Trends in Genetics

Review

Etiology of Human Genetic Disease on the Fly

Clement Y. Chow[1] and Lawrence T. Reiter[2,3,4,*]

The model organism *Drosophila melanogaster* has been at the forefront of genetic studies since before the discovery of DNA. Although human disease modeling in flies may still be rather novel, recent advances in genetic tool design and genome sequencing now confer huge advantages in the fly system when modeling human disease. In this review, we focus on new genomic tools for human gene variant analysis; new uses for the Drosophila Genetic Reference Panel (DGRP) in detection of background alleles that influence a phenotype; and several examples of how multigenic conditions, both complex disorders and duplication and/or deletion syndromes, can be effectively studied in the fly model system. Fruit flies are a far cry from the quaint genetic model of the past, but rather, continue to evolve as a powerful system for the study of human genetic disease.

Trends

Thousands of new GAL4 lines make suppressor and/or enhancer screening and disease models more precise in flies than ever before.

New genomic tools, such as *Minos*-mediated integration cassette (MiMIC) and Trojan-GAL4, allow for quick analysis of human disease-associated variants for function and pathogenicity in flies.

Using background variants to identify genes that can influence phenotypic outcomes is a new concept in disease modeling.

Copy number variants and novel disease-associated variants in humans have been difficult to model quickly in mammals, but are relatively easy to model in Drosophila.

Drosophila melanogaster: Exciting Recent History or a Historic Model Organism?

The model organism *Drosophila melanogaster* (fruit fly) has a long and impressive history, including some of the first studies of chromosomal inheritance [1], pattern formation [2], nervous system development, and even human disease modeling [3,4]. The fly genome was one of the first to be completely sequenced [5], setting the stage for direct comparisons with human genes that cause a variety of disease phenotypes [6]. Now, with an extensive collection of genetic tools available for manipulating temporal and tissue-specific gene expression in flies, we are poised to embark on a new era of human disease modeling in Drosophila. These tools, including RNAi, open reading frame (ORF) collections for most fly genes, and next-generation transposon strategies to modify endogenous loci, are more adaptable to better model human disease compared with previously available methods. A recent review by Wangler *et al.* highlights the power of both forward and genetic screening in the Drosophila system and why researchers are reinvesting in flies to conduct research directly translatable to human genetic conditions [7].

Our intent here is not to be exhaustive in our review of the ways in which the Drosophila system can directly inform the etiology of human genetic disease, but rather to highlight new strategies to direct disease modeling and genetic modifier screening in flies. We focus on new methods for the study of human disease gene variant function, new genetic screening tools for modifiers of disease-associated phenotypes, and broad approaches to the study of copy number variants (CNVs) associated with human disease.

The Bipartite Revolution

Targeted gene expression in a temporal, tissue-specific, or even single gene-specific fashion has been a hallmark of Drosophila research since the early 1990s [8]. In the almost 25 years since the first application of the yeast GAL4 DNA-binding protein and Upstream Activator Sequence (UAS), or **GAL4/UAS (**see Glossary) system, in flies, this simple bipartite system has

[1]Department of Human Genetics, The University of Utah School of Medicine, Salt Lake City, UT 84112, USA
[2]Department of Neurology, University of Tennessee Health Science Center, Memphis, TN 38163, USA
[3]Department of Pediatrics, University of Tennessee Health Science Center, Memphis, TN 38163, USA
[4]Department of Anatomy and Neurobiology, University of Tennessee Health Science Center, Memphis, TN 38163, USA

*Correspondence: lreiter@uthsc.edu (L.T. Reiter).

