ANNUAL REVIEWS

Annual Review of Genomics and Human Genetics

How Natural Genetic Variation Shapes Behavior

Natalie Niepoth and Andres Bendesky

Zuckerman Mind Brain Behavior Institute and Department of Ecology, Evolution, and Environmental Biology, Columbia University, New York, NY 10027, USA;
email: natalie.niepoth@columbia.edu, a.bendesky@columbia.edu

Annu. Rev. Genom. Hum. Genet. 2020. 21:437–63

First published as a Review in Advance on April 13, 2020

The *Annual Review of Genomics and Human Genetics* is online at genom.annualreviews.org

https://doi.org/10.1146/annurev-genom-111219-080427

Keywords

animal behavior, genetic variation, population genetics, quantitative trait locus mapping, evolution, ethology

Abstract

Nervous systems allow animals to acutely respond and behaviorally adapt to changes and recurring patterns in their environment at multiple timescales—from milliseconds to years. Behavior is further shaped at intergenerational timescales by genetic variation, drift, and selection. This sophistication and flexibility of behavior makes it challenging to measure behavior consistently in individual subjects and to compare it across individuals. In spite of these challenges, careful behavioral observations in nature and controlled measurements in the laboratory, combined with modern technologies and powerful genetic approaches, have led to important discoveries about the way genetic variation shapes behavior. A critical mass of genes whose variation is known to modulate behavior in nature is finally accumulating, allowing us to recognize emerging patterns. In this review, we first discuss genetic mapping approaches useful for studying behavior. We then survey how variation acts at different levels—in environmental sensation, in internal neuronal circuits, and outside the nervous system altogether—and then discuss the sources and types of molecular variation linked to behavior and the mechanisms that shape such variation. We end by discussing remaining questions in the field.

1. INTRODUCTION

Forward genetics screen: the artificial introduction of mutations throughout the genome to identify the mutations and genes that affect a trait of interest

Reverse genetics: an approach to analyze the function of a specific gene by altering its coding sequence or expression

Heritability: the fraction of phenotypic variance in a trait within a population that is attributable to genetic variation

Animal behavior is characterized by its complexity. It is generated by the integration of sensory cues with internal states to direct motor output via precise signaling in sophisticated neuronal circuits. These circuits are remarkably malleable and are constantly remodeled developmentally and by experience and learning, allowing animals to adapt to both recurring patterns and changes in their environment. Behavior is further influenced by innate variation in neuronal anatomy and function. Thus, behaviors are plastic within individuals throughout their lives as well as variable among individuals.

Behavior can be difficult to measure, particularly in natural settings, where the conditions animals experience over their lifetimes are difficult to control. Even under controlled laboratory environments, behavior is notoriously susceptible to subtle environmental perturbations (29). These challenges make it difficult to measure the environmental and genetic variables that influence behavior. Therefore, our knowledge of the genetic underpinnings of behavior lags behind what we currently know about morphology, physiology, and disease risk. However, technological and statistical methods for studying genetic contributions to behavior are advancing quickly, unlocking new opportunities. Though much is still unknown, patterns in the field have begun to emerge. We have reached an opportune moment to study these patterns and make inferences about the larger processes that govern the evolution of behavior.

This review surveys examples of natural genetic variants that modulate behavior within and among populations of a species and that contribute to differences in behavior among species. We focus on three common targets of genetic influences on behavior: sensation of environmental cues, higher-order processing in the central nervous system, and interactions with environmental molecules outside of the nervous system. We then discuss the molecular types of variants observed and how these variants arise and are maintained in populations. We conclude with a summary of emerging patterns in the field and outstanding questions.

2. MEASURING THE GENETIC BASIS OF BEHAVIOR IN NATURAL POPULATIONS

Analyses of the physiological roles of genes on behavior started with the pioneering forward genetics screens of Seymour Benzer in *Drosophila melanogaster* in the 1970s. More recently, powerful reverse genetics tools have been applied to laboratory animals—mainly nematodes, flies, and mice—to study specific genes. However, these forward and reverse genetics approaches reveal little about the genetic bases of behavioral variation in nature. For example, the Mouse Genome Informatics database contains more than 10,000 examples of artificial mutations that affect behavior in laboratory mice (17). If these mutations were to arise spontaneously in nature, many would be too detrimental for survival and thus quickly removed from populations. Therefore, these mutations likely do not represent the types of variation that occur and segregate in nature.

What is clear, however, is that variation in most natural behaviors has a substantial genetic component. A sweeping meta-analysis of 17,804 traits from 2,748 twin study publications showed that the heritability of behavioral traits in humans is comparable to that of nonbehavioral traits (117). Most human behavioral traits studied are 30–60% heritable, though this estimate may be skewed by ascertainment and publication biases; for example, geneticists may choose to measure the heritability of traits they think are likely to have a genetic component. In nonhuman animals, heritability is also similar among behavioral, life history, and morphological traits (135), although these estimates may suffer from similar biases.

Pinpointing the specific genes and variants that contribute to trait heritability is a central goal of behavioral genetics. Different approaches have been classically used to parse this genetic

Table 1 Methods for identifying genetic loci linked to behavior

Method	Scale of comparison	Mapping resolution	Sample size required	Advantages	Disadvantages
GWAS	Within populations	High (to scale of LD)	Large	Can study in nature	Has biases from population stratification, cannot capture associations with rare variants, also measures indirect effects
QTL mapping	Within pedigrees, between populations or sister species	Low or moderate (depending on the number of recombination events)	Moderate	Can use controlled laboratory conditions; family structure avoids stratification	Potentially overestimates effect sizes, can generate false negatives from closely linked variants with opposite effects, captures only effects of variants present in founders of cross
Population differentiation scan	Between populations or closely related species	High (to scale of LD)	Small	Can study in nature, provides evidence of selection	Is agnostic to phenotypic traits
Comparative gene expression	Within and between populations or species	Transcript level	Small	Offers temporal and tissue-specific insights into cellular function	Is agnostic to genetic variation; many correlated genes will not contribute directly to behavior

Abbreviations: GWAS, genome-wide association study; LD, linkage disequilibrium; QTL, quantitative trait locus.

component of behavior into contributory variants and the genes they affect. In recent years, quantitative genetics research has encouragingly shifted from candidate gene studies to analyses of variants throughout the whole genome. Unbiased approaches such as genome-wide association studies (GWASs), quantification of population differentiation, and quantitative trait locus (QTL) mapping are powerful methods for discovering genetic loci associated with variation in behavior (**Table 1**).

2.1. Genome-Wide Association Studies

GWASs test for association between a trait (such as a specific behavior) and genetic variants that are common in populations (frequency > ~5%). However, most common variants have exceedingly small effects, so large samples—on the order of tens to hundreds of thousands of subjects—are usually necessary to gain sufficient power (70, 121). In recent years, public consortia and private companies have compiled genetic samples and phenotypic information from more than 4 million individuals (e.g., 5, 20, 108). This data accessibility has enabled the analysis of dozens of human behavioral traits, and in the past 10 years, more than 700 publications have reported the results of GWASs for specific human behaviors (18). These publications demonstrate that nearly all behaviors are highly polygenic: An individual behavior within a population is influenced by many genetic loci, each of mostly small effect.

Candidate gene study: a test of the association between variants in one or a few predefined genes and a trait

GWASs are not designed to detect associations with rare variants, yet nearly all human genetic variants are rare: More than 90% of variants segregating in human populations have minor allele

frequencies below 1% (83). Because rare variants tend to have larger effects than common variants (53, 93), they may be important contributors to the behaviors of specific individuals or families where these variants segregate. The extent of the contribution of rare variants to the heritability of behavior within populations is still an open question.

F_{ST}: a statistical measure of the difference in genetic variance within versus between populations

Genetic associations in traditional GWAS designs measure not only direct genetic effects of variants on people in the study but also the indirect contributions of these variants through their effects on relatives with whom the subjects interact (78, 157). For example, variants may affect the behavior of subjects indirectly by modulating parental behaviors. Future studies that jointly analyze the genotypes of test subjects and their parents will help to alleviate these problems and provide better estimates of direct and indirect genetic effects—both of which are necessary for a comprehensive understanding of behavioral variation (156). GWASs are also subject to false-positive associations resulting from population stratification and assortative mating, which are difficult to fully control for using statistical methods (156). For example, studies of the use of chopsticks in a city would likely capture genetic variants common in Asians even if none of these variants directly affect chopstick use (82).

2.2. Population Differentiation

Populations often differ in their behavior, yet it is very difficult to estimate the contribution of genetic differences to such behavioral variation. People with mixed ancestry from these populations can be used to find associations between local genomic ancestry and behavior in an approach called admixture mapping. Because genomic segments from different populations are present in admixed individuals, the argument is that this curtails problems of environmental and genetic interaction confounders that occur when estimating the effect of genes in two separate populations. However, ancestry proportions in admixed individuals have been found to correlate with socioeconomic status (21), which can result from familial and cultural contingencies as well as discrimination and can severely confound genetic analyses. The proportion of ancestry can also be correlated with many other environmental variables, such as diet, religion, and education, further confounding behavioral inferences from genetics.

Another powerful population-genetics approach to study differences among populations involves scanning the genomes of two or more populations that differ in a behavior of interest. Regions that are particularly differentiated between populations (i.e., outliers in F_{ST}-based statistics) may contain variants that explain trait variation. For example, a genome-wide F_{ST} scan between two populations of warblers identified a candidate gene contributing to the choice in winter migration to either Central or South America (143). Because populations often differ in more than one trait, differentiated regions may affect not the behavior of interest but rather a correlated trait. Nevertheless, these F_{ST}-based methods can be useful in identifying candidate genomic regions for local adaptation that can be further studied using other approaches.

2.3. Quantitative Trait Locus Mapping

An alternative method to probe the genetic basis of behavioral variation is QTL mapping. The goal of QTL mapping is to identify loci that cosegregate with a trait in families or in experimental crosses. The family structure of these mapping populations generally avoids issues with population stratification. Laboratory crosses allow for careful behavioral measurements while affording control of environmental conditions and biological variables such as age. However, QTL mapping crosses usually originate from small numbers of founding animals, limiting the number of haplotypes that can be analyzed and failing to fully capture natural allele frequency distributions. Moreover, because many experimental crosses comprise only a few generations where meiotic

recombination can take place, mapping resolution is usually low. Advanced intercross schemes greatly improve mapping resolution (124), and QTLs can also be fine-mapped using follow-up targeted crosses. Multiple genes and specific genetic variants affecting behavior have been found using QTL mapping and follow-up fine-mapping, particularly in *Caenorhabditis elegans* and *Drosophila melanogaster* (9, 36, 102). Even when QTL mapping does not lead to the identification of causative variants, important inferences can be made about the genetic architecture of behavior, including estimation of pleiotropy and sex-specific effects as well as quantification of effect sizes (7, 42, 149).

2.4. Comparative Gene Expression

Unbiased whole-transcriptome analysis technologies, such as RNA sequencing (RNA-seq), allow for correlating gene expression with variation in a behavior within populations or between populations or species. While this is a common approach, it is not always a successful method for identifying causal genes. For instance, comparing expression in relevant tissues between two species with different innate behaviors can identify thousands of differentially expressed genes. Thus, complementary methods such as QTL mapping or experimental manipulations are usually required to narrow down the list of correlated genes to the most likely candidates (e.g., 7). For example, an *Aedes aegypti* mosquito olfactory receptor (OR) was first identified through RNA-seq comparisons between mosquitoes attracted to humans and those attracted to other animals; the expression of this gene was then found to correlate with behavior in a cross between these two types of mosquitos (101).

