Genes and Genomic Searches

C. P. Kyriacou and E. Tauber, University of Leicester, Leicester, UK

Introduction

Behavioral genetics can be argued to be the oldest branch of genetics and can loosely trace its ancestry back to Francis Galton's book 'Hereditary Genius,' which was first published in 1869. In this epic study, Galton (1822–1911) examined the male relatives of highly distinguished Victorian men and observed that the larger the genetic distance between family members, the lower the frequency of outstanding mental abilities. Galton's work became central to the 'eugenics' movement of the first half of the twentieth century, which was later so tragically perverted by the Nazis. Nor did the American psychiatric establishment distinguish itself in this regard, with thousands of people sterilized and institutionalized, often on the flimsiest 'evidence' of genetic mental 'inferiority.' From this extreme genetic determinism of the 1920s sprang behaviorism, the extreme environmentalism of the psychologist John Broadus Watson (1878–1958) that carried almost everything before it, but with the odd exception, notably the studies on the genetic basis of learning in rodents from the laboratories of Edward Chase Tolman (1886–1959) and Robert Chaote Tryon (1901–1967). Perhaps it is from this period that behavioral genetics, as an experimental discipline, was finally born. However, it was not until the dust of the Second World War settled that a handful of zoologists and psychologists began serious work on the genetic basis of animal behavior. A subgroup of these, the ethologists Konrad Lorenz (1903–1989), Niko Tinbergen (1907–1988), and Karl von Frisch (1886–1982) studied instinctive species-specific behavior in vertebrates and insects, with the implication that these motor programs had an underlying genetic basis. They were to share a Nobel Prize for Medicine and Physiology in 1973. Yet it was many years before the first 'behavioral' gene was identified at the molecular level (see Hay's (1985) textbook for more on the history of this subject).

The behavioral geneticists of the 1950s, using inbred lines or selection experiments, studied the genetic architecture underlying behavioral phenotypes such as mating or open field activity, by making a series of genetic crosses, usually in mice, rats, or flies. They could even map differences in behavior between strains of flies to specific chromosomes. Studies, such as those of Fulker, used the methods of quantitative genetics to provide some information about the evolutionary history of the behavioral trait in question. However, it was not until this kind of formal genetic analysis was blended with molecular biology in the 1990s that progress was made in identifying individual genes that contributed, at least partially, to complex behavioral phenotypes. Thus, in the 1960s, the best one could do if one wished to study single gene effects on behavior was to take a morphological mutant in the fly such as *ebony* or *yellow*, *or* a neurological mutant mouse such as *waltzer* or *twirler*, and study various behavioral phenotypes in the hope that something interesting might emerge. Sometimes it did and sometimes not.

The Birth of Neurogenetics – Genetic Screens for Behavioral Phenotypes

In the mid- 1960s, Seymour Benzer (1921–2007) suggested a novel 'bottom up' approach whereby a single mutation was made randomly within the genome of a model organism, and then behavior was screened for interesting phenotypes. His organism of choice was the fruitfly. Not only did it have a life cycle of only 10 days, making genetic analysis relatively rapid (compared to several months in mice), but the genetic map of the fly was already well understood, and the behavior of a fly seemed genuinely interesting. Benzer's idea was to feed the flies a powerful mutagen, and then screen for behavioral mutants using various fly-specific genetic tricks. The underlying mutation would alter one nucleotide base pair, and if that altered a codon, the amino acid change might generate a phenotypic difference. Using simple yet ingenious behavioral screens involving flight, movement, vision, courtship, etc., Benzer's students soon identified many mutants that would do strange things, not fly, not mate, not see, shake violently, and they were given colorful names like *drop-dead, coitus interruptus, ether-a-go-go*, etc. These behavioral mutations could be mapped to the genome, thereby identifying the corresponding gene as a position on a chromosome. In addition, 'fate mapping,' a technique that Benzer extended to behavior, allowed an approximate identification of the likely neuronal (or otherwise) tissues in which the mutated behavioral gene was having its primary effect. The molecular analysis of these genes came much later, with the advent of cloning and germline transformation techniques in the 1980s.

Benzer's students soon progressed to identifying mutations in more complex and interesting behaviors, for example in learning and memory and circadian rhythms. Flies learn to associate specific odors with electric shocks,