Trends in Genetics, June 2017, Vol. 33, No. 6 http://dx.doi.org/10.1016/j.tig.2017.03.007
© 2017 Elsevier Ltd. All rights reserved.

been developed well beyond its original modest intentions. Early studies of human disease gene modifiers often took advantage of the *Glass Multimer Reporter* eye-specific GAL4 driver, *gmr*-GAL4. The *gmr*-GAL4 driver is expressed in a subset of eye progenitor cells of the third-instar larval eye disc. The adult flies emerge with smaller, misshapen, or 'rough' eyes, often with disordered ommatidia. This adult eye phenotype can easily be scored under a dissecting microscope. The effect of a modifier gene can be determined simply by observing a change in eye size or roughness compared with the eyes when no modifier gene is present. Suppressor and/or enhancer screens for genetic interactors in Rett syndrome, amyotrophic lateral sclerosis (ALS), Alzheimer's disease, and other single-gene disorders of the nervous system have been conducted to identify new interacting proteins contributing to pathogenesis in these diseases [9–11]. More recently, new tools have emerged allowing for more precise placement of the UAS transgene in locations that better reflect endogenous expression patterns for a particular human ortholog. In this way, the direct effects of point mutations orthologous to those that are pathogenic in humans or even transgenes expressing the human version of the gene can be analyzed for toxicity or pathway disruption. Even lethal alleles can be studied using the GAL4/UAS system combined with the GAL80 suppressor to express genes only in a defined time window. In addition, there are now thousands of cell type- and temporal-specific GAL4 drivers freely available through the various stock centers[i], with continued development of thousands of new neuron-specific drivers and gene enhancer-specific drivers [12,13]. In combination with new and improved **clustered regularly interspaced short palindromic repeats (CRISPR)/*Cas9*** methods in flies and collections of UAS-RNAi[ii] and UAS-ORF[iii] projects, we are poised to begin a new era of human disease-specific investigations in flies that better represent the actual disease state. Additional tools that make these strategies even more powerful include the ease of use of CRISPR and **PhiC31** integration in the fly genome, making it possible to directly replace the gene of interest under the endogenous promoter and/or enhancer elements [4].

Jumping Genes and Sequenced Genomes

Although genome-editing tools, such as CRISPR/*Cas9*, have been developed in Drosophila, the primary mode of transgenic manipulation remains DNA transposon insertions. Recently, Hugo Bellen and his lab at The Baylor College of Medicine developed a broad targeting *Minos* DNA transposon called the ***Minos*-mediated integration cassette** (MiMIC) for genome-wide protein trapping at the endogenous locus [14]. These vectors contain a PhiC31 recombination-mediated cassette exchange (RMCE) component that allows for the swapping of internal sequences of the inserted element [13]. The advantage of this methodology is that the internal cassette of the MiMIC element can be swapped out *in trans* for a *Trojan*-GAL4 exon [13], which simultaneously kills expression of the gene and generates a new reporter line that expresses the GAL4 protein in the pattern of the endogenous gene. This GAL4 line can be used to express a reporter such as UAS-*GFP*, or can be used to substitute a fly or human variant of the gene of interest for analysis using GAL4/UAS (Box 1).

Box 1. RMCE Technology

Figure I (Key Figure) provides an example of how to use the RMCE method for the study of human variants *in vivo*. The first step is to identify a homolog to the human gene being studied, in this case *EBF3,* which is homologous to the fly *knot* gene (Figure IA). The RMCE method exploits a highly efficient DNA transposon, called MiMIC, which, unlike previous generation P-element transposons, integrates randomly into the fly genome. The preferred MiMIC element is located in a coding region intron near the start of the gene (red arrow in Figure IB). The internal portion of the MiMIC element contains a visible marker (*yellow+*) and a gene trap cassette with an EGFP reporter [14]. The strength of this system lies in the *att*P sites flanking the internal portion of the element, which allow for exogenous *att*B-flanked DNA sequences to recombine with the inserted MiMIC element, exchanging the internal portion for any *att*B-flanked sequence desired. RMCE can be performed *in vivo* using publically available stocks via genetic crossing schemes, eliminating the need for traditional embryo injection approaches. The approach used here involves a combination of expressing GAL4 protein in the pattern of the *knot* gene while simultaneously killing the endogenous protein [13]. Using the *Trojan*-GAL4 method,

Glossary

Balancer chromosomes: specially engineered chromosomes used in Drosophila that inhibit homologous recombination; are marked with a visible dominant marker, such as curly wings or stubbly bristles; and are typically lethal when homozygous. These chromosomes allow one to 'balance' a particular mutation, which may be homozygous lethal, to fix that mutation in a particular strain of flies and follow that mutation during genetic crossing schemes.