RNA-seq can be a powerful method in carefully designed comparisons of specific organs or tissues, brain regions, or cell types. For example, RNA-seq was used to identify the molecular bases of the evolution of infrared sensors in snakes and vampire bats. TRPA1 channels of pit vipers and some species of boas and pythons harbor mutations that make them very heat sensitive (46). Moreover, these channels are expressed at much higher levels in the trigeminal ganglia—which innervate the heat-sensitive pit—than in the dorsal root ganglia, which transmit other somatosensory information (46). TRPA1 is not upregulated in the trigeminal ganglia in snakes without infrared-detecting pits (46). Vampire bats have evolved high trigeminal ganglion expression of an isoform of the heat-sensing TRPV1 that is particularly heat sensitive. This high TRPV1 isoform expression was also identified by measuring alternative splicing using RNA-seq in vampire bats and fruit bats (45).

2.5. Artificial Selection on Behavior

The methods described above are typically applied to study behavioral variation within or among natural populations. Behavioral differences can also be exaggerated through artificial selection over many generations to create strains with extreme behavior. The genetic differences among selected lines can then be probed by genome-wide F_{ST} scans, through QTL mapping, or by comparing gene expression in relevant tissues or brain areas. Molecular signatures of selection can also be searched for within a selected line. These types of artificial selection studies have pointed to genetic regions and specific genes implicated in aggression in various species, including flies (35), rats (1, 55), and foxes (79).

3. IDENTIFYING GENES THAT MODULATE BEHAVIOR

There are no genes that specify behavior (122); rather, genetic variation modulates biochemical and cellular pathways and shapes neuronal circuits that ultimately give rise to behavior. Genetic

variation can therefore affect behavior by acting at different levels: by altering sensory perception (Section 3.1), by modulating higher-order circuits of the central nervous system (Section 3.2), or by affecting metabolic processes outside of the nervous system (Section 3.3).

As more genetic variants are identified, important evolutionary questions can begin to be answered. For example, are particular classes of genes or biological processes more often implicated in behavioral variation and evolution? And how often does independent evolution of similar behaviors converge on the same molecular pathways? Though we still know very little about the genetic mechanisms underlying most behaviors, the case studies highlighted in the following sections provide some of the first clues to the answers.

3.1. Genetic Variation in Sensory Systems Alters Behavioral Responses to External Cues

Before environmental information is processed by the nervous system, signals must be detected by sensory receptors that are often housed in specialized sensory organs. Genetic variation that alters sensory detection—for example, by affecting the function of these receptors—can cause a direct behavioral shift by disrupting signal input. Alternatively, variation downstream from receptors can affect how a stimulus is processed after it is detected.

Mutations in sensory receptors are characteristic of behavioral shifts in natural populations. Cockroaches taste sugar using hair-like sensory structures protruding from their mouthparts that house neurons expressing taste receptors. Most cockroaches, like other insects, are attracted to sugars, including glucose; however, several populations of cockroaches have recently evolved glucose aversion. Glucose is a component of many commercial cockroach baits designed to poison the animals; therefore, behavioral attraction to glucose has negative fitness consequences in populations of cockroaches under this selection regime. The taste neurons that sense sugars—and mediate attraction—in wild cockroaches have become less sensitive to glucose in cockroaches from populations that avoid glucose, while the neurons that detect bitter compounds—and mediate aversion—have become more sensitive to glucose (145). Thus, mutations affecting sensory neurons have changed the valence of glucose from attractive to aversive.

In humans, food and drink preference is modulated by smell. Some people are genetically predisposed to detecting the taste of cilantro as unpleasantly soapy, which affects diet choice and cilantro preference. A GWAS identified a genetic region significantly associated with this soapy-taste detection contained within a cluster of OR genes on chromosome 11. One such OR, OR6A2, has a high binding specificity for several aldehydes that give cilantro its characteristic odor (39).

Variation in food preference among species can arise from species-specific adaptations in sensory receptors. Most birds, including chickens, turkeys, and finches, have lost the ability to sense sugars, as they lack the sweet taste receptor gene *TAS1R2* (4). Hummingbirds, however, are specialists that feed exclusively on nectar and have regained sugar sensation by repurposing the umami receptor (a dimer encoded by the genes *TAS1R1* and *TAS1R3*). Mutations in these genes transform the receptor from one that detects savory amino acids into one that detects sugars, thereby permitting the characteristic specialization of nectar-feeding behavior in these birds (4).

Some animals communicate using chemicals called pheromones when they signal within species and kairomones when they signal between species. In moths and other insects, females produce and secrete sex pheromones that attract males. Female Asian corn borer moths (*Ostrinia furnacalis*) produce the sex pheromones (*E*)-12- and (*Z*)-12-tetradecenyl acetate, whereas female European corn borer moths (*Ostrinia nubilalis*) produce slightly different isomers, (*E*)-11- and (*Z*)-11-tetradecenyl acetate (65). *O. furnacalis* and *O. nubilalis* males are attracted to the distinct pheromone blend from conspecific females due to a nonsynonymous mutation in the

O. furnacalis pheromone receptor gene *Or3* that reduces *O. furnacalis* male response to the *O. nubilalis* pheromone 14-fold (85). Therefore, genetic variation affecting pheromone receptors can mediate interspecies specificity in mate attraction.

Genetic variation modulating chemical communication also affects behavioral interactions between distantly related species. Some species of nematodes commensally infest live insects; the insects provide the nematode with dispersal opportunities, a food source, and, after the insect dies, a substrate for the nematode to continue its life cycle (75). Some natural populations of the nematode *Pristionchus pacificus* associate with the oriental scarab beetle (*Anomala orientalis*) (54). These beetles produce the chemical (*E*)-11-tetradecenyl acetate, and certain strains of *P. pacificus* nematodes are highly attracted to this kairomone (58). Differences in (*E*)-11-tetradecenyl acetate attraction among *P. pacificus* strains map to variation not in a sensory receptor but rather in the protein kinase EGL-4 (58), a component of the cGMP signaling pathway that regulates olfaction in *C. elegans* (87).

In *C. elegans*, pheromones that accumulate in high local population densities bind to the pheromone receptors SRG-36 and SRG-37 to stimulate progression into an alternative diapause-like state called dauer (103). Under high-density selection regimes in the laboratory, two *C. elegans* strains have independently acquired resistance to dauer progression from nearly identical deletions affecting both *srg-36* and *srg-37* (103). While no wild strain has been identified that harbors deletions affecting both genes, 18% of wild *C. elegans* strains from around the globe harbor a putative loss-of-function deletion within *srg-37* that is identical by descent (86). Interestingly, there is enrichment of this allele in *C. elegans* populations that have colonized a rotting fruit niche (which provides bacteria that worms eat), suggesting that there may be particularly strong selection against dauer formation in worms exploiting resources that support reproductive growth (86).

Loss of sensitivity to environmental molecules may therefore be adaptive in certain habitats and can underlie the evolution of behavior. In highveld mole rats, multiple genetic changes affect both the protein sequence of the TRPA1 channel and the expression of a second channel protein, NALCN, in sensory neurons. These mutations confer insensitivity to the painful substance allyl isothiocyanate, a defensive compound produced by some insects and plants, and thus shapes mole rat behavior by permitting both feeding on pungent food sources and coexistence with aggressive stinging ants (38).

Decision-making must integrate both environmental signals and internal states such as hunger; therefore, sensation of internal cues can be as important as that of external cues. In *C. elegans*, the decision to abandon an area with food is modulated by noncoding variation affecting the G protein–coupled catecholamine receptor gene *tyra-3*. The receptor encoded by this gene is expressed in sensory neurons yet binds internal biogenic amines (tyramine), suggesting that the gene modulates responses to the environment by integrating internal information (9).

3.2. Genetic Variation Alters Behavior by Modulating Central Nervous System Circuitry

Behavior can also be modulated by variation affecting higher-order nervous system processing rather than sensory perception. Pioneering work on the genetic modulation of social behavior implicated variation in neuropeptide receptors in the brain, but more recent examples show that variation in other classes of neuronal molecules is also important for generating the diversity of social behavior observed both within and between species.

The G protein–coupled receptors of the neuropeptides arginine vasopressin (AVP) and oxytocin (OT) are classic examples of genes underlying natural variation in mating systems among

species of voles (160). Molecular approaches comparing the monogamous prairie vole (*Microtus ochrogaster*) and the promiscuous montane vole (*Microtus montanus*) implicated species-specific distribution patterns of AVP and OT receptors in the brain in many behavioral aspects of monogamy, including pair bonding, paternal care, and mate guarding (64). Interestingly, QTL mapping in a different rodent clade—comparing the monogamous oldfield mouse (*Peromyscus polionotus*) and the promiscuous prairie deer mouse (*Peromyscus maniculatus*)—also implicated the AVP system in variation in parental care (7). However, in *Peromyscus* mice, *cis*-regulatory variation affecting the expression of the AVP ligand in the hypothalamus, rather than the AVP receptor, is linked to the elaborate nest building characteristic of monogamous parents.

Cis-regulatory variation: genetic variation that alters the expression of a physically linked gene

Introgression: the transfer of genetic material from one population or species to another by a hybridization event and subsequent backcrossing

Insects also have diverse social structures; for example, some populations of the sweat bee *Lasioglossum albipes* produce solitary nests, whereas others are eusocial. A GWAS approach linked noncoding variation in syntaxin 1a (syx1a), a protein that mediates synaptic vesicle release, to intraspecific variation in eusociality. A single intronic polymorphism of *syx1a* altered its expression in in vitro assays, consistent with in vivo expression differences between social and solitary bees (77).

C. elegans nematodes also vary in their social behaviors. For example, they differ in their propensity to aggregate with each other, a behavior that is influenced by pheromonal communication as well as by environmental variables such as food availability and oxygen levels (48, 57, 98). Differences in this behavior were initially found to be strongly affected by a single amino-acid difference in the receptor *npr-1*, a gene homologous to the mammalian neuropeptide-Y receptor (32). A later study, based on analyses of more than 200 wild *C. elegans* strains, strongly suggested that this mutation arose during the domestication of *C. elegans* in the laboratory (102). Further quantitative genetics approaches implicated naturally occurring polymorphisms affecting the expression of EXP-1, a receptor for the neurotransmitter GABA, in the propensity of *C. elegans* to aggregate with each other (8).

As described in Section 3.1, females of some species attract conspecific males by emitting sex pheromones; however, changes in the valence of sex pheromones—whether they are perceived as attractive or aversive—can arise from genetic differences affecting central neural circuitry rather than peripheral sensory perception. Two species of *Drosophila*, *D. melanogaster* and *D. simulans*, are closely related yet are largely reproductively isolated due to differences in pheromone signaling between the species. *D. melanogaster* females, but not *D. simulans* females, produce the sex pheromone 7,11-heptacosadiene, which is highly attractive to *D. melanogaster* males but aversive to *D. simulans* males (129). Interestingly, sensory neurons respond similarly to 7,11-heptacosadiene in *D. melanogaster* and *D. simulans*, but differences in how the signal is propagated in downstream circuits in the fly brain explain this variation in behavioral response (129).

D. melanogaster and *D. simulans* diverge not only in their pheromone signaling but also in the male courtship songs that attract conspecific females. Differences in aspects of the courtship song between laboratory strains of these species mapped to a retroelement insertion into the intron of the *slowpoke* (*slo*) gene, which encodes a calcium-activated potassium channel expressed throughout the central nervous system (36). The existence or prevalence of this mutation in natural populations is unknown.