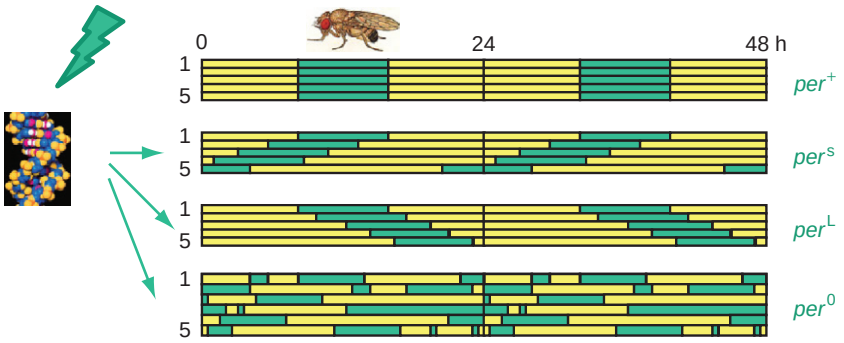

Figure 1 Identification of the *period* gene in *Drosophila melanogaster.* Chemical mutagenesis of DNA resulted in three single nucleotide changes within the *per* gene giving rise to short, long period and arrhythmic animals. The locomotor activity of a fly is double-plotted on the horizontal axis, for 5 days (vertical axis). Wild-type per^{+} flies, are active (blue) or sleep (yellow) in 24 h rhythms, so they start activity and end it at the same time each day. The short, per^{S} mutant, has fast 19 h rhythms so activity begins and ends 5 h earlier on every successive day (the 'actogram' moves to the left). The per^{L} mutant has long 29 h cycles, so the activity trace moves to the right, while the per^{0} mutant is arrhythmic. Adapted from Konopka RJ and Benzer S (1971) Clock mutants of *Drosophila melanogaster. Proceedings of the National Academy of Sciences of the United States of America* 68: 2112–2116, with permission from · · ·.

and avoid these in future. Mutations in genes *dunce, amnesiac, rutabaga, turnip, zucchini* fail these memory tests. One of Benzer's students, Ron Konopka, developed a method to measure circadian (24 h) locomotor activity rhythms in flies, and subsequent mutagenesis identified three alternative mutant alleles of a single gene termed *period (per)*, which produced short- or long-period, or arrhythmic behavioral cycles (**Figure 1**). Sometime later, it was discovered that *per* could be deleted entirely from the genome, yet the fly appeared happy, healthy, but arrhythmic – in other words, *per* was not a vital gene – it was as true a behavioral gene as any could be. Thus, the identification of *per* comes at the birth of the field known as neurogenetics, which has flourished ever since. Indeed, the fly story mentioned here has been significantly enhanced by similar studies of phenotype-driven 'forward genetics' in mice (for further reading on flies, see Nitabach and Taghert (2008)). In 1994, Joe Takahashi and his group used chemical mutagenesis to identify a variant that disrupted circadian locomotor behavior. They called this mutant mouse *Clock* (*Clk*) and it provided the entree into the molecular basis of the vertebrate circadian mechanism, which incidentally turns out to be highly conserved between flies and mice.

Transposon Mutagenesis

Chemical mutagenesis usually changes one base pair at a time, but mutagenesis can also be accomplished by hopping a mobile piece of DNA (a transposable element, TE) into another gene and disrupting it to cause a behavioral phenotype. Many behavioral mutants in flies are caused by such TEs, including some clock mutants. These types of approaches have one neat advantage over chemical methods in that they can be used as molecular tags to clone the surrounding areas (the behavioral gene into which they have hopped). These flanking DNA sequences surrounding the transposon can be identified and entered into the fly genome database to find the disrupted fly gene (http://flybase.org/). One disadvantage however is that transposons, unlike chemical mutagens, do not interrogate the genome randomly, but tend to prefer certain sequence compositions for their insertion. On the other hand, chemical mutagenesis has the disadvantage that time-consuming genetic mapping followed by positional cloning is usually the only way to identify the molecular lesion, as in the case of mouse *Clk.*

RNAi – RNA Interference, a Revolution in Genome Screening

A few years ago, small double-stranded RNA molecules (dsRNA) were discovered by Fire and Mello, which was to earn them a Nobel Prize in 2006. These molecules have the ability to interfere with the translation of any mRNA that has a similar sequence and provides a means for 'knocking down' gene expression. To downregulate a particular mRNA, a double-stranded RNA molecule corresponding to the gene must be made and introduced into the organism. For example, a short sequence from *per* could be used to make an inverted repeat of that sequence that will allow the two complementary sequences to base pair and form a dsRNA molecule. This can then be transformed into the fly in a way that will target it to cells that express *per.* These short molecules will then pair with the endogenous *per* mRNA and block translation. A number of centers around the world have generated dsRNA molecules providing RNAi for every gene in the fly genome, all 14 000. One can order a fly strain that carries a dsRNA of interest, and then by crossing this line to another strain

that carries an activator of this dsRNA, fused to a sequence that targets the activator to a tissue of choice, your favorite gene can be knocked down tissue-specifically. Systematic screening of all RNAi lines for behavioral phenotypes is usually too laborious, unless the behavioral screen can be made 'high throughput.' Instead, it is possible to use a cellular model as a behavioral readout, as was done by Amita Sehgal, who was able to screen a molecular RNAi library within cell lines in order to identify new genes that were important for entraining the cellular circadian clock to light–dark cycles.