Clustered regularly interspaced short palindromic repeats combined with the *Streptococcus pyogenes* protein Cas9 (CRISPR/Cas9): a system that involves using the location of short palindromic repeats in the genome, a guide RNA that matches a particular target of interest, and the Cas9 endonuclease that can cut the DNA at that particular locus. The system is set up so that, by using different guide RNAs, one can change the endogenous sequence of the DNA by generating a deletion or a single base-pair change.

Drosophila Genetic Reference Panel (DGRP): a collection of >200 fully sequenced, wild-derived *Drosophila melanogaster* strains. The DGRP comprises natural genetic variation used to study genetic variation and complex traits. More information and a simple association study web interface can be found online[v].

Drosophila Synthetic Population Resource (DSPR): a collection of 1700 recombinant inbred lines (RILs) of *D. melanogaster.* These RILs were derived from 15 genetically diverse founder lines. Complete genomic sequence data are available for the founder lines, which can be used in a similar manner to the DGRP. More information is available online[vi].

F1 hybrid: the offspring of a genetic cross from two distinct genetic parents.

GAL4/UAS: by combining the GAL4 *trans*-activating protein with its DNA target, the upstream activator sequence (UAS) in a single fly, we can perform the targeted expression of transgenes in the fly system in a temporal- and cell-specific manner.

Genetic variation: the natural variation that occurs between individuals in a population. This one can swap an in-frame terminator sequence tied to a GAL4 protein (*T2A*-GAL4) [13], resulting in the creation of a GAL4 driver that can be used to drive expression of a fly- or human disease-associated gene or gene variant in a null background for the gene of interest (Figure IB). The truncated knot protein contains both T2A and GAL4. GAL4 protein is released when the T2A protein self cleaves. The final step involves integrating human cDNA variants, in this case for EBF3, using a different site in the genome that contains a well-insulated *att*B site (Figure IC). Once integrated, these cDNAs for wild-type and variant versions of the *EBF3* gene can be expressed in the endogenous *knot* pattern by crossing the UAS-*cDNA* lines with the *knot-T2A-GAL4* line (Figure IC). One can now analyze the effects of known human variants for pathogenicity *in vivo* in the fly model.