Differences in social and mating behaviors may involve the coordination of many genetic variants that are selected upon through subtle changes in allele frequencies at many loci simultaneously (known as polygenic selection). For example, male cichlid fishes of Lake Malawi build bowers that attract females: Certain species dig pit bowers, whereas others build castles. Genetic variants across each of the 22 linkage groups in 20 diverse species of pit diggers and castle builders have elevated F_{ST} values, suggesting that the divergence in genetic architecture between the species is highly complex yet consistent across species, perhaps due to introgression (155). F_1 hybrids

between a pit-digging species and a castle-building species display both bower-building behaviors sequentially: During the pit-digging epoch, a suite of alleles inherited from the pit-digging parent become upregulated in the F_1 brain, while during the castle-building epoch, alleles inherited from the castle-building parent become upregulated. The temporal specificity of this allele-specific expression indicates that modular synchronization of transcriptomic responses can underlie the display of highly complex behaviors (155).

3.3. Variation in Genes Outside the Nervous System Affects Behavior

Variation affecting genes that function outside of the nervous system can also modulate behavior (**Figure 1**). For example, variation in metabolism can affect what and how much an animal chooses to eat or drink. Among mammals, variation in the copy number of the gene for amylase, a digestive enzyme that breaks down starch, correlates with starch preference. The more copies of the amylase gene a species has, the more starch it tends to eat as part of its diet. It is unclear,

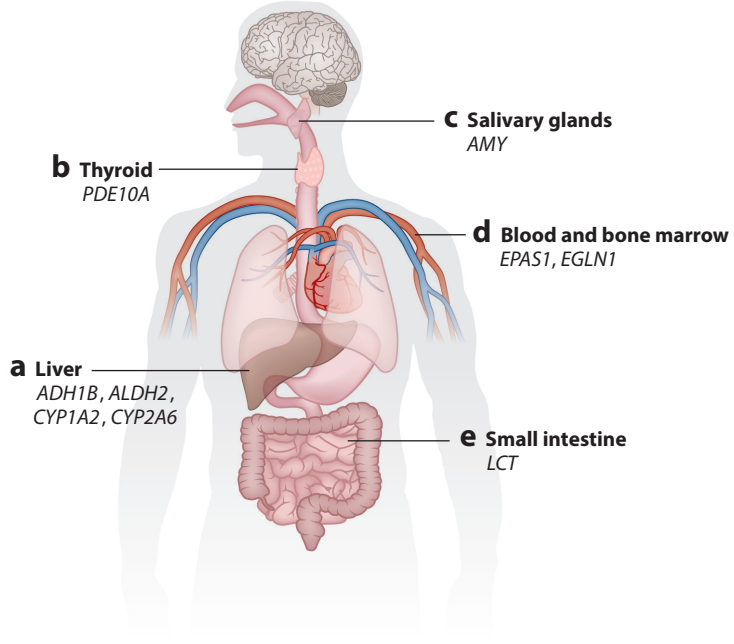

Figure 1

Non-neuronal genes that affect the behavior of humans and other vertebrates. (*a*) Polymorphisms in *ADH1B* and *ALDH2* affect the rate at which alcohol is metabolized in the liver, affecting alcohol dependence (23, 96); noncoding single-nucleotide polymorphisms (SNPs) near the caffeine-metabolizing-enzyme gene *CYP1A2* are associated with increased coffee drinking and caffeine consumption (28); and coding SNPs in the nicotine-metabolizing-enzyme gene *CYP2A6* are associated with cigarette-smoking behavior (140). (*b*) Polymorphisms in the *PDE10A* gene of Bajau people cause increased spleen size, likely by modulating hormones released by the thyroid, thereby allowing specialized diving behaviors (63). (*c*) Copy number variation of the *AMY* gene, whose product metabolizes starch, is correlated with the amount of starch mammals eat (99, 116). (*d*) Variants affecting *EPAS1* and *EGLN1* permit high-altitude adaptation and perhaps habitat preference in humans, *Peromyscus* mice, ducks, and other vertebrates (e.g., 13, 47, 128, 154). (*e*) Noncoding variants cause lactase persistence by prolonging expression of the *LCT* gene in the small intestine into adulthood (11).

EVOLUTION SHAPES BEHAVIOR, BUT BEHAVIOR ALSO SHAPES EVOLUTION

Behavior is modulated by genetic variation, but behavior itself also shapes evolutionary processes. Behavior can promote selection via behavioral drive (100) by exposing individuals to new environments and new selective pressures, thereby accelerating evolution. For example, behavioral flexibility allows individuals to shift their preferences in habitat or food choice, which may influence the evolution of other morphological, metabolic, or behavioral traits that facilitate adaptation to these environments.

Behavioral flexibility also allows individuals to shield themselves from environmental variation, thus reducing selective pressure via behavioral inertia, also known as the Bogert effect (16). Rather than adapting to a changing environment, for example, individuals can search for and move to microhabitats with more preferable conditions, thereby shielding themselves from directional selection. For example, behavioral inertia reduces selection on salamanders because they move between microhabitats to regulate temperature and moisture levels (40). Behavioral drive and inertia may also work simultaneously, as in the case of a tropical lizard in which a microhabitat shift to boulders has driven skull and limb evolution (behavioral drive) while enabling thermoregulation that releases selection pressure against colder environments (behavioral inertia) (105).

however, whether species-specific preference for starch drove the evolution of amylase copy number variation or whether copy number variation preceded starch preference (116) (see the sidebar titled Evolution Shapes Behavior, but Behavior Also Shapes Evolution). In humans, individual differences in salivary amylase level and copy number affect the perception of texture and perhaps even flavor and likely affect starchy food preference (34, 99).

A similar example in humans is the repeated evolution of the ability to digest lactose into adulthood, which affects how much dairy people eat. Most mammals cannot easily digest lactose after weaning; however, variants in and around the lactase gene drive its continued expression in the small intestine into adulthood, facilitating lactose digestion (137). This phenotype arose independently in different pastoral human populations that came to rely on dairy as an important source of nutrition. The haplotype containing variants that contribute to lactose tolerance is identical by descent in Europeans and Indians (43) but different in Africans (142). Both lactose-tolerance haplotypes have experienced a selective sweep over the past 7,000 years (142).

The amount of alcohol and coffee that people drink is strongly modulated by genetic variation. Alcohol is broken down into acetaldehyde and from there into acetic acid by the liver enzymes alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH), respectively. Variants in paralogs of each enzyme, specifically ADH1B and ALDH2, are associated with alcohol consumption and are the common genetic variants with the strongest known effects on human behavior (96). Because people with low-activity *ADH1B* alleles accumulate toxic acetaldehyde more slowly than people with high-activity alleles, they are less prone to the acetaldehyde symptoms characteristic of alcohol excess, including nausea and headache, and have a threefold-higher risk of developing alcoholism (12). Low-activity *ALDH2* variants, which lead to an accumulation of acetaldehyde, have an even stronger effect on alcohol consumption and dependence. These variants are common only in some East Asian populations, whereas *ADH1B* variants are also common in Europeans and Africans (147). Similarly, coffee drinking is shaped by variation in genes involved in the metabolism of caffeine (27).

Selective sweep: a phenomenon in which the frequency of an allele in a population increases rapidly due to strong positive selection

More than 1 billion people worldwide smoke cigarettes, and this behavior is largely mediated by dependence on nicotine, a highly addictive component of tobacco. Variation in the cytochrome P450 2A6 (CYP2A6) liver enzyme, which is essential for nicotine metabolism, is associated with the number of cigarettes smoked per day (92). People who carry versions of CYP2A6 with reduced

activity (and therefore slower metabolism of nicotine) smoke fewer cigarettes per day and usually find it easier to quit smoking (120). However, variation in genes expressed in the central nervous system also modulates smoking: Nearly all nicotinic acetylcholine receptor (nAChR) genes expressed in the brain exhibit variation associated with smoking behaviors (10, 92).

Parallel evolution: an evolutionary process by which a trait evolves in separate taxa via similar molecular mechanisms (e.g., by mutations in the same genes or pathways)

Adaptation to extreme environments can also arise from selection on variation affecting genes outside of the nervous system. Different human populations, including Tibetans and Andeans, have independently colonized extremely high-altitude environments and are genetically adapted to low-oxygen (hypoxic) conditions. High-altitude-adapted Tibetans carry a variant of *EPAS1*, which encodes a transcription factor regulating the production of hemoglobin and the development of new blood vessels, that helps them use oxygen more efficiently at high altitudes (154). Both Tibetans and Andeans have signatures of positive selection on the *EGLN1* gene, whose product interacts with EPAS1 (13). Ducks adapted to high altitudes also carry variants of *EPAS1* and *EGLN1* at higher frequencies than lowland ducks (47), and *Peromyscus* mice harbor molecular signatures of selection at *Epas1* (128), suggesting that high-altitude habitat choice may converge on similar genetic mechanisms in distant species. Mutations in the *C. elegans* homolog of *EGLN1*, *egl-9*, strongly affect preference for high or low oxygen (aerotaxis) (22), suggesting that variation in hypoxia-related genes could also affect vertebrate preference for different oxygen concentrations.

The choice of plants that herbivores eat is influenced by attraction, preference, and resistance to plant defensive compounds. Cardiac glycoside compounds produced by the milkweed plant are toxic for many herbivorous species. However, parallel evolution of cardiac glycoside resistance has permitted feeding on milkweed across many orders of insects (165). For example, the monarch butterfly (*Danaus plexippus*) feeds on milkweed in its larval stage and sequesters the toxic chemical to deter predators as a butterfly. Cardiac glycoside resistance in many species is conferred by three mutations that alter three amino acids in the protein pump Na^+,K^+-ATPase, the molecular target of cardiac glycoside. Two studies have recently used phylogenetic comparative approaches and genetic engineering to prove that the order in which these three mutations evolved matters (69, 138), highlighting how genetic interactions (epistasis) can constrain the paths through which behaviors evolve.

Two *Drosophila* species have independently evolved specialization to the toxic noni fruit *Morinda citrifolia*: a population of *D. yakuba* from the island of Mayotte and the noni specialist *D. sechellia* from the nearby Seychelles archipelago (95, 153). Variation in genomic regions linked to noni fruit adaptation overlaps more often than expected by chance between the Mayotte and Seychelles noni specialists, suggesting a parallel molecular basis to this specialization (153). Mayotte *D. yakuba* showed strong signatures of selection in several detoxification genes compared with mainland *D. yakuba* generalists, including a major toxin tolerance locus previously identified in *D. sechellia* (153). Species-specific attraction to noni fruit in *D. sechellia* is influenced not only by detoxification genes but also by variants affecting olfactory receptor tuning to noni fruit volatile chemicals (118, 119). Transgenic experiments demonstrate that OR22a, which mediates long-range attraction to these volatiles, contains three naturally occurring amino-acid substitutions that each increase sensitivity to noni volatiles (3).

Social behavior is modulated by visual, auditory, mechanical, and chemical signaling between partners. Genetic variation that alters social signals can therefore strongly affect the behavior of animals receiving the signal. Divergence in female sex pheromone synthesis can acutely alter male attraction and promote divergence of male preference over longer timescales. In the European corn borer moth, two populations have begun to diverge in their pheromone signaling, leading to reproductive isolation. Female moths from the *E* population produce a blend of pheromone containing 98% (*E*)-11-tetradecenyl acetate and 2% (*Z*)-11-tetradecenyl acetate, whereas the *Z* population produces 3% (*E*)-11-tetradecenyl acetate and 97% (*Z*)-11-tetradecenyl acetate (84).

This divergence is caused by multiple nonsynonymous substitutions in a single fatty-acyl reductase gene involved in the synthesis of precursors to (*E*)-11- and (*Z*)-11-tetradecenyl acetate (84).

In nature, most *C. elegans* individuals are hermaphrodites with the ability to self-fertilize, while males occur at a frequency of less than 1%. An Australian strain of *C. elegans* exhibits male–male mating behavior caused by a natural loss-of-function mutation in a single gene (*plep-1*) expressed in the excretory pore; males homozygous for the *plep-1* mutation attract copulations from other males (112). This result shows that even behaviors that appear complex can sometimes arise from mutations in single genes.