RNAi experiments provide an example of gene product driven 'reverse genetic' approaches. Another reverse method is to knock out or eliminate a gene. In circadian biology, most of the murine clock genes that are homologous to the fly genes were identified by sequence similarity and then targeted by gene knockouts (KOs) to examine any phenotypes. Thus, a KO of the mouse homolog of fly *cyc* (called *Bmal1*) made by Bradfield and colleagues gives complete arrhythmicity, revealing the striking functional conservation of the two species genes.

What Do Gene Sequences Tell Us?

From the mid-1980s, it became possible to molecularly clone fly genes, identify their DNA sequences, and translate them into their putative proteins *in silico* using the genetic code. When the *per* gene was first sequenced in the mid-1980s, it looked like nothing else in the databases – it encoded a 'pioneer protein.' Over the years, a number of other proteins were identified in various organisms that shared a particular sequence domain with PER called PAS. This domain was important for protein–protein interactions and was found in many proteins that were environmental sensors, and particularly responsive to light, oxygen, and voltage. This makes a certain sense as PER must have evolved in response to environmental light–dark cycles. This PAS domain of PER was used in a reverse genetic approach as a trap to identify a protein partner of PER called TIMELESS (TIM). At about the same time, a forward genetics mutagenesis produced a *tim* mutant which was arrhythmic. It turns out that PER and TIM are partner molecules in the fly clock mechanism. They are transcribed into mRNA early at night in clock cells and then translated into proteins in the cytoplasm during the night (**Figure 2**). Late at night they dimerise via the PAS domain of PER and move into the nucleus. There, they (PER–TIM) interact with the transcription factor CLK (see above – it is found in the fly as well as the mouse) and negatively regulate their own genes by sequestering CLK and its partner CYCLE (CYC, also initially defined by mutagenesis, both CYC and CLK have PAS domains). Later on, around dawn, PER and TIM degrade, releasing their block on CLK and CYC, which are now free to move back onto *per* and *tim* and reactivate transcription (**Figure 2**). This relentless molecular cycle of *per* and *tim* mRNA and their proteins thus requires the two negative factors PER and TIM, and the two positive factors CLK and CYC, within the negative feedback loop that underlies the circadian mechanism, both in flies, and with some minor modifications, in mammals.

Gene sequences can be translated *in silico* into a protein sequence, which can then be compared with thousands of other sequences of known function in a protein database. For example, if the protein has a kinase domain, it will phosphorylate another protein, possibly leading to changes in its stability. If it is a transcription factor, it will be turning on or off other downstream genes. If it is a signaling protein, it will be involved in a transduction cascade, and so on. This information is crucial for understanding the underlying functional biology of the behavioral phenotype and informs and guides future experimentation.

Cellular Biology of Behavior

We linger on biological rhythms as they provide the best example we have of forward genetics being used to identify clock components. However, once a gene is identified, so is the protein, and using reagents such as antibodies or hybridization probes for the endogenous mRNA, a precise determination of exactly which tissues express the gene and protein, and when, can be made. This opens up the cellular as well as the molecular biology of the behavioral phenotype, and needless to say, in circadian rhythms, or learning, and courtship in flies, these approaches have been refined to an art form. Almost any gene can be expressed or misexpressed in almost any tissue of the fly, and this permits a panoramic exploration of the biology of behavior. So, for example, the critical clock neurons in the fly have been identified, and misexpressing clock genes or apoptotic genes (that cause cell death), within subsets of these neurons has revealed separate oscillators that control 'morning' and 'evening' behavior. In courtship, misexpression of a male-specific splice form of the gene *fruitless* (*fru*), in different neurons within the antennal regions, can convert a phenotypic female into a 'she-male,' who will inappropriately court other females. Careful examination of these regions of the central and peripheral nervous system by Billeter and colleagues reveal sex-specific anatomical differences in the shapes and the numbers of some of these *fru*- expressing neurons.

Neurogenetic Disease Models

It is also possible to subvert the fly and use it to study behavior indirectly. For example, Huntington's disease