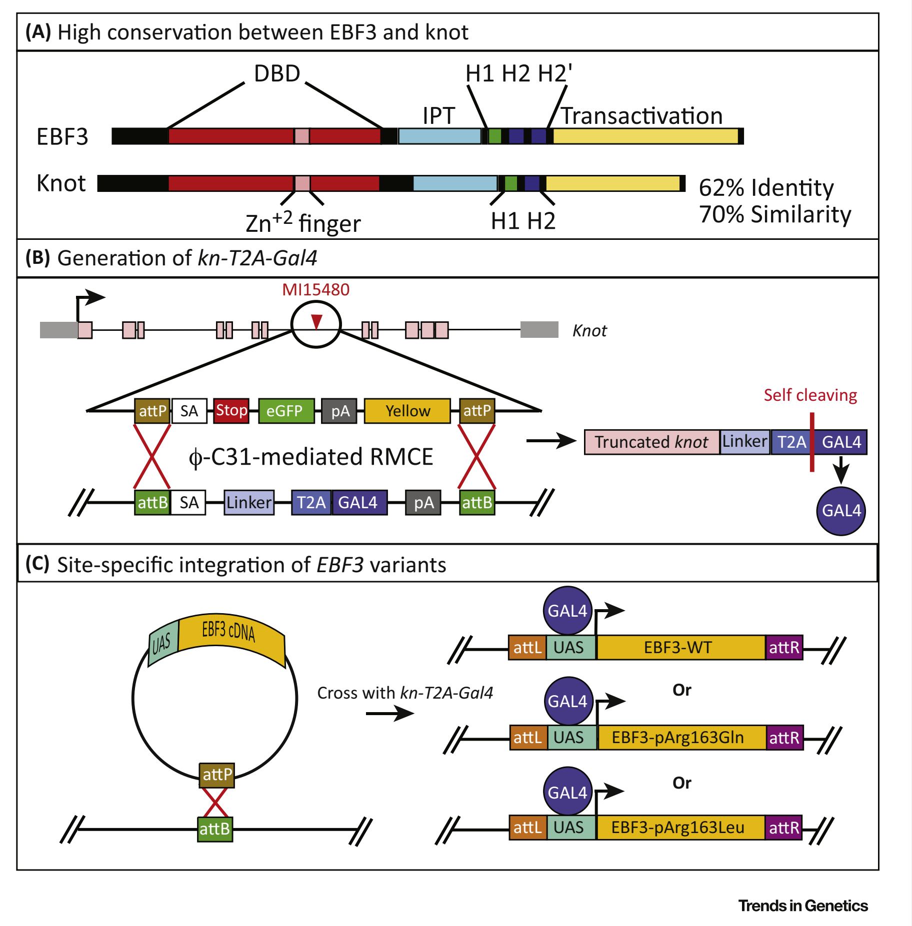

Figure I. Human Variant Analysis Using RMCE in Flies (modified from [17])

variation includes both SNPs and larger deletions or duplications that can contribute to phenotypic variation observed in different individuals in a given population.

Genome-wide association studies (GWAS): genetic studies that use the physical association on a chromosome between a given trait and a DNA marker, such as a SNP or small indel.

***Minos* mediated integration cassette (MiMIC):** a randomly integrating DNA transposon with an exchangeable *att*P-flanked internal cassette.

PhiC31 integration: borrowed from the bacteriophage *phiC31*, this technology involves using both the site-specific integrase enzyme and the specific DNA sites needed to trade segments of DNA *in vivo* that are flanked by phiC31 integration sites (typically *att*B in the transgene and *att*P at the integration site).

Sensitized background: a genetic background that already shows a phenotype or is mutated for a given gene to aid in the identification of modifiers of the phenotype.

Single nucleotide polymorphism (SNP): often used in GWAS studies as a marker of the locus for a given trait.

Z-score: the number of standard deviations from the mean for a given raw score.

Recent work has focused on the use of MiMIC insertions for the study of putatively pathogenic human variants. These variant studies have been facilitated not only by new tools for genetic manipulation, but also by bioinformatic tools that make the identification of pathogenic variants and orthologous genes in multiple model systems relatively simple for all investigators. An excellent example of one of these emerging tools is the MARRVEL database[iv]. These new approaches to the study of human gene function and pathogenicity in the fly are at the forefront of human molecular genetics. They allow for the direct analysis of gene function and variant pathogenicity for diseases where new human variants have been identified. In particular, Drosophila genetic technologies now allow us to test large-effect Mendelian disease genes and smaller-effect, rare alleles from complex human diseases indemnified by genome-wide technologies. Several recent examples illustrate the potential of this approach: (i) analysis of a pathogenic variant in *DNM1L* causing infantile encephalopathy [15]; (ii) mechanistic studies of a recurrent *de novo* variant in *ATAD3A* associated with distinct neurological phenotypes [16]; (iii) description of a new neurological syndrome resulting from *de novo* variants in *EBF3* [17]; and (iv) validation of a newly identified allele of *TM2D3* pathogenic for Alzheimer's disease [18]. These studies illustrate that Drosophila tools and analysis capabilities are catching up with the modern, postgenomic, clinical genetics needed to efficiently analyze and understand new gene variants associated with disease.

Squeezing the Genome: Exploring the Phenotypic Influence of Normal Variants

To truly understand the complexity of human disease phenotypes, we need to take into account and begin to investigate the effects of genetic variation on disease outcomes [19,20]. At the heart of this dilemma is understanding how genetic variation might be acting on a given disorder in a specific population to modify disease outcome.