3.4. Synthesis: Lessons from Genetic Mapping of Behavioral Diversity

In the introduction to this section, we posed two major questions: What types of genes and biological processes does variation most often impact, and do similar behaviors evolve through similar or distinct molecular mechanisms? We do not yet have enough information to fully answer these questions, but we can identify two major patterns. First, genetic variants affecting sensory receptors are very common. Additionally, genetic variation affecting the expression or protein sequence of other classes of genes in the brain, such as neuropeptide and neurotransmitter receptors, also characterizes behavioral divergence. In most cases, however, the specific mechanisms underlying these genetic effects are not well understood, even if a general biological pathway can be implicated in the behavior. Second, despite an expectation that parallel evolution of behavior—the independent evolution of a behavior based on changes in the same genes or pathways—might be rare due to the complexity of the genetic and neuronal bases of behavior, there are many examples of parallelism both within and among species. For example, the same hypoxia-inducible factor pathway is involved in adaptation to high altitude in humans, deer mice, and ducks, and variation in the vasopressin system affects monogamous behaviors in both voles and deer mice. Additionally, the courtship song of Hawaiian crickets has evolved in three independent pairs of species through changes at overlapping QTLs, suggesting parallel evolution (14).

Importantly, demonstrating that particular genetic variants influence behavior through their effects on specific genes (quantitative trait genes) is very challenging. The gold standard is the reciprocal hemizygosity test (and related tests) (6, 134), which is rarely performed outside the powerful genetic model organisms *Drosophila* and *C. elegans*. Thus, many studies implicate genes based on protein-coding changes, allele-specific expression differences, proximity to mapped variants, and experimental manipulations. Each of these approaches has limitations, but together they can provide more convincing evidence for the effects of particular genes on behavior.

4. THE SOURCE AND MAINTENANCE OF GENETIC VARIATION CONTRIBUTING TO BEHAVIORAL DIVERSITY

Genetic variation ranges from that affecting single nucleotides, to additions or deletions of thousands to millions of bases, to large-scale chromosomal rearrangements. Each type of variant can affect the regulation of genes or alter protein sequence. Furthermore, they can have long-term evolutionary consequences due to the acquisition of novel function or linkage of genetic variants across generations from suppression of recombination. In this section, we outline the types of genetic variation found to affect behavior and consider how they appear and are maintained in populations.

4.1. Types of Molecular Variation That Modulate Behavior

Mutations affecting a small genetic region, such as single-nucleotide polymorphisms (SNPs) and short insertions or deletions (indels), are a major source of variation that can affect protein

sequence or gene regulation. Other, larger-scale mutations, such as supergenes and gene expansions, are also prominent contributors to behavioral evolution.

4.1.1. Regulatory versus coding mutations. Mutations can alter the temporal and spatial regulation of gene expression or modify protein-coding sequences themselves. Both types of molecular changes have been shown to contribute to behavioral diversity. Protein-sequence changes may be particularly important in the evolution of sensory receptor tuning to various environmental cues. For example, coding variation in taste receptors permits hummingbird attraction to sugar (4), while sweet taste receptor genes in many carnivorous mammals have been pseudogenized (67, 90) (**Figure 2**). However, protein-coding changes in genes that are widely expressed in the central nervous system can detrimentally disrupt essential networks, while regulatory variants alter gene expression more modularly (144). Therefore, regulatory mutations rather than coding mutations are likely the primary type of variants affecting genes that are broadly expressed in the brain or are essential for neural development (144). For example,

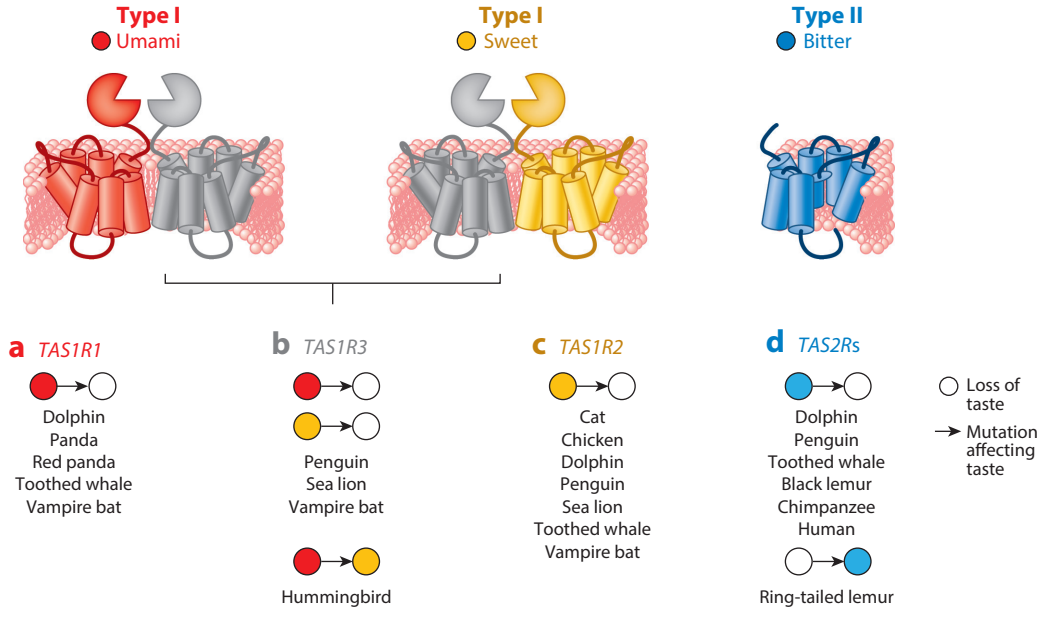

Figure 2

Molecular evolution of vertebrate taste receptor genes. In most vertebrates, the proteins TAS1R1 and TAS1R3 dimerize to form the umami taste receptor, while TAS1R2 and TAS1R3 dimerize to form the sweet taste receptor. Type II taste receptors (TAS2Rs) primarily detect bitter taste. Genetic variation across the animal kingdom has altered the protein structure and function of these genes, causing loss of function (pseudogenization or other coding changes) or gain of novel functionality, which alters taste perception. (*a*,*c*) Independent pseudogenization of the *TAS1R1* and *TAS1R2* genes caused loss of umami and sweet taste perception, respectively, in many taxa (41, 60, 67, 89, 90, 130, 162–164). (*b*) Pseudogenization of the *TAS1R3* receptor conferred loss of both umami and sweet taste perception in penguins, sea lions, and vampire bats (4, 67, 162, 163), while coding mutations in *TAS1R3* in hummingbirds transformed the TAS1R1/TAS1R3 heterodimer from detecting umami to detecting sweet tastes. (*d*) Pseudogenization of *TAS2R* receptors in dolphins, penguins, and toothed whales caused loss of bitterness perception (41, 67, 162). Lemur-specific amino-acid substitutions in TAS2R16 changed the receptor response to arbutin from agonism to inverse agonism, thereby reducing sensitivity to salicin bitterness. However, ring-tailed lemurs regained the ability to recognize arbutin as a TAS2R16 agonist via coding mutations affecting the *TAS2R16* sequence (66). Independent mutations to *TAS2R38* eliminate sensitivity to the bitter compound phenylthiocarbamide in some chimpanzees and humans (151).

Table 2 Supergenes that affect animal behavior

Animal	Locus	Haplotype length	Behaviors affected	Nonbehavioral traits affected	References
White-throated sparrow (*Zonotrichia albicollis*)	*ZAL2*	98 Mb	Parental care, singing, mate preference, aggression, courtship	Plumage color	139, 141
Ruff (*Philomachus pugnax*)	*Faeder/Satellite*	4.5 Mb	Mating strategy, territoriality	Plumage color, body size	80, 81
Fire ant (*Solenopsis invicta*)	*SB/Sb*	13 Mb	Tolerance of multiple queens, aggression	Body size, queen fecundity	72, 148
House mouse (*Mus musculus*)	*t* haplotype	40 Mb	Dispersal, migration	Spermatogenesis manipulation (meiotic drive)	97, 126, 131

regulatory variation affecting *slo*, a gene expressed ubiquitously in the fly brain, contributes to the evolution of the *Drosophila* courtship song (36).

4.1.2. Supergenes. Variation affecting behavior tends to be spread across the genome, where chromosomal segregation and recombination unlink variants that have beneficial effects on traits. Genomic rearrangements that prevent recombination ensure that a block of the genome is inherited together and can therefore spread in a population. These supergenes can accumulate further genes and variants. Supergenes have strong effects on multiple behaviors across species (**Table 2**). In fire ants, a 13-Mb supergene contributes to variation in social organization (148). Contained within each of the two nonrecombining supergene alleles, social B (*sB*) and social b (*sb*), are specific variants of the gene *Gp-9*, which encodes an odorant-binding protein that dictates whether colonies will accept multiple queens (125). The supergene alleles also confer a difference in colony-level aggression (125). Honeybees from highland and lowland populations in East Africa have rampant gene flow between them, with the exception of two haplotype blocks on two chromosomes that result from inversions (146). Many genes within these supergene-like haplotypes influence honeybee behavior that may be adaptive in these divergent environments; for example, one haplotype contains nearly all of the octopamine receptor genes in the honeybee genome, and these genes play essential roles in learning and foraging behavior (146).

Variation in mating behavior in the ruff, a wading bird, is also caused by linked variation within a 4.5-Mb inversion. The ruff has three male morphs (independents, faeders, and satellites) that differ in behavior, color, and size, representing three lekking strategies during which males aggregate and compete for access to females. Independents have retained the ancestral genotype (no inversion), while the faeder allele arose from an initial inversion and the satellite allele likely originated from an unlikely recombination event between faeder and independent alleles (81). Another avian species, the white-throated sparrow, contains a chromosomal rearrangement at the *ZAL2* locus. Genetic variants within the *ZAL2* inversion ($ZAL2^{m}$) increase the expression of estrogen receptor α (*ESR1*) in specific brain regions, which causes heightened aggression (104). Sparrows containing the $ZAL2^{m}$ allele also have alternate territorial song, nestling-provisioning, and mate-guarding behaviors compared with sparrows that do not contain the inversion (59, 166).

4.1.3. Gene expansions. While behavioral diversity due to supergenes acts within species, a major source of evolutionary divergence among species is large-scale gene expansion. Gene duplication can relax selective constraint on one of the copies and allow the gain of novel

functionality (neofunctionalization) from new mutations in paralogs. Across species, large expansions or contractions of gene repertoires can shape species-specific behavior. There are many well-documented expansions of sensory gene repertoires, suggesting that sensory gene evolution has been a steadfast process powering behavioral evolution. Sensory gene radiation across mammals has occurred to the greatest extent in olfactory and vomeronasal receptors (110, 158). Because of combinatorial olfactory perception for most odorants—a regime under which individual receptors participate in the detection of specific odors but are neither necessary nor sufficient—olfactory genes may be particularly mutable across deep evolutionary timescales and can be prime sources of genetic variation affecting behavior.

Hundreds of gains and losses of OR genes have occurred across different lineages of reptiles and mammals (110). For example, primates have fewer than 400 functional OR genes, while dogs and rodents have two and three times as many OR genes, respectively. Variation in the number of functional OR genes among different lineages appears to be driven by ecological adaptation. In birds and reptiles, for example, patterns of OR expansion correlate with the ecological requirements of the lineage. In diverse bird species, specific OR family expansions coincide with aquatic adaptations (water birds), vocal learning, and land specialization (73). The expansion of *OR5*, *OR8*, and *OR9* occurred in both predatory birds and alligators, suggesting an adaptive role for those genes in carnivory (73). Surprisingly, large expansions of these genes are actually linked to herbivory in mammals (62). Though neutral evolutionary processes likely contribute somewhat to the rapid duplication and pseudogenization of ORs (109), it has been shown that ORs in great apes are under selective constraints (44). Furthermore, the correlation between OR evolution and ecological requirements suggests that at least some families of OR genes are likely under positive selection in diverse animal taxa and that neofunctionalization of ORs may play an important role in behavioral adaptation.