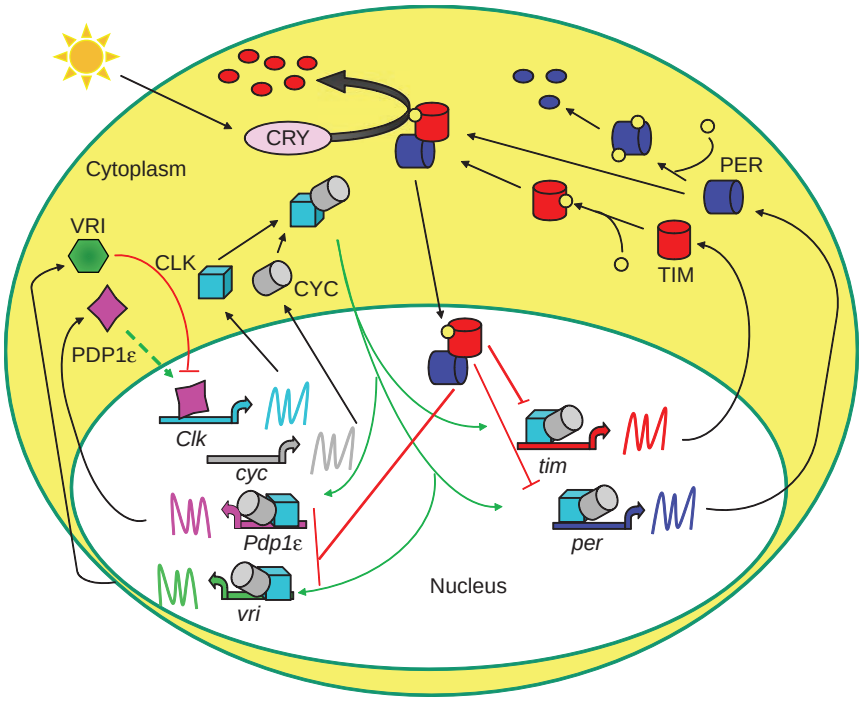

Figure 2 Forward genetics defines the molecular basis for the fly's intracellular circadian clock. The genes (italics) and corresponding proteins (Roman) are color coded. The *period* and *timeless* genes are activated by the CLK and CYC transcription factors (green arrows). The mRNAs (shown as single stranded squiggles) are exported to the cytoplasm where PER and TIM are translated. PER is phosphorylated (small yellow circles) by DBT kinase (encoded by the *doubletime* gene, not shown), which earmarks it for degradation (small blue circles). TIM is also phosphorylated (small yellow circles) by the kinase encoded by the *shaggy* (*sgg*) gene (not shown). Late at night TIM prevents DBT from phosphorylating PER so PER levels build up, and TIM–PER enter the nucleus and negatively regulate the CLK–CYC dimer (red lines), thereby repressing *per/tim* transcription as well. Thus *per* and *tim* mRNAs and proteins cycle in abundance during the circadian cycle. The *Clk* gene is itself positively (green arrows) and negatively (red lines) regulated by VRI and possible PDP1ε (dotted green arrow) leading to cycles in *Clk* and CLK abundance. Thus the *Clk* and *per/tim* feedback loops are interconnected, leading to additional stability. Both *vri* and *Pdp1ε* genes are also positively regulated by CLK–CYC (green arrows) and negatively by PER–TIM (red arrows). The blue-light photoreceptor Cryptochrome (CRY) is activated by light at dawn, and physically interacts with TIM, causing its degradation (small red circles). PER is thus exposed to DBT and degraded, thereby releasing the repression on the *per* and *tim* genes (this also occurs in constant darkness via another molecular route not involving CRY). The CLK–CYC dimer can now restart the molecular cycle by activating *per* and *tim* transcription. The roles of all these genes in the circadian clock were initially identified by forward genetics (i.e., mutagenesis) except for *Pdp1ε*, which was identified initially as a cycling transcript in fly heads. The *vri* and *sgg* genes were identified via a clever transposon mutagenesis whereby a specially constructed TE landed close to each gene, and was activated to overexpress the adjacent *vri* or *sgg* mRNA in clock neurons, revealing disruptive effects on circadian behavior.

(HD) is caused by an expansion of a polyglutamine tract (polyQ) within the huntingtin protein that is toxic to the human nervous system and causes devastating neurobehavioral impairments. When this expanded mammalian polyQ region is expressed in the eye of a fly, the eye degenerates, providing a cellular model for HD. Benzer screened 7000 TE lines and found several that could suppress the mutant Huntington's eye phenotype. Two of these lines had TEs inserted into genes that encoded chaperone domains, which are found in proteins that can prevent the misfolding of proteins that are under stress, be it mutational or environmental. Thus, the fly eye can be used as a substitute for more laborious behavioral screening and implicate gene products that might be used in future therapeutic interventions. Indeed, the fly has provided a surprisingly good model for dissecting neurodegenerative disorders, and not just those related to expanded polyQ repeats (there are nine polyQ diseases known in humans). Alzheimer's, Parkinson's, Fragile-X, and Angelman's syndrome are just some of the other neurogenetic diseases that are being studied with the fly.