To study relevant modifier genes, we need to use unbiased forward genetic screens of natural genetic variation to reveal these modifying loci. The traditional Drosophila forward genetic screen uses chemical or mutagenic approaches to generate mutations across the genome in a **sensitized background**. However, these broad suppressor and/or enhancer mutagenesis screens do not reflect the effects that natural genetic variation has on a pathway or phenotype. In particular, mutagenic agents are designed to generate specific types of mutation, at a specific frequency, across the genome. GAL4/UAS screening methods are also designed to select for large-effect mutations. By contrast, natural genetic variation often comprises numerous gene variants with smaller individual effects [19,20]. Natural genetic variation studies rely on standing variation present in a given population and produced by nature. Given that these natural variants are often not loss-of-function alleles, they can have different, unexpected effects on a disease phenotype [19,20]. Modifiers identified by mutagenesis screens will not necessarily overlap with modifiers identified by natural genetic variation methods. Thus, new approaches have been developed to incorporate natural genetic variation in our Drosophila models of human disease.

Several resources have been generated to allow the incorporation of genetic variation into Drosophila models of human disease. Some of these resources include the **Drosophila Genetic Reference Panel** (DGRP) [21] and the **Drosophila Synthetic Population Resource** (DSPR) [22]. The DGRP is the most widely used tool for the study of genetic variation in Drosophila and the main tool used by the two authors of this review. The DGRP is a collection of approximately 200 strains derived from a natural population in North Carolina, USA [21]. Each strain has been extensively inbred and represents a single wild-derived genome from that population. However, the power of this type of resource is not only the number of strains that can be studied, but also that the genome sequences of those strains are available, and, thus, the variants of each strain, allowing for genotype–phenotype correlations. Combined with an easy to use interface for **genome-wide association studies** (GWAS)[v], these strains enable the identification of candidate genes for any quantifiable measurable trait in flies (Box 2).

While the DGRP and similar resources can be powerful for genetic analysis, a short discussion of a few caveats is warranted. Given that the DGRP was collected from a single population, during a single season, it is necessarily only a snapshot of the variation in that population. It certainly is not meant to represent all possible variation. There have also been attempts to capture more worldwide genetic diversity in Drosophila strains [23]. It is important to note that the DGRP contains homozygous, inbred strains and, thus, lethal alleles are not present. Homozygosity does not fully model the possible epistasis in heterozygous individuals, although this can be overcome by performing crosses or creating outbred populations [24]. The inbreeding process also selects for strains that do not carry alleles that result in infertility,

Box 2. Screening for Modifiers with the DGRP

The DGRP is a collection of >200 wild-caught *Drosophila* lines that were inbred for 20 generations per each line and then completely sequenced [21]. These lines represent the normal background variation in a given population (in this case, a population of flies that lived in North Carolina). This variation in the population can have a direct effect on disease phenotypes in humans, accounting for the variation in phenotypic expressivity. Figure I illustrates how we use the DGRP to identify the normal background modifier alleles that influence a given phenotype of interest. Each of the >200 lines is crossed with flies with the appropriate human disease model phenotype and then the severity of the phenotype scored. For example, one can express a disease-associated variant in the eye of the fly with gmr-GAL4 and look for DGRP alleles that make this phenotype better or worse on a grading scale. In Figure I, two different modifier alleles are illustrated: SNP101 is a single allele that was present in two or more of the lines and consistently made the phenotype worse, while the two adjacent alleles, SNP167 and SNP254, were always present together when the rough eye phenotype was suppressed (less rough) in this example. The DGRP strain collection represents over 4.5 million sequenced SNPs, which can be used for this modifier analysis, conferring single-SNP and typically single-gene resolution for these background modifier effects.