Aquatic mammals, such as whales, are characterized by a reduction in ORs relative to their land ancestors, concordant with the evolution of other sensory modalities, such as echolocation (76, 109). Similarly, a reduction of functional OR genes in primates may be related to their acquisition of three-dimensional color vision (trichromacy) due to adaptive variation in pigments (called opsins) that allow vision in vertebrates (71). In primates, color vision likely has important consequences for behaviors such as foraging, mate choice, predator avoidance, and navigation (71). The most light-sensitive type of opsin—responsible for vision in dimly lit conditions—is the rod opsin gene rhodopsin (*RH1*), and most vertebrate taxa possess just a single *RH1* gene. However, three deep-sea teleost lineages have independently gained additional copies of *RH1* (91, 107), suggesting that these expansions have permitted these lineages to live in the deep sea. The deep-sea silver spinyfin in particular has expanded its RH1 repertoire to 38 rod opsins, the largest number known in any vertebrate (107). Protein regeneration and simulation have shown that these spinyfin RH1s are tuned to a wide range of light wavelengths that encompass the bioluminescence spectrum of the deep sea, suggesting this expansion may allow spinyfins to better perceive bioluminescent signals important for adaptation.

Vomeronasal receptors that bind pheromones are also among the fastest-evolving genes in mammals and have gone through huge expansions in some species of rodents and loss of all functional genes in catarrhine primates and dogs (159, 161). Mice have not only more than 1,000 ORs but also more than 350 vomeronasal receptors (V1Rs and V2Rs) that allow specialized olfaction of pheromones essential for regulating social behaviors such as mating, parenting, and aggression (74, 152). The Lake Victoria cichlid fish *Haplochromis chilotes* also has an expanded repertoire of vomeronasal type II receptor-like genes (*OlfC* genes), which has been suggested to contribute to its extraordinary feeding behavior diversification by allowing for the detection of a wide range of amino acids (111).

4.2. Sources of Variation Contributing to Behavior

Genetic variation fundamentally arises through mutation and spreads within and between populations through migration and mating. New mutations, standing genetic variation, and gene flow between populations and species are important sources of variation that contribute to behavioral evolution.

4.2.1. New mutations. The ultimate source of genetic variation is new mutations. While most mutations are deleterious and disappear quickly, some are maintained in the population at low frequencies. A third, rare outcome is the selective sweep, whereby a beneficial mutation spreads rapidly due to positive selection. A classic example of a selective sweep of a behavior-modulating variant is the spread of lactase persistence alleles in Europeans (Section 3.3). These alleles were not detected in ancient DNA samples from early Neolithic Europeans, suggesting that they arose recently (19). In horses, a mutation in the gene *DMRT3*, encoding a transcription factor that affects the differentiation of spinal cord interneurons, likely arose within the last 10,000 years (133). While most horses with the ancestral *DMRT3* allele have a limited locomotive repertoire (walk, trot, or gallop), horses containing this recent variant of *DMRT3* exhibit unusual gaited locomotive patterns (2). This variant was artificially selected for by humans, presumably based on its interesting effect on horse locomotion, producing tens of gaited horse breeds that exist today.

4.2.2. Standing genetic variation. Selection can alter the allele frequencies of either new mutations or preexisting genetic variation in the population. Standing genetic variants may persist at low frequencies in the population in the absence of selection and then segregate at intermediate frequencies in response to soft sweeps, genetic drift, or balancing selection. Due to cryptic genetic variation, variants that confer small or no phenotypic effects in particular environments can allow for behavioral adaptation when environments change (115).

Selection on standing genetic variation underlies variation in schooling behavior between marine and freshwater stickleback fishes. Sticklebacks from marine populations overwhelmingly carry the ancestral allele of the gene *Eda*, but the alternate allele persists in the population at low frequencies and has repeatedly become fixed in many independent populations that have colonized freshwater habitats (24). Marine and freshwater sticklebacks differ in various aspects of their schooling behavior, including the angles of their bodies during schooling. Differences in this body position map to variation at the *Eda* locus (52), and follow-up transgenic experiments confirmed the functional effect of *Eda* expression on schooling behavior variation (51).

It has been argued that soft sweeps on standing genetic variation are more common in human adaptation (including behavioral adaptation) than hard sweeps following new beneficial mutations (127). For example, the *PDE10A* allele that increases spleen size and helps breath-holding diving in the Bajau is present in 37% of Bajau people but less than 7% of people in closely related populations (63). In humans, genes involved in central nervous system development appear to be particularly enriched for adaptation from standing genetic variation (127).

4.2.3. Gene flow and adaptive introgression. Hybridization with other populations or other species can also introduce genetic variation that affects behavior. Neanderthals and Denisovans colonized Europe and Asia hundreds of thousands of years before modern humans left Africa (56). When modern humans expanded out of Africa, they mixed with Neanderthals and Denisovans, and gene flow from those archaic humans provided modern humans with genetic variants that facilitated their adaptation to their new environments. For example, the *EPAS1* gene in Tibetans, which now permits high-altitude living (Section 3.3), was introgressed from Denisovans (61).

Present-day Europeans also bear genomic signatures of gene flow with Neanderthals, and introgressed Neanderthal DNA affects many behavioral traits, including sleeping patterns, mood, and smoking (31).

4.3. Maintenance of Behavioral Variation Within Species

Evolutionary processes, including stabilizing selection and purifying selection, tend to reduce genetic variation in populations over time. Alternatively, balancing selection and selfish genetic elements are two mechanisms by which genetic variation can be maintained within populations and species.

4.3.1. Balancing selection. Local adaptation of populations to different environments can give rise to multiple behavioral strategies that are maintained in the species as a whole by balancing selection (88). Because behavior is particularly adaptable to local environments (Section 5.1), balancing selection may affect behavior more than it affects other traits. Indeed, genome-wide scans in *Drosophila* have identified an enrichment of neuronal genes under balancing selection, including genes involved in sensory perception, olfactory behavior, aggression, circadian behaviors, and neurodevelopment (25, 30).

Variation affecting the brain expression pattern of the neuropeptide vasopressin receptor gene *Avpr1a* modulates social and mating behaviors in different species of voles (Section 3.2). Moreover, multiple alleles of the *Avpr1a* gene are actively maintained within populations of the prairie vole (*Microtus ochrogaster*) by balancing selection (113). While male prairie voles are socially monogamous, they are not sexually exclusive: Extrapair copulations are common, though the degree of sexual infidelity varies among individuals (114). Innate variation in this behavior likely reflects a fitness trade-off between paternity gained from mating with multiple partners and paternity lost from poor mate guarding. Alleles within and near the *Avpr1a* gene affect its expression in the retrosplenial cortex (a brain region), and this variation in expression influences the degree of male sexual fidelity (113). Therefore, multiple *Avpr1a* alleles that cause specific levels of neural expression drive population-level variation in mating behavior. Allele frequencies of SNPs affecting *Avpr1a* expression were highly skewed toward an excess of intermediate frequencies, suggesting that these variants are actively maintained in the population through balancing selection (113).

Similarly, the length of microsatellites in regulatory regions of *Avpr1a* and the oxytocin receptor (*Oxtr*) in the bank vole (*Myodes glareolus*) influences the expression of these genes in the brain, leading to intraspecific differences in social behavior (94). Polymorphisms in the length of the *Avpr1a* microsatellite are maintained through balancing selection, as the fitness of each genotype depends on both population density and sex (94).

Balancing selection is also characteristic of several genes that modulate foraging behavior in *C. elegans*. In these animals, the G protein–coupled receptor SRX-43 detects pheromones that signal population density and subsequently suppresses foraging behavior (49); however, some individuals possess genetic variants that lower the sensitivity of SRX-43, and these worms tend to continue foraging even at high population densities. Polymorphisms affecting SRX-43 sensitivity therefore give rise to two foraging strategies that have different fitness effects in different environments, suggesting that they are maintained by balancing selection. Worms with high sensitivity to pheromones that bind SRX-43 have a competitive advantage over worms with low sensitivity in food-limited environments, but this advantage disappears in environments with patchy food, where the exploratory low-sensitivity worms are better able to find food patches and thus outperform high-sensitivity worms (49, 50).

4.3.2. Selfish genetic elements. Selfish genetic elements exploit a range of molecular tactics to bias their own transmission, including segregation distortion, transposition, and male sterility. Interestingly, selfish genes and chromosomes can also ensure their transmission by altering behavior.

The selfish motivation of sex chromosomes can favor the evolution of behaviors that promote their own transmission. In some animal taxa, including Lepidoptera, birds, and some fish and reptiles, a ZW chromosome system determines the sex of offspring. In this system, the W chromosome is inherited only by females, as females carry one copy of the Z chromosome and one copy of the W chromosome, while males carry two copies of the Z chromosome. Recent theoretical models suggest that the selfish evolutionary interest of the W chromosome can drive the evolution of female mating preferences for harmful male traits, such as sexually selected ornamental handicaps and even male parental care (106), therefore reducing the fitness of male offspring while comparatively increasing that of female offspring and improving selfish transmission of W chromosomes to females.

Behavior can evolve as a strategy to resolve intragenomic conflict. Some natural populations of the wild house mouse (*Mus musculus domesticus*) carry a selfish genetic element called the *t* haplotype. This 40-Mb haplotype, which is linked by inversions, induces the death of non-*t* sperm, thus promoting its inheritance to future generations (97, 131). However, the *t* haplotype is homozygous lethal (37), so its fitness declines as its frequency increases in the population and it cannot sweep to fixation. Male *t* heterozygotes are also poor sperm competitors (136); therefore, the *t* haplotype is most fit in smaller populations, where multiple copulation by females (polyandry) is more limited. Theory and modeling suggest that the *t* haplotype should be selected to increase migration propensity away from natal populations and thus increase its chance of transmission in smaller satellite populations where *t* frequency is low (126). Measurements in the field have shown that the emigration propensity of *t* carriers is indeed higher than it is in mice that do not carry the *t* haplotype (126), suggesting that either the *t* haplotype itself contains elements that promote emigration or the rest of the genome has evolved a behavioral resolution to genetic conflict.

5. CONCLUSIONS

5.1. The Evolvability of Behavior

Is behavior more evolutionarily labile than other traits? Phylogenetic patterns across species indeed suggest that behavior may be particularly evolvable (15). For example, in primates, the phylogenetic signal—the conservation of a trait among lineages across evolutionary time—is typically lower for behavioral traits such as diet choice, sociability, and foraging patterns than for morphological and life-history traits (68). A phylogeny of *Polyrhachis* ants contains many evolutionary transitions of highly intricate social nest-weaving behavior (123), suggesting that even complex behaviors can readily evolve in different species.

Sensory receptors are encoded by some of the most evolutionarily labile genes in the animal kingdom (Sections 3.1 and 4.1.3), perhaps allowing for rapid evolution of signal perception while bypassing potential negative pleiotropy of genetic change to higher-order circuits. The types of natural genetic variation that affect behavior are nonetheless incredibly diverse: An individual behavior may be modulated by many types of genes either inside or outside of the nervous system (Section 3). However, certain systems may be more adaptable than others, promoting evolutionary parallelism (Section 3.4). Population-level mechanisms that maintain genetic diversity can provide the variation necessary for rapid evolution. In particular, balancing selection likely plays an essential role in maintaining behavioral variation by preserving multiple alleles in a species

(Section 4.3.1). Furthermore, standing and cryptic genetic diversity provides an adaptive substrate for selection when environmental pressures change (Section 4.2.2).