Mammalian Screens

Obviously, systematic genomic searches for behavioral genes are time consuming and expensive, and hence the prominence of *Drosophila* as the major model system. Nevertheless, large-scale mouse screens for many behavioral phenotypes such as learning and memory, circadian rhythms, psychostimulant responses, vision, and stress

responses have been underway for some time. In addition, many mouse genes have now been KO'd and can be directly screened for behavioral defects. This would seem the perfect way to look for behavioral genes in mammals, but there is an associated problem. In mammals, many genes have paralogs, that is, copies of themselves somewhere else in the murine genome that duplicated from the ancestral gene millions of years ago. As evolutionary time goes by, paralogs will take on overlapping but related functions. In fact, a mouse *Clk* KO gave very subtle effects on the circadian phenotype, compared to the original *Clk* mutation, which, when homozygous, was dramatically arrhythmic. The *Clk* KO phenotype was compensated by a paralog, whereas the chemically induced *Clk* mutation was a dominant gain-of-function allele that basically gummed up the clockworks. It is interesting to speculate that if the *Clk* KO had been the only means to screen interesting circadian genes in the mouse, the *Clk* gene would probably have remained undiscovered in this context. This kind of result in which the KO mutant is not as dramatic as a chemically induced mutant may turn out to be quite widespread in mammals. In flies, most genes are single copy, so this problem of compensating paralogs in fly gene KOs does not usually rear its ugly head.

Transcriptomics

There are other ways of screening genomes for behavioral genes, and all are based on reverse genetics approaches. Transcriptomics is a popular method for detecting change in mRNA levels that correlate with altered behavior. For example, Dierick and Greenspan selected for highly aggressive male flies over a number of generations from a base population. They then isolated the head mRNA from the aggressive and the control males, and after copying it into cRNA, hybridized it to a commercial gene chip or microarray. On this microarray were placed the DNA corresponding to the entire fly transcriptome (~13 500 sequences, **Figure 3**). If any one of these sequences (arrayed as DNA spots) gave a higher or lower intensity hybridization signal in the aggressive compared to the control flies, it would suggest that the mRNA for that particular gene was up or downregulated. About 80 genes were differentially expressed in the aggressive flies, one of them, *Cyp6a20a*, encoded a cytochrome P450. To validate the microarray results, a mutant strain for this gene was obtained and was found to be significantly more aggressive, consistent with the microarray observation that the selected aggressive flies were downregulated in this particular mRNA species (**Figure 3**). Thus, a transcriptomic screen had identified a gene for aggression, which was subsequently found to be expressed in nonneuronal cells that are associated with pheromone receptors, indicating that olfaction plays a prominent role in these agonistic encounters.

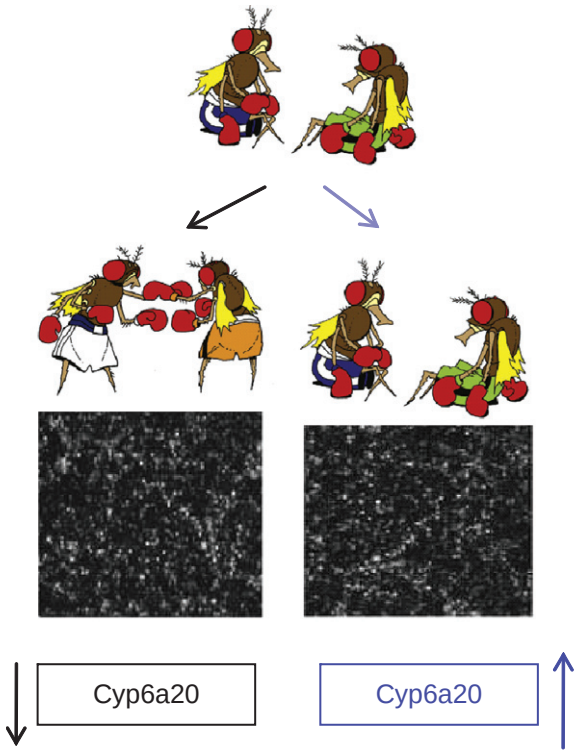

Figure 3 Transcriptomic screen for aggression genes in *Drosophila*. A base population was selected for highly aggressive flies (black arrow) or simply maintained as neutral flies (blue arrow). Microarrays were independently interrogated with mRNA from the heads of aggressive and neutral flies, and a number of genes were differentially regulated (seen as dark or light spots, each spot corresponding to a particular gene sequence). From these candidate aggression genes, one, Cyp6a20, is downregulated in aggressive flies and a mutation in this gene which reduces mRNA levels, gives increased aggression (loosely based on Dierick and Greenspan (2006) *Nature Genetics* 38: 1023-1031; cartoon of fighting flies reprinted with permission from Dierick H (2008) *Curr Biol* 18: R161–163.).

Similar transcriptomic analyses have been used to identify ~150 genes whose mRNAs cycle in abundance with a circadian period in the fly's head, or several hundred similarly cycling genes from the suprachiasmatic nuclei of the mouse, the organ that determines murine behavioral rhythms. Unlike the example of aggression, it is not differences in behavior that are being assayed here, but a molecular phenotype that has behavioral implications.