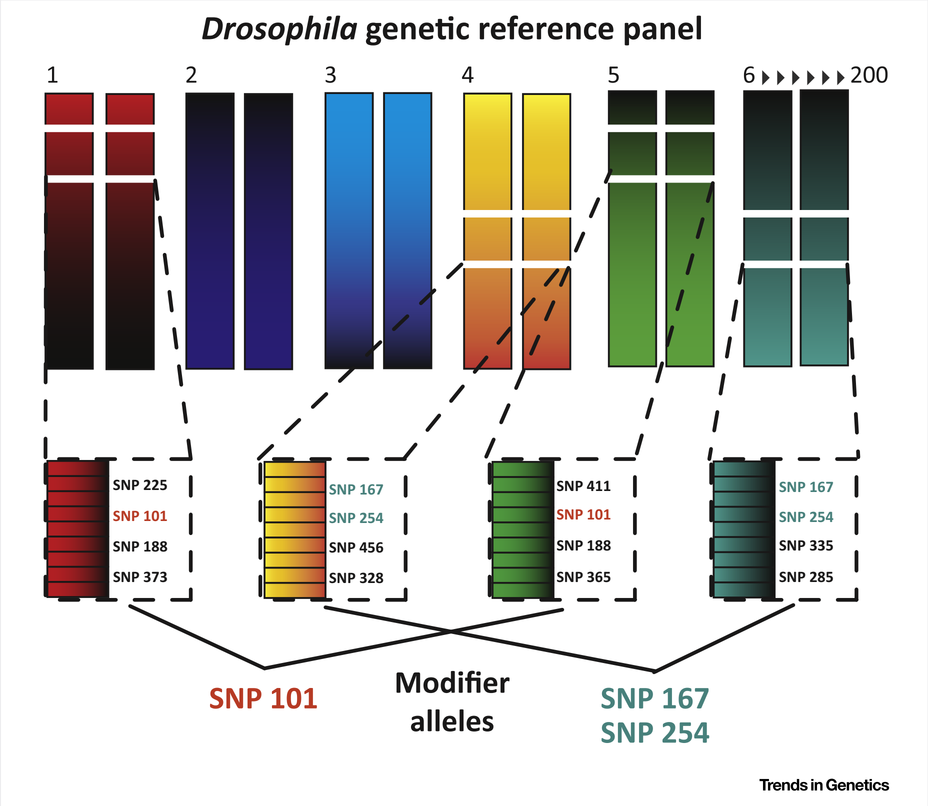

Figure I. SNP Variants Across the DGRP Used for Modifier Screening

again reducing the load of possible genetic variants. Nevertheless, the DGRP remains a powerful tool that can still reveal the role of natural genetic variation in both disease and complex phenotypes.

There are two approaches that have been devised to study the effect of genetic variation on a disease phenotype. The first is to measure a phenotype that does not require genetic crosses. This is by far the most common approach and focuses on phenotypes that can be measured directly from the DGRP strains [23–27] or require some environmental and/or dietary manipulation [28–31]. For example, a recent study measured the variation in susceptibility to endoplasmic reticulum (ER) stress in the DGRP strains [28]. Survival was quantified with a survival statistic for a cohort of flies from each DGRP strain. The study found that there was incredible variation in susceptibility to ER stress across the DGRP strains [28]. By using natural genetic variation in the DGRP, the study also identified several novel members of the ER stress response.

A similar, recently used, approach involved the combination of several quantifiable behavioral phenotypes to identify genetic influences shared across these behaviors. This approach was used to study a complex genetic disorder heavily influenced by natural variation: autism spectrum disorder (ASD). The types of behavior associated with human ASD can be approximated well in flies if distilled to their basic components: social interaction, social communication and repetitive behaviors. All three of these traits can be measured and quantified in each of the approximately 200 DGRP lines, allowing for the identification of both major alleles that act across the behaviors and those that are identified by generation of a **Z-score** for all three behaviors before GWAS. This type of natural variant analysis is not possible in many model systems and is certain to result in new avenues of research for complex traits.