5.2. Emerging Patterns and Outstanding Questions

There is an extraordinary diversity of behavior across the animal kingdom, and we still have much to learn about the genetic contributions to such diversity. However, a few general patterns are beginning to emerge. In general, many genes and many genetic variants contribute to specific behaviors, and these variants can affect gene regulation or protein sequence. Tentatively, protein-coding changes appear to be enriched in genes that interact with environmental molecules to modulate behavior, such as those encoding sensory receptors and enzymes. Furthermore, there are many examples of genetic variation affecting sensory systems, but it is not yet clear whether this represents a primary source of adaptation or is merely a system where genetic effects can be more easily detected or dissected.

An important remaining question is to what extent the genetic architecture of behavior differs from those of nonbehavioral traits. Unlike other quantitative traits, such as metabolite concentrations or gene-expression levels, behavioral traits are not discrete molecules that can be measured, but are rather more arbitrary constructs whose magnitude and scale depend on how they are defined and measured. Thus, it is difficult to quantitatively compare the number and effect size of loci associated with behavior with those of other traits. Qualitatively, however, the genetic architecture of behavior appears to be similar to those of other traits: Multiple loci of small effects usually contribute to variation in behavior within species and among closely related species.

On the other hand, emerging evidence suggests that balancing selection is a particularly important evolutionary force shaping the function of the brain and behavioral patterns compared with other traits (25, 30). In addition, certain molecular events, such as large-scale changes in particular classes of genes and the contribution of supergenes, appear to be particularly prominent in behavioral evolution. How these forces and molecular mechanisms constrain or facilitate behavioral evolution remains an open question.

5.3. The Future of Behavioral Genetics

Whereas genetic mapping approaches have yielded many loci linked to behavioral traits, a pressing issue in behavioral genetics is how to identify the genes affected by the variants linked to behavioral variation. A common approach is to assume that the gene closest to the peak of linkage or association is the causative (quantitative trait) gene. Benchmarking using well-curated molecular traits indicates that 70% of causative genes are closest to peaks of association in GWASs (132), but this proximity might be lower for behavioral traits because neuronal genes tend to have highly elaborate regulatory mechanisms (33, 150). Thus, other lines of evidence are necessary to implicate specific genes in trait variation. The gold standard is the reciprocal hemizygosity test (134), but this test is difficult to perform in animals that lack powerful genetic tools.

With the increasing number of behavioral GWASs in humans and the development of polygenic scores to predict traits, some might be tempted to use such scores to study the genetic bases of behavioral differences among populations. However, because of gene–gene and gene–environment/culture interactions, population stratification, and lack of knowledge about causal variants (in most cases we know only of associated haplotypes in specific populations), translating polygenic scores estimated in one population to another is highly problematic (26).

Convincing cases of genetic contributions to differences in behavior between human populations have identified peaks of genetic differentiation (e.g., 63, 154). These studies also find evidence

that variation in the trait within a population is associated with polymorphisms within these peaks of genetic differentiation that fall near genes implicating specific biological functions (63). These three pillars—loci strongly differentiated among populations of interest, association of loci with behavior within populations, and functional evidence supporting links between variants or genes and behavior—are good guideposts for future population-genetics-based studies in humans and other species.

Our knowledge of behavioral variation has been historically limited to select groups (i.e., to laboratory model species or, in human genetics, to European populations). However, novel and low-cost technologies now allow geneticists to study essentially any species, which can help to answer questions about preferred targets of behavioral diversity (parallelism) and to discover new genes that affect behavior. New methods for gene editing in nonmodel organisms may also advance our understanding of the biological mechanisms underlying variation in behavior. Expanding genetic analyses of behavior to other human populations will have substantial impacts on psychiatric genetics and on public health throughout the world.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

A.B. is supported by National Institutes of Health grant HD084732, the Searle Scholars Program, a Klingenstein-Simons Fellowship, and a Sloan Foundation Fellowship. We thank Patrick McGrath, Josh Greene, and Hopi Hoekstra and members of her laboratory for suggestions on earlier versions of this article.

LITERATURE CITED

1. Albert FW, Carlborg Ö, Plyusnina I, Besnier F, Hedwig D, et al. 2009. Genetic architecture of tameness in a rat model of animal domestication. *Genetics* 182:541–54
2. Andersson LS, Larhammar M, Memic F, Wootz H, Schwochow D, et al. 2012. Mutations in *DMRT3* affect locomotion in horses and spinal circuit function in mice. *Nature* 488:642–46
3. Auer TO, Khallaf MA, Silbering AF, Zappia G, Ellis K, et al. 2020. Olfactory receptor and circuit evolution promote host specialization. *Nature* 579:402–8
4. Baldwin MW, Toda Y, Nakagita T, O'Connell MJ, Klasing KC, et al. 2014. Evolution of sweet taste perception in hummingbirds by transformation of the ancestral umami receptor. *Science* 345:929–33
5. Baselmans BML, Jansen R, Ip HF, van Dongen J, Abdellaoui A, et al. 2019. Multivariate genome-wide analyses of the well-being spectrum. *Nat. Genet.* 51:445–51a
6. Bendesky A, Bargmann CI. 2011. Genetic contributions to behavioural diversity at the gene-environment interface. *Nat. Rev. Genet.* 12:809–20
7. Bendesky A, Kwon Y-M, Lassance J-M, Lewarch CL, Yao S, et al. 2017. The genetic basis of parental care evolution in monogamous mice. *Nature* 544:434–39
8. Bendesky A, Pitts J, Rockman MV, Chen WC, Tan MW, et al. 2012. Long-range regulatory polymorphisms affecting a GABA receptor constitute a quantitative trait locus (QTL) for social behavior in *Caenorhabditis elegans*. *PLOS Genet*. 8:e1003157
9. Bendesky A, Tsunozaki M, Rockman MV, Kruglyak L, Bargmann CI. 2011. Catecholamine receptor polymorphisms affect decision-making in *C. elegans*. *Nature* 472:313–18
10. Berrettini W, Yuan X, Tozzi F, Song K, Francks C, et al. 2008. α-5/α-3 nicotinic receptor subunit alleles increase risk for heavy smoking. *Mol. Psychiatry* 13:368–73

11. Bersaglieri T, Sabeti PC, Patterson N, Vanderploeg T, Schaffner SF, et al. 2004. Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* 74:1111–20
12. Bierut LJ, Goate AM, Breslau N, Johnson EO, Bertelsen S, et al. 2012. *ADH1B* is associated with alcohol dependence and alcohol consumption in populations of European and African ancestry. *Mol. Psychiatry* 17:445–50
13. Bigham A, Bauchet M, Pinto D, Mao X, Akey JM, et al. 2010. Identifying signatures of natural selection in Tibetan and Andean populations using dense genome scan data. *PLOS Genet.* 6:e1001116
14. Blankers T, Oh KP, Shaw KL. 2019. Parallel genomic architecture underlies repeated sexual signal divergence in Hawaiian Laupala crickets. *Proc. Biol. Sci.* 286:20191479
15. Blomberg SP, Garland T, Ives AR. 2003. Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution* 57:717–45
16. Bogert CM. 1949. Thermoregulation in reptiles, a factor in evolution. *Evolution* 3:195–211
17. Bult CJ, Blake JA, Smith CL, Kadin JA, Richardson JE. 2019. Mouse Genome Database (MGD) 2019. *Nucleic Acids Res.* 47:D801–6
18. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, et al. 2019. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* 47:D1005–12
19. Burger J, Kirchner M, Bramanti B, Haak W, Thomas MG. 2007. Absence of the lactase-persistence-associated allele in early Neolithic Europeans. *PNAS* 104:3736–41
20. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, et al. 2018. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562:203–9
21. Campbell DD, Parra MV, Duque C, Gallego N, Franco L, et al. 2012. Amerind ancestry, socioeconomic status and the genetics of type 2 diabetes in a Colombian population. *PLOS ONE* 7:e33570
22. Chang AJ, Bargmann CI. 2008. Hypoxia and the HIF-1 transcriptional pathway reorganize a neuronal circuit for oxygen-dependent behavior in *Caenorhabditis elegans*. *PNAS* 105:7321–26
23. Clarke TK, Adams MJ, Davies G, Howard DM, Hall LS, et al. 2017. Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK biobank ($N = 112117$). *Mol. Psychiatry* 22:1376–84
24. Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G Jr., Dickson M, et al. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307:1928–33
25. Comeron JM. 2014. Background selection as baseline for nucleotide variation across the *Drosophila* genome. *PLOS Genet.* 10:e1004434
26. Coop G. 2019. Reading tea leaves? Polygenic scores and differences in traits among groups. arXiv:1909.00892 [q-bio.GN]
27. Cornelis MC, Byrne EM, Esko T, Nalls MA, Ganna A, et al. 2015. Genome-wide meta-analysis identifies six novel loci associated with habitual coffee consumption. *Mol. Psychiatry* 20:647–56
28. Cornelis MC, Monda KL, Yu K, Paynter N, Azzato EM, et al. 2011. Genome-wide meta-analysis identifies regions on 7p21 (AHR) and 15q24 (CYP1A2) as determinants of habitual caffeine consumption. *PLOS Genet.* 7:e1002033
29. Crabbe JC, Wahlsten D, Dudek BC. 1999. Genetics of mouse behavior: interactions with laboratory environment. *Science* 284:1670–72
30. Croze M, Wollstein A, Božičević V, Živković D, Stephan W, Hutter S. 2017. A genome-wide scan for genes under balancing selection in *Drosophila melanogaster*. *BMC Evol. Biol.* 17:15
31. Dannemann M, Kelso J. 2017. The contribution of Neanderthals to phenotypic variation in modern humans. *Am. J. Hum. Genet.* 101:578–89
32. de Bono M, Bargmann CI. 1998. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. *Cell* 94:679–89
33. de la Torre-Ubieta L, Stein JL, Won H, Opland CK, Liang D, et al. 2018. The dynamic landscape of open chromatin during human cortical neurogenesis. *Cell* 172:289–304
34. De Wijk RA, Prinz JF, Engelen L, Weenen H. 2004. The role of α-amylase in the perception of oral texture and flavour in custards. *Physiol. Behav.* 83:81–91
35. Dierick HA, Greenspan RJ. 2006. Molecular analysis of flies selected for aggressive behavior. *Nat. Genet.* 38:1023–31