These types of studies require considerable statistical aplomb in order to separate false hits from real ones, and validation of candidate mRNAs is required, either by independent molecular methods or with the use of mutants, as in the fly aggression example cited earlier. However, as an entrée into the molecular basis of a behavioral phenotype, transcriptomics have the added

flexibility that even nonmodel organisms can be studied. What is required in these cases is the generation of the microarray (gene chip) carrying thousands of cDNA spots, each one corresponding to a different gene made from the RNA of the relevant organism (see **Figure 3**). This is followed by the interrogation of the chip with the RNA from the individuals that show differences in the phenotype, be it behavioral or molecular. Any positive hits on the slide can then be sequenced to identify the corresponding differentially expressed gene.

Applying Molecular Genetics to Identify Natural Genetic Variation

Generating mutants by forward genetics approaches is rather like hitting the animal on the head. The screen usually involves a drastic change in the phenotype for the new mutant to be noticed. However, the gene sequences that are identified by mutagenesis can then become the focus for studies of natural genetic variation. Thus, a natural polymorphism in the *tim* gene of *D. melanogaster* was shown to be spreading from southern Europe into northern Europe, under directional selection. The new mutation had originally occurred a few thousand years ago in a single fly in southern Italy and the new mutation had spread slowly northwards. This new *tim* allele provided the fly with a more adaptive behavioral response to the seasonal environments experienced in Europe, compared to the ancestral *tim* allele, which had evolved in sub-Saharan Africa, in which there is much less of a seasonal challenge. Without the *tim* sequence (identified by forward genetics and mutagenesis), there could have been no reverse genetics whereby the natural polymorphism in *tim* could be placed within a functional and evolutionary context.

Natural genetic variation can also be used to dispense with mutagenesis completely, and provide a gentler approach for searching for behavioral genes. The exponential increase in the available DNA sequence data and the identification of specific sequence regions ('markers') make this a tractable proposition. Quantitative Trait Loci (QTL) mapping is a natural continuation of the types of studies that biometrical behavioral geneticists were doing in the 1980s, before the molecular revolution really took off. This method can commence with two inbred parental lines, ideally (but not necessarily) showing a different phenotype (**Figure 4**). The two lines are crossed, and recombination in the F1 is captured in the F2 generation, which is then itself inbred for a number of generations. Each recombinant inbred line is therefore a unique mosaic of the two parental strains, and various algorithms are available to correlate the behavior of each recombinant inbred line with the genetic marker information (**Figure 4**). This method can also be directly applied to individuals from an F2 or a backgross generation without further inbreeding.

The development of molecular markers does not necessarily require a sequenced genome, so QTL mapping can be extended to nonmodel organisms. Most of the studies, however, are in model-organisms such as *Drosophila* and mouse. Not only are there stable recombinant inbred lines available for this type of work, but also the available genome sequences permit detailed mapping of QTLs and potentially may identify single loci mediating the behavior. For example, an extensive QTL study of circadian behavior in the mouse revealed 14 loci that were involved in regulating various rhythmic parameters. However, most of these QTLs did not include known circadian clock genes.

QTL mapping is limited to loci that are variable (polymorphic) between the parental lines. Genes encoding critical components for Darwinian fitness will probably be under strong directional selection, which reduces genetic variation, and so these loci may not be uncovered by QTL mapping. However, a powerful aspect of QTL mapping, which is unrivalled by the other methods we have described earlier, is the opportunity to scan simultaneously for the interaction (epistasis) between different loci: indeed, using our circadian example, a substantial amount of epistasis across the mouse genome among circadian loci was revealed.

Identifying the causative genes within a QTL is still a major challenge, as these genomic regions are large (tens to hundreds of kilobases) and typically include many genes. Finer mapping of these large regions can be extremely painful, but this process can be accelerated if some candidate genes are lurking therein. As yet, few behavioral QTL studies have revealed the underlying gene(s). In a study of emotionality in mice, a modified QTL screen using outbred stocks indicated that a regulator of G-protein signaling, *Rgs2* contributed a small proportion to the behavioral variation (~5%). Although this might seem less than overwhelming, *Rgs2* null mutants studied by Willis-Owen and Flint did show altered anxiety responses, thereby validating the QTL. Thus, QTL's may provide the candidate genes through subtle natural variation, but for validation, mutants (KOs, RNAi knockdowns or chemically or transposon induced, see above) will be required.