The second, newer approach, is a sensitized screen, where one asks how DGRP backgrounds influence the variation of a phenotype associated with a disease mutation [32–34]. This approach requires at least one genetic cross. Essentially, one is asking which natural variants interact with the primary disease-causing mutation: a classic gene × gene interaction. This approach has been used to test dominant and recessive diseases in the DGRP. The approach to studying a dominant mutation is simpler. To study dominant mutations in the DGRP, one can utilize a single **F1** cross to put the mutant gene into the DGRP background. This generates F1 flies between the 'donor strain' carrying the dominant mutation and each DGRP strain. The F1 progeny that are measured are heterozygous for their respective DGRP parent and the 'donor strain'. This approach has been used to study how **genetic variation** impacts a model of retinal degeneration [34]. A transgene expressing a mutant rhodopsin protein was crossed into the DGRP with a single cross. The mutant rhodopsin causes retinal degeneration in Drosophila and is a model of human retinal degeneration. Indeed, retinal degeneration, as measured by eye size, ranged from very severe to nearly no degeneration in the DGRP.

This approach can also be applied to diseases caused by recessive loss-of-function mutations. Rather than backcrossing a recessive mutation into each of the approximately 200 DGRP strains, GAL4/UAS and the extensive collection of UAS-RNAi lines available are used to create a 'recessive' loss of function genotype in one F1 cross. A donor strain is constructed such that it carries the GAL4 driver construct, the UAS-RNAi construct, and a GAL80 repressor of the UAS sequence. The GAL-80 construct acts to inhibit RNAi expression in the donor strain. This gives a 'healthy' donor strain. In one F1 cross with the DGRP, knockdown flies are created and are identified by scoring against the GAL80 carrying the **balancer chromosome**. This approach is currently being used to study several metabolic diseases in the Drosophila model system.

Whether you are measuring a direct phenotype in the DGRP or crossing in a dominant or recessive mutation, the ultimate goal is to identify the modifier genes that cause phenotypic variation. To identify these modifiers, several different association approaches are used. These approaches take the variable phenotype, and determines, **single nucleotide polymorphism** (SNP) by SNP, whether there are any polymorphisms that appear to associate with the phenotype. The smaller, less complex genome in Drosophila makes this an easier and more powerful approach than association analysis in human populations. Furthermore, the small to non-existent linkage disequilibrium (LD) blocks in flies [26] typically allow for base-pair resolution. An online tool that facilitates this analysis has been developed[v]. These approaches produce rank order-associated SNPs in or around candidate genes. Depending on the ultimate goal, some groups focus follow-up studies on specific associated SNPs and other groups prefer to focus on candidate genes that are tagged by associated SNPs. In many studies, modifiers have always included a few that make biological sense and several novel candidate modifiers, which are particularly exciting. These quantitative screens in the DGRP are just a launching point for more in-depth functional studies. Follow-up studies should take advantage of the other techniques discussed in this review to conduct gene-level analysis that place the new modifiers into the pathways of interest. In the end, describing the extensive variation will not be sufficient:, studies will need to demonstrate the functional consequences of these modifiers.

This approach to studying the influence of genome variation on disease models and disease-relevant phenotypes will lead to new knowledge of even well-studied disorders. These types of result will reveal the nature of how genetic variation modifies disease outcomes across individual variability and how this knowledge might be applied to develop more precise, personalized medicine.

E pluribus unum: Many Genes, but just One Fly

Recurrent CNVs encompassing multiple genes are increasingly more relevant to human disease, but are difficult to model in mammalian systems. Now commonly referred to as genomic disorders [35], CNVs can cause syndromes that result from the deletion or duplication of one or more genes contained within the CNV. Modeling how dosage changes in multiple genes result in disease can be difficult. Often, these syndromes are studied one gene at a time, making it difficult to understand how the genes may interact. These studies are often carried out in mouse models, where cost and time prohibit large genetic interaction studies. Given that flies only have four chromosomes, three of which are typically used for genetic manipulations, four or more genes can be studied simultaneously, using balancer chromosomes and female germline-specific recombination [36].