36. Ding Y, Berrocal A, Morita T, Longden KD, Stern DL. 2016. Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature* 536:329–32
37. Dunn LC, Gluecksohn-Schoenheimer S. 1950. Repeated mutations in one area of a mouse chromosome. *PNAS* 36:233–37
38. Eigenbrod O, Debus KY, Reznick J, Bennett NC, Sánchez-Carranza O, et al. 2019. Rapid molecular evolution of pain insensitivity in multiple African rodents. *Science* 364:852–59
39. Eriksson N, Wu S, Do CB, Kiefer AK, Tung JY, et al. 2012. A genetic variant near olfactory receptor genes influences cilantro preference. *Flavour* 1:22
40. Farallo VR, Wier R, Miles DB. 2018. The Bogert effect revisited: Salamander regulatory behaviors are differently constrained by time and space. *Ecol. Evol.* 8:11522–32
41. Feng P, Zheng J, Rossiter SJ, Wang D, Zhao H. 2014. Massive losses of taste receptor genes in toothed and baleen whales. *Genome Biol. Evol.* 6:1254–65
42. Flint J. 2003. Analysis of quantitative trait loci that influence animal behavior. *J. Neurobiol.* 54:46–77
43. Gallego Romero I, Basu Mallick C, Liebert A, Crivellaro F, Chaubey G, et al. 2011. Herders of Indian and European cattle share their predominant allele for lactase persistence. *Mol. Biol. Evol.* 29:249–60
44. Gilad Y, Bustamante CD, Lancet D, Pääbo S. 2003. Natural selection on the olfactory receptor gene family in humans and chimpanzees. *Am. J. Hum. Genet.* 73:489–501
45. Gracheva EO, Cordero-Morales JF, González-Carcacía JA, Ingolia NT, Manno C, et al. 2011. Ganglion-specific splicing of TRPV1 underlies infrared sensation in vampire bats. *Nature* 476:88–92
46. Gracheva EO, Ingolia NT, Kelly YM, Cordero-Morales JF, Hollopeter G, et al. 2010. Molecular basis of infrared detection by snakes. *Nature* 464:1006–11
47. Graham AM, McCracken KG. 2019. Convergent evolution on the hypoxia-inducible factor (HIF) pathway genes *EGLN1* and *EPAS1* in high-altitude ducks. *Heredity* 122:819–32
48. Gray JM, Karow DS, Lu H, Chang AJ, Chang JS, et al. 2004. Oxygen sensation and social feeding mediated by a *C. elegans* guanylate cyclase homologue. *Nature* 430:317–22
49. Greene JS, Brown M, Dobosiewicz M, Ishida IG, Macosko EZ, et al. 2016. Balancing selection shapes density-dependent foraging behaviour. *Nature* 539:254–58
50. Greene JS, Dobosiewicz M, Butcher RA, McGrath PT, Bargmann CI. 2016. Regulatory changes in two chemoreceptor genes contribute to a *Caenorhabditis elegans* QTL for foraging behavior. *eLife* 5:e21454
51. Greenwood AK, Mills MG, Wark AR, Archambeault SL, Peichel CL. 2016. Evolution of schooling behavior in threespine sticklebacks is shaped by the *Eda* gene. *Genetics* 203:677–81
52. Greenwood AK, Wark AR, Yoshida K, Peichel CL. 2013. Genetic and neural modularity underlie the evolution of schooling behavior in threespine sticklebacks. *Curr. Biol.* 23:1884–88
53. Hernandez RD, Uricchio LH, Hartman K, Ye C, Dahl A, Zaitlen N. 2019. Ultrarare variants drive substantial *cis* heritability of human gene expression. *Nat. Genet.* 51:1349–55
54. Herrmann M, Mayer WE, Sommer RJ. 2006. Nematodes of the genus *Pristionchus* are closely associated with scarab beetles and the Colorado potato beetle in western Europe. *Zoology* 109:96–108
55. Heyne HO, Lautenschläger S, Nelson R, Besnier F, Rotival M, et al. 2014. Genetic influences on brain gene expression in rats selected for tameness and aggression. *Genetics* 198:1277–90
56. Higham T, Douka K, Wood R, Ramsey CB, Brock F, et al. 2014. The timing and spatiotemporal patterning of Neanderthal disappearance. *Nature* 512:306–9
57. Hodgkin J, Doniach T. 1997. Natural variation and copulatory plug formation in *Caenorhabditis elegans*. *Genetics* 146:149–64
58. Hong RL, Witte H, Sommer RJ. 2008. Natural variation in *Pristionchus pacificus* insect pheromone attraction involves the protein kinase EGL-4. *PNAS* 105:7779–84
59. Horton BM, Hudson WH, Ortlund EA, Shirk S, Thomas JW, et al. 2014. Estrogen receptor α polymorphism in a species with alternative behavioral phenotypes. *PNAS* 111:1443–48
60. Hu Y, Wu Q, Ma S, Ma T, Shan L, et al. 2017. Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. *PNAS* 114:1081–86
61. Huerta-Sánchez E, Jin X, Asan Bianba Z, Peter BM, et al. 2014. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* 512:194–97

62. Hughes GM, Boston ESM, Finarelli JA, Murphy WJ, Higgins DG, Teeling EC. 2018. The birth and death of olfactory receptor gene families in mammalian niche adaptation. *Mol. Biol. Evol.* 35:1390–406
63. Ilardo MA, Moltke I, Korneliussen TS, Cheng J, Stern AJ, et al. 2018. Physiological and genetic adaptations to diving in Sea Nomads. *Cell* 173:569–80
64. Insel TR. 2010. The challenge of translation in social neuroscience: a review of oxytocin, vasopressin, and affiliative behavior. *Neuron* 65:768–79
65. Ishikawa Y, Takanashi T, Kim C, Hoshizaki S, Tatsuki S, Huang Y. 1999. *Ostrinia* spp. in Japan: their host plants and sex pheromones. *Entomol. Exp. Appl.* 91:237–44
66. Itoigawa A, Hayakawa T, Suzuki-Hashido N, Imai H. 2019. A natural point mutation in the bitter taste receptor TAS2R16 causes inverse agonism of arbutin in lemur gustation. *Proc. Biol. Sci.* 296:2861904
67. Jiang P, Josue J, Li X, Glaser D, Li W, et al. 2012. Major taste loss in carnivorous mammals. *PNAS* 109:4956–61
68. Kamilar JM, Cooper N. 2013. Phylogenetic signal in primate behaviour, ecology and life history. *Philos. Trans. R. Soc. B* 368:20120341
69. Karageorgi M, Groen SC, Sumbul F, Pelaez JN, Verster KI, et al. 2019. Genome editing retraces the evolution of toxin resistance in the monarch butterfly. *Nature* 574:409–12
70. Karlsson Linnér R, Biroli P, Kong E, Meddens SFW, Wedow R, et al. 2019. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nat. Genet.* 51:245–57
71. Kawamura S. 2016. Color vision diversity and significance in primates inferred from genetic and field studies. *Genes Genom*. 38:779–91
72. Keller L, Parker JD. 2002. Behavioral genetics: a gene for supersociality. *Curr. Biol.* 12:180–81
73. Khan I, Yang Z, Maldonado E, Li C, Zhang G, et al. 2015. Olfactory receptor subgenomes linked with broad ecological adaptations in Sauropsida. *Mol. Biol. Evol.* 32:2832–43
74. Kimchi T, Xu J, Dulac C. 2007. A functional circuit underlying male sexual behaviour in the female mouse brain. *Nature* 448:1009–14
75. Kiontke K, Fitch DHA. 2005. The phylogenetic relationships of *Caenorhabditis* and other rhabditids. In *WormBook*, ed. *C. elegans* Res. Commun. **https://doi.org/10.1895/wormbook.1.11.1**
76. Kishida T, Kubota S, Shirayama Y, Fukami H. 2007. The olfactory receptor gene repertoires in secondary-adapted marine vertebrates: evidence for reduction of the functional proportions in cetaceans. *Biol. Lett.* 3:428–30
77. Kocher SD, Mallarino R, Rubin BER, Yu DW, Hoekstra HE, Pierce NE. 2018. The genetic basis of a social polymorphism in halictid bees. *Nat. Commun.* 9:4338
78. Kong A, Thorleifsson G, Frigge ML, Vilhjalmsson BJ, Young AI, et al. 2018. The nature of nurture: effects of parental genotypes. *Science* 359:424–28
79. Kukekova AV, Johnson JL, Xiang X, Feng S, Liu S, et al. 2018. Red fox genome assembly identifies genomic regions associated with tame and aggressive behaviours. *Nat. Ecol. Evol.* 2:1479–91
80. Küpper C, Stocks M, Risse JE, Dos Remedios N, Farrell LL, et al. 2015. A supergene determines highly divergent male reproductive morphs in the ruff. *Nat. Genet.* 48:79–83
81. Lamichhaney S, Fan G, Widemo F, Gunnarsson U, Thalmann DS, et al. 2015. Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). *Nat. Genet.* 48:84–88
82. Lander ES, Schork NJ. 1994. Genetic dissection of complex traits. *Science* 265:2037–48
83. Lappalainen T, Scott AJ, Brandt M, Hall IM. 2019. Genomic analysis in the age of human genome sequencing. *Cell* 177:70–84
84. Lassance JM, Groot AT, Liénard MA, Antony B, Borgwardt C, et al. 2010. Allelic variation in a fatty-acyl reductase gene causes divergence in moth sex pheromones. *Nature* 466:486–89
85. Leary GP, Allen JE, Bunger PL, Luginbill JB, Linn CE, et al. 2012. Single mutation to a sex pheromone receptor provides adaptive specificity between closely related moth species. *PNAS* 109:14081–86
86. Lee D, Zdraljevic S, Cook DE, Frézal L, Hsu J, et al. 2019. Selection and gene flow shape niche-associated variation in pheromone response. *Nat. Ecol. Evol.* 3:1455–63

87. L'Etoile ND, Coburn CM, Eastham J, Kistler A, Gallegos G, Bargmann CI. 2002. The cyclic GMP-dependent protein kinase EGL-4 regulates olfactory adaptation in *C. elegans*. *Neuron* 36:1079–89
88. Levene H. 1953. Genetic equilibrium when more than one ecological niche is available. *Am. Nat.* 87:331–33
89. Li R, Fan W, Tian G, Zhu H, He L, et al. 2010. The sequence and de novo assembly of the giant panda genome. *Nature* 463:311–17
90. Li X, Li W, Wang H, Cao J, Maehashi K, et al. 2005. Pseudogenization of a sweet-receptor gene accounts for cats' indifference toward sugar. *PLOS Genet*. 1:e3
91. Lin JJ, Wang FY, Li WH, Wang TY. 2017. The rises and falls of opsin genes in 59 ray-finned fish genomes and their implications for environmental adaptation. *Sci. Rep.* 7:15568
92. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, et al. 2019. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat. Genet.* 51:237–44
93. Long T, Hicks M, Yu H-C, Biggs WH, Kirkness EF, et al. 2017. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nat. Genet.* 49:568–78
94. Lonn E, Koskela E, Mappes T, Mokkonen M, Sims AM, Watts PC. 2017. Balancing selection maintains polymorphisms at neurogenetic loci in field experiments. *PNAS* 114:3690–95
95. Louis J, David JR. 1985. Ecological specialization in the *Drosophila melanogaster* species subgroup: a case study of *D. sechellia*. *Acta Oecol*. 7:215–29
96. Luczak S, Glatt SJ, Wall TJ. 2006. Meta-analyses of *ALDH2* and *ADH1B* with alcohol dependence in Asians. *Psychol. Bull.* 132:607–21
97. Lyon MF. 2003. Transmission ratio distortion in mice. *Annu. Rev. Genet.* 37:393–408
98. Macosko EZ, Pokala N, Feinberg EH, Chalasani SH, Butcher RA, et al. 2009. A hub-and-spoke circuit drives pheromone attraction and social behaviour in *C. elegans*. *Nature* 458:1171–75
99. Mandel AL, Peyrot des Gachons C, Plank KL, Alarcon S, Breslin PAS. 2010. Individual differences in *AMY1* gene copy number, salivary α-amylase levels, and the perception of oral starch. *PLOS ONE* 5:e13352
100. Mayr E. 1963. *Animal Species and Evolution*. Cambridge, MA: Harvard Univ. Press
101. McBride CS, Baier F, Omondi AB, Spitzer SA, Lutomiah J, et al. 2014. Evolution of mosquito preference for humans linked to an odorant receptor. *Nature* 515:222–27
102. McGrath PT, Rockman MV, Zimmer M, Jang H, Macosko EZ, et al. 2009. Quantitative mapping of a digenic behavioral trait implicates globin variation in *C. elegans* sensory behaviors. *Neuron* 61:692–99
103. McGrath PT, Xu Y, Ailion M, Garrison JL, Butcher RA, Bargmann CI. 2011. Parallel evolution of domesticated *Caenorhabditis* species targets pheromone receptor genes. *Nature* 477:321–25
104. Merritt JR, Grogan KE, Zinzow-Kramer WM, Sun D, Ortlund EA. 2020. A behavioral polymorphism caused by a single gene inside a supergene. bioRxiv 2020.01.13.897637. **https://doi.org/10.1101/2020.01.13.897637**
105. Muñoz MM, Losos JB. 2018. Thermoregulatory behavior simultaneously promotes and forestalls evolution in a tropical lizard. *Am. Nat.* 191:E15–26
106. Muralidhar P. 2019. Mating preferences of selfish sex chromosomes. *Nature* 570:376–79
107. Musilova Z, Cortesi F, Matschiner M, Davies WIL, Patel JS, et al. 2019. Vision using multiple distinct rod opsins in deep-sea fishes. *Science* 364:588–92
108. Nakka P, Pattillo Smith S, O'Donnell-Luria AH, McManus KF, Agee M, et al. 2019. Characterization of prevalence and health consequences of uniparental disomy in four million individuals from the general population. *Am. J. Hum. Genet.* 105:921–32
109. Nei M, Niimura Y, Nozawa M. 2008. The evolution of animal chemosensory receptor gene repertoires: roles of chance and necessity. *Nat. Rev. Genet.* 9:951–63
110. Niimura Y, Nei M. 2007. Extensive gains and losses of olfactory receptor genes in mammalian evolution. *PLOS ONE* 2:e708
111. Nikaido M, Suzuki H, Toyoda A, Fujiyama A, Hagino-Yamagishi K, et al. 2013. Lineage-specific expansion of vomeronasal type 2 receptor-like (OlfC) genes in cichlids may contribute to diversification of amino acid detection systems. *Genome Biol. Evol.* 5:711–22

