma: the role of environmental modifiers. *Curr. Opin. Immunol.* 7, 670–678
13 Via, S. *et al.* (1995) Adaptive phenotypic plasticity: consensus and controversy. *Trends Ecol. Evol.* 10, 212–217
14 Falconer, D.S. and Mackay, T.F.C. (1996) *Introduction to Quantitative Genetics* (4th edn), Longman
15 Lynch, M. and Walsh, B. (1998) *Genetics and Analysis of Quantitative Traits,* Sinauer
16 McClearn, G.E. and DeFries, J.C. (1973) *Introduction to Behavioral Genetics,* Freeman
17 Roff, D.A. and Mousseau, T.A. (1987) Quantitative genetics and fitness: lessons from *Drosophila*. *Heredity* 58, 103–118
18 Sokal, R.R. and Rohlf, F.J. (1981) *Biometry,* Freeman
19 Falconer, D.S. (1960) Selection of mice for growth on high and low planes of nutrition. *Genet. Res.* 1, 91–113

Chromatin regulation and sex determination in *Caenorhabditis elegans*

Edward E. Large and Laura D. Mathies

Department of Genetics, North Carolina State University, Raleigh, NC 27695-7614, USA

Glioma-associated oncogene (GLI) transcription factors function downstream of the hedgehog signal transduction pathway to regulate the development of many animals. Although the nematode *Caenorhabditis elegans* lacks a hedgehog pathway, it does have a GLI protein that represses male development in favor of hermaphrodite development. As we discuss here, recent findings implicate two conserved transcription-repressor complexes in the repression of male-specific genes. This research indicates a possible conserved role for these complexes in either GLI-directed gene repression or sex determination.

Sex determination in *Caenorhabditis elegans*

Caenorhabditis elegans develop into one of two sexes: hermaphrodites, which are self-fertile females, or cross-fertile males (Figure 1). The disparate phenotypic and behavioral differences between males and hermaphrodites enabled researchers to screen rapidly for sex-determination mutants. As a result, sex determination was the first genetic pathway to be described in *C. elegans* [1]. A recent study by Grote and Conradt revealed a new facet of sex determination in *C. elegans*, with the discovery that novel and previously defined transcription-repressor complexes also function to control sexual development [2].

The sex-determination pathway is a negative regulatory cascade that culminates in the activation of the transformer-1 (*tra-1*) gene in hermaphrodites. *tra-1* encodes two protein isoforms, TRA-1A and TRA-1B; TRA-1A is responsible for all known functions of *tra-1* and is most closely related to zinc-finger transcription factors, glioma-associated oncogene (GLI) in vertebrates and cubitus interruptus (CI) in *Drosophila* [3]. The CI and GLI proteins in *Drosophila* and vertebrates, respectively, both function

Corresponding author: Mathies, L.D. (ldmathies@ncsu.edu).
Available online 20 April 2007.