One potentially informative approach is to apply microarrays to the kinds of genetic crosses that we have discussed, and correlate behavior with differential gene expression in the segregating generations. The net result can be described as a gene network in which large numbers of genes interact to produce the phenotype. This kind of analysis has become popular recently, and some believe that the future of behavioral genetics may lie in understanding these networks. One possible problem may be that these networks may not be very robust, in that a

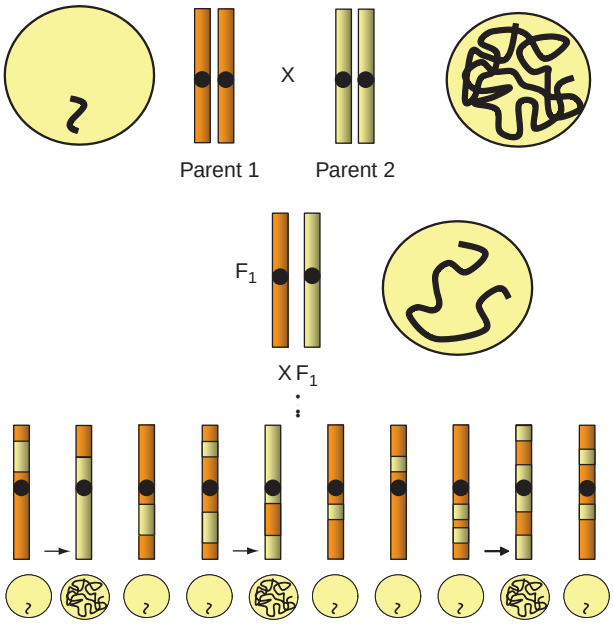

Figure 4 Mapping quantitative trait loci (QTL) involved in foraging behavior in *Drosophila* larvae (hypothetical example). Two parental strains that behave like 'sitters' (left) and do not move around very much on an agar plate, or 'rovers' (right), which do move considerably more, are crossed. The F1 progeny show intermediate behavior. The F1 are crossed for a few generations and then recombinant inbred lines (RILs) are generated by full-sib crosses. Each RIL is a mosaic of the parental genome, which can be identified by molecular markers. The behavior of each RIL is scored. The arrow indicates RILs that inherit a fragment from parent 2, and show the parental rover phenotype. This fragment is likely to carry a QTL affecting foraging behavior. Modified from Mauricio R (2001) *Nature Reviews Genetics* 2: 370–381, with apologies to Marla Sokolowski.

slight change in the conditions in which the behavior is measured ('noise') could significantly affect the overall topography of the network, recruiting new genes or losing others.

Human Molecular Neurogenetics

We cannot end without some comment on the development of neurogenetics with that most difficult of model organisms, *Homo sapiens*. The major tool that is used here is the linkage study. Briefly, if we take a family pedigree within which is segregating a behavioral phenotype of interest and consider the underlying causative behavioral mutation, the flanking genomic regions will likely contain another variant (perhaps a SNP, single nucleotide polymorphism, in a nearby gene) that always cosegregates with the behavioral mutation because the two loci are so close together that they remain undisturbed by genetic recombination. The two loci are thus in linkage and the two variants are in linkage disequilibrium, and thus the SNP in this case becomes a marker for the behavioral mutation. This is the basic principle behind linkage studies, and they have had their successes in human behavioral genetics.

A classic case involves that of a large Dutch family in which some of the boys showed unusually high levels of violence and antisocial behavior, including arson, attempted rape, and other impulsive displays such as exhibitionism. The mutation was tracked down by linkage analysis to an X-linked gene encoding monoamine oxidase A (MAOA), an enzyme that is used to break down neurotransmitters. Later studies in other families were to show that boys carrying milder mutations that produced less active versions of MAOA would not show any of these problems unless they had been subjected to abuse during childhood. These studies show beautifully how the social environment can modulate the expression of a mutant phenotype. Indeed, in the field of maternal behavior in rodents, there exists some stunning work that documents how environment can alter the heritable expression of a gene. Rat pups that receive minimal maternal care from their mothers do the same to their offspring because their gene encoding the receptor for the steroid stress response hormone, glucocorticoid, has been epigenetically modified through methylation of the DNA sequence (Fish et al., 2004). This environmentally triggered modification of the gene is passed on to the next generation, providing, superficially, a quasi-Lamarckian type of inheritance.

Another remarkable linkage study showed that a family that was segregating a dominant, autosomal disorder in which the affected individuals would wake up early and also fall asleep extremely early (Advanced Sleep Phase Syndrome or 'larks') contained a mutation in one of the

four copies of the human *Per* gene (*hPer2*). In fact, this human clock mutation was very similar to the original per^s mutation found by Konopka in his short-period fly mutant. In both the fly and the human variants, a key Serine amino acid that is phosphorylated had been replaced, and mutants of both species showed fast-running clocks. In a 24 h world, both the fast-running mutants adapted by advancing their sleep–wake cycles by several hours and becoming 'larks.'