112. Noble LM, Chang AS, McNelis D, Kramer M, Yen M, et al. 2015. Natural variation in *plep-1* causes male-male copulatory behavior in *C. elegans*. *Curr. Biol.* 25:2730–37
113. Okhovat M, Berrio A, Wallace G, Ophir AG, Phelps SM. 2015. Sexual fidelity trade-offs promote regulatory variation in the prairie vole brain. *Science* 350:1371–74
114. Ophir AG, Phelps SM, Sorin AB, Wolff JO. 2008. Social but not genetic monogamy is associated with greater breeding success in prairie voles. *Anim. Behav.* 75:1143–54
115. Paaby AB, Rockman M V. 2014. Cryptic genetic variation: evolution's hidden substrate. *Nat. Rev. Genet.* 15:247–58
116. Pajic P, Pavlidis P, Dean K, Neznanova L, Romano RA, et al. 2019. Independent amylase gene copy number bursts correlate with dietary preferences in mammals. *eLife* 8:e44628
117. Polderman TJC, Benyamin B, De Leeuw CA, Sullivan PF, Van Bochoven A, et al. 2015. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* 47:702–9
118. Prieto-Godino LL, Rytz R, Bargeton B, Abuin L, Arguello JR, et al. 2016. Olfactory receptor pseudo-pseudogenes. *Nature* 539:93–97
119. Prieto-Godino LL, Rytz R, Cruchet S, Bargeton B, Abuin L, et al. 2017. Evolution of acid-sensing olfactory circuits in drosophilids. *Neuron* 93:661–76
120. Ray R, Tyndale RF, Lerman C. 2009. Nicotine dependence pharmacogenetics: role of genetic variation in nicotine-metabolizing enzymes. *J. Neurogenet.* 23:252–61
121. Ripke S, Neale BM, Corvin A, Walters JTR, Farh K-H, et al. 2014. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511:421–27
122. Robinson GE, Fernald RD, Clayton DF. 2008. Genes and social behavior. *Science* 322:896–900
123. Robson SKA, Kohout RJ, Beckenbach AT, Moreau CS. 2015. Evolutionary transitions of complex labile traits: silk weaving and arboreal nesting in *Polyrhachis* ants. *Behav. Ecol. Sociobiol.* 69:449–58
124. Rockman MV, Kruglyak L. 2009. Recombinational landscape and population genomics of *Caenorhabditis elegans*. *PLOS Genet.* 5:e1000419
125. Ross KG, Keller L. 1998. Genetic control of social organization in an ant. *PNAS* 95:14232–37
126. Runge JN, Lindholm AK. 2018. Carrying a selfish genetic element predicts increased migration propensity in free-living wild house mice. *Proc. Biol. Sci.* 285:20181333
127. Schrider DR, Kern AD. 2017. Soft sweeps are the dominant mode of adaptation in the human genome. *Mol. Biol. Evol.* 34:1863–77
128. Schweizer RM, Velotta JP, Ivy CM, Jones MR, Muir SM, et al. 2019. Physiological and genomic evidence that selection on the transcription factor *Epas1* has altered cardiovascular function in high-altitude deer mice. *PLOS Genet.* 15:e1008420
129. Seeholzer LF, Seppo M, Stern DL, Ruta V. 2018. Evolution of a central neural circuit underlies *Drosophila* mate preferences. *Nature* 559:564–69
130. Shi P, Zhang J. 2006. Contrasting modes of evolution between vertebrate sweet/umami receptor genes and bitter receptor genes. *Mol. Biol. Evol.* 23:292–300
131. Silver LM. 1985. Mouse *t* haplotypes. *Annu. Rev. Genet.* 19:179–208
132. Stacey D, Fauman EB, Ziemek D, Sun BB, Harshfield EL, et al. 2019. ProGeM: a framework for the prioritization of candidate causal genes at molecular quantitative trait loci. *Nucleic Acids Res.* 47:e3
133. Staiger EA, Almén MS, Promerová M, Brooks S, Cothran EG, et al. 2017. The evolutionary history of the *DMRT3* 'Gait keeper' haplotype. *Anim. Genet.* 48:551–59
134. Stern DL. 2014. Identification of loci that cause phenotypic variation in diverse species with the reciprocal hemizygosity test. *Trends Genet.* 30:547–54
135. Stirling DG, Réale D, Roff DA. 2002. Selection, structure and the heritability of behaviour. *J. Evol. Biol.* 15:277–89
136. Sutter A, Lindholm AK. 2015. Detrimental effects of an autosomal selfish genetic element on sperm competitiveness in house mice. *Proc. R. Soc. B* 282:20150974
137. Swallow DM. 2003. Genetics of lactase persistence and lactose intolerance. *Annu. Rev. Genet.* 37:197–219
138. Taverner AM, Yang L, Barile ZJ, Lin B, Peng J, et al. 2019. Adaptive substitutions underlying cardiac glycoside insensitivity in insects exhibit epistasis in vivo. *eLife* 8:e48224

139. Thomas JW, Caceres M, Lowman JJ, Morehouse CB, Short ME, et al. 2008. The chromosomal polymorphism linked to variation in social behavior in the white-throated sparrow (*Zonotrichia albicollis*) is a complex rearrangement and suppressor of recombination. *Genetics* 179:1455–68
140. Thorgeirsson TE, Gudbjartsson DF, Surakka I, Vink JM, Amin N, et al. 2010. Sequence variants at *CHRNB3-CHRNA6* and *CYP2A6* affect smoking behavior. *Nat. Genet.* 42:448–53
141. Thorneycroft HB. 1975. A cytogenetic study of the white-throated sparrow, *Zonotrichia albicollis* (Gmelin). *Int. J. Org. Evol.* 29:611–21
142. Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, et al. 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* 39:31–40
143. Toews DPL, Taylor SA, Streby HM, Kramer GR, Lovette IJ. 2019. Selection on *VPS13A* linked to migration in a songbird. *PNAS* 116:18272–74
144. Torgerson DG, Boyko AR, Hernandez RD, Indap A, Hu X, et al. 2009. Evolutionary processes acting on candidate *cis*-regulatory regions in humans inferred from patterns of polymorphism and divergence. *PLOS Genet*. 5:e1000592
145. Wada-Katsumata A, Silverman J, Schal C. 2013. Changes in taste neurons support the emergence of an adaptive behavior in cockroaches. *Science* 340:972–75
146. Wallberg A, Schoning C, Webster MT, Hasselmann M. 2017. Two extended haplotype blocks are associated with adaptation to high altitude habitats in East African honey bees. *PLOS Genet*. 13:e1006792
147. Walters RK, Polimanti R, Johnson EC, McClintick JN, Adams MJ, et al. 2018. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat. Neurosci.* 21:1656–69
148. Wang J, Wurm Y, Nipitwattanaphon M, Riba-Grognuz O, Huang YC, et al. 2013. A Y-like social chromosome causes alternative colony organization in fire ants. *Nature* 493:664–68
149. Weber JN, Peterson BK, Hoekstra HE. 2013. Discrete genetic modules are responsible for complex burrow evolution in *Peromyscus* mice. *Nature* 493:402–5
150. Won H, de la Torre-Ubieta L, Stein JL, Parikshak NN, Huang J, et al. 2016. Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature* 538:523–27
151. Wooding S, Bufe B, Grassi C, Howard MT, Stone AC, et al. 2006. Independent evolution of bitter-taste sensitivity in humans and chimpanzees. *Nature* 440:930–34
152. Wu Z, Autry AE, Bergan JF, Watabe-Uchida M, Dulac CG. 2014. Galanin neurons in the medial preoptic area govern parental behaviour. *Nature* 509:325–30
153. Yassin A, Debat V, Bastide H, Gidaszewski N, David JR, Pool JE. 2016. Recurrent specialization on a toxic fruit in an island *Drosophila* population. *PNAS* 113:4771–76
154. Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZXP, et al. 2010. Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* 329:75–78
155. York RA, Patil C, Abdilleh K, Johnson ZV, Conte MA, et al. 2018. Behavior-dependent *cis* regulation reveals genes and pathways associated with bower building in cichlid fishes. *PNAS* 115:E11081–90
156. Young AI, Benonisdottir S, Przeworski M, Kong A. 2019. Deconstructing the sources of genotype-phenotype associations in humans. *Science* 365:1396–400
157. Young AI, Frigge ML, Gudbjartsson DF, Thorleifsson G, Bjornsdottir G, et al. 2018. Relatedness disequilibrium regression estimates heritability without environmental bias. *Nat. Genet.* 50:1304–10
158. Young JM, Massa HF, Hsu L, Trask BJ. 2010. Extreme variability among mammalian V1R gene families. *Genome Res.* 20:10–18
159. Young JM, Trask BJ. 2007. V2R gene families degenerated in primates, dog and cow, but expanded in opossum. *Trends Genet*. 23:212–15
160. Young LJ, Wang Z, Insel TR. 1998. Neuroendocrine bases of monogamy. *Trends Neurosci*. 21:71–75
161. Zhang J, Webb DM. 2003. Evolutionary deterioration of the vomeronasal pheromone transduction pathway in catarrhine primates. *PNAS* 100:8337–41
162. Zhao H, Li J, Zhang J. 2015. Molecular evidence for the loss of three basic tastes in penguins. *Curr. Biol.* 25:R141–42
163. Zhao H, Xu D, Zhang S, Zhang J. 2012. Genomic and genetic evidence for the loss of umami taste in bats. *Genome Biol. Evol.* 4:73–79

164. Zhao H, Zhou Y, Pinto CM, Charles-Dominique P, Galindo-González J, et al. 2010. Evolution of the sweet taste receptor gene *Tas1r2* in bats. *Mol. Biol. Evol.* 27:2642–50
165. Zhen Y, Aardema ML, Medina EM, Schumer M, Andolfatto P. 2012. Parallel molecular evolution in an herbivore community. *Science* 337:1634–37
166. Zinzow-Kramer WM, Horton BM, McKee CD, Michaud JM, Tharp GK, et al. 2015. Genes located in a chromosomal inversion are correlated with territorial song in white-throated sparrows. *Genes Brain Behav*. 14:641–54