These two spectacular and successful examples are rarities within the behavioral field, because an enormous and largely unsuccessful effort has been mounted over the past two decades in identifying some of the genes that contribute to common complex phenotypes, particularly those involving psychpathology, schizophrenia, uni- and bi-polar depression, alcoholism, etc. The net result of hundreds of such studies, many large scale and expensive, has been disappointing. A number of studies have found associations between genes such as *neuregulin*, *dysbindin*, and the gene encoding COMT (catechol-o-methyltransferase), and schizophrenia, for example, yet for every study that identifies such a candidate gene, there appear to be several others than cannot confirm this association. This has led some to question whether this kind of approach will ever be successful in isolating these loci, and some imaginative alternative hypotheses about the genetic and evolutionary basis of schizophrenia have been proposed, particularly by Tim Crow. He has suggested that epigenetic, not genetic, modifications are responsible, thereby explaining why no genetic factors have been consistently identified. This epigenetic modification is invoked to involve the *protocadherin* genes (encoding cell surface adhesion molecules) located on the X and Y chromosomes within a chromosomal rearrangement that distinguishes humans from the great apes and other primates. This rearrangement may have played a role in both the evolution of language and in its distortion in schizophrenia (hallucinations and delusions, i.e. hearing voices). Crow's ingenious epigenetic theory fits in well with the known environmental modulation of this pathology, yet a stringent experimental molecular analysis is difficult with human subjects.

Future Prospects

Neurogenetics is now a mature discipline that straddles behavior, evolution, neurobiology, and genetics. Technical developments such as RNAi have extended the field beyond the model organisms of fly, mouse, zebrafish, and nematode worm. A marine biologist, for example, might be interested in using RNAi to knock down a *per* homolog in a crab to study whether this manipulation disrupted the crustacean's 12 h tidal rhythms. Perhaps a gene originally identified within the fly that affects memory, if knocked down in a honeybee, might affects the workers ability to associate the sun compass with a food source? The technology now exists to potentially manipulate genes in organisms that have no formal genetics, so these rather more interesting eco-behavioral phenotypes will become open for neurogenetic analysis. Natural genetic variation will continue to be studied through QTL-type approaches, although many challenges still remain in dissecting out loci that contribute small yet significant components of behavioral variation, via reverse-genetics, where a gene sequence originally identified through mutagenesis then becomes the substrate for examination of natural polymorphisms. Complex behavior in humans as well as animals will have complex underlying genetic architectures, so whether the QTL or linkage and association approaches will make major contributions to dissecting out natural genetic variation remains to be seen. One prediction is that the epigenetic modification of behavioral genes that we have touched on briefly will become a major field of study in the ensuing years. Needless to say, those will be very exciting times.

See also: Honeybees; *Nasonia* Wasp Behavior Genetics.

Further Reading

Bilen J and Bonini NM (2005) *Drosophila* as a model for human neurodegenerative disease. *Annual Review of Genetics* 39: 153–171.

Billeter J-C, Rideout EJ, Dornan AJ and Goodwin SF (2006) Control of male sexual bhehaviour in *Drosophila* by the sex determination pathway. *Current Biology* 16: R766–R776.

Brunner HG, Nelen M, Breakefield XO, Ropers HH and van Oost BA (1993) Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. *Science* 262: 578–580.

Caspi A, McClay J, Moffitt TE et al. (2002) Role of genotype in the cycle of violence in maltreated children. *Science* 297: 851–854.

Crow TJ (2007) How and why genetic linkage has not solved the problem of psychosis: Review and hypothesis. *American Journal of Psychiatry* 164: 13–21.

Dierick HA and Greenspan RJ (2006) Molecular analysis of flies selected for aggressive behavior. *Nature Genetics* 38: 1023–1031.

Fish EW, Shahrokh D, Bagot R et al. (2004) Epigenetic programming of stress responses through variations in maternal care. *Annals of the New York Academy of Sciences* 1036: 167–180.

Fulker DW (1966) Mating speed in male *Drosophila melanogaster*: A psychogenetic analysis. *Science* 153: 203–205.

Hay DA (1985) *Essentials of Behaviour Genetics*. Victoria: Blackwell.

Kyriacou CP, Peixoto AA, Sandrelli F, Costa R and Tauber E (2008) Clines in clock genes: Fine-tuning circadian rhythms to the environment. *Trends in Genetics* 24: 124–132.

Nitabach MN and Taghert PH (2008) Organisation of the *Drosophila* circadian control circuit. *Current Biology* 18: R84–R93.

Panda S, Hogenesch JB and Kay SA (2002) Circadian rhythms from flies to human. *Nature* 417: 329–335.

Sathyanarayanan S, Zeng X, Kumar S et al. (2008) Identification of novel genes involved in light-dependent CRY degradation through

a genome-wide RNAi screen. *Genes & Development* 22: 1522–1533.

Shimomura K, Low-Zeddies SS, King DP et al. (2001) Genome-wide epistatic interaction analysis reveals complex genetic determinants of circadian behavior in mice. *Genome Research* 11: 959–980.

Takahashi JS, Hong H-K, Ko CH, and McDearmon EL (2008) The genetics of mammalian circadian order and disorder: Implications for physiology and disease. *Nature Reviews Genetics* 9: 764–775.

Willis-Owen SAJ and Flint J (2006) The genetic basis of emotional behaviour in mice. *European Journal of Human Genetics* 14: 721–728.