An excellent example of this approach is illustrated by an investigation of multiple genes thought to be responsible for heart defects in Down syndrome (Trisomy 21). All possible combinations of these three genes were expressed in the fly heart using *hand*-GAL4 until the critical pair of genes was identified [37]. A mouse model was then produced using these two genes and the predicted heart defect was found, saving considerable time and resources compared with carrying out this analysis in the mouse model alone [37]. The Reiter laboratory is currently using a similar approach to investigate the interaction among genes in the 15q Duplication syndrome to tease out the multigenic effects from those duplicated genes. In the near future, using these multigenic approaches in flies may be the ‘first pass’ for the identification of genes responsible for CNV disorders, even if these are rare and occur at extremely low frequency in the human population.

Concluding Remarks

The primary message to the genetics community is that *D. melanogaster* is far from a quaint genetic model of the past, but rather, continues to evolve as a powerful system for the study of human genetic disease. As we continue to model more complex mutagenic conditions in flies, their utility only increases for understanding gene function and the influence of genetic background (see Outstanding Questions). Flies are becoming even more powerful because of new combinatorial approaches to the study of complex traits, such as autism, intellectual disability, and other human conditions considered difficult to explore using traditional GWAS or patient-centered methodologies. Their low cost, high yield, and tremendous number of available tools means that we can expect great progress in using the fly to understand human disease.

Outstanding Questions

What disorders are appropriate for disease modeling in flies beyond the nervous system?

How will synergistic background variant effects translate to mammals? What if the same modifier alleles do not exist in mammals?

How will having redundancy for many genes in humans versus one copy in flies complicate the influence of background effects on phenotype?

Will these new models of human disease be useful for drug screening?

If a variant is not pathogenic in flies, but was found in humans, what are the next steps in understanding this variant?

Acknowledgments

The authors wish to thank Hsiao-Tuan Chao and Hugo Bellen for allowing us to modify their previously published figure, as well as Trudy MacKay and Robert Anholt for developing the DGRP collection.

Resources

[i]http://flystocks.bio.indiana.edu/
[ii]http://fgr.hms.harvard.edu/fly-in-vivo-rnai
[iii]http://flyorf.ch/
[iv]http://marrvel.org
[v]http://dgrp2.gnets.ncsu.edu/
[vi]http://wfitch.bio.uci.edu/∼dspr/

References

1. Morgan, T.H. *et al.* (1923) *The Mechanism of Mendelian Heredity.* (2nd edn), H. Holt & Co
2. Nusslein-Volhard, C. and Wieschaus, E. (1980) Mutations affecting segment number and polarity in Drosophila. *Nature* 287, 795–801
3. McGurk, L. *et al.* (2015) Drosophila as an in vivo model for human neurodegenerative disease. *Genetics* 201, 377–402
4. Ugur, B. *et al.* (2016) Drosophila tools and assays for the study of human diseases. *Dis. Model. Mech.* 9, 235–244
5. Adams, M.D. *et al.* (2000) The genome sequence of *Drosophila melanogaster*. *Science* 287, 2185–2195
6. Reiter, L.T. *et al.* (2001) A systematic analysis of human disease-associated gene sequences in *Drosophila melanogaster*. *Genome Res.* 11, 1114–1125
7. Wangler, M.F. *et al.* (2015) Fruit flies in biomedical research. *Genetics* 199, 639–653
8. Brand, A.H. and Perrimon, N. (1993) Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118, 401–415
9. Casci, I. and Pandey, U.B. (2015) A fruitful endeavor: modeling ALS in the fruit fly. *Brain Res.* 1607, 47–74
10. Fernandez-Funez, P. (2015) Modeling the complex pathology of Alzheimer's disease in Drosophila. *Exp. Neurol.* 274, 58–71
11. Xu, Z. *et al.* (2015) Studying polyglutamine diseases in Drosophila. *Exp. Neurol.* 274, 25–41
12. Jenett, A. *et al.* (2012) A GAL4-driver line resource for Drosophila neurobiology. *Cell Rep.* 2, 991–1001
13. Diao, F. *et al.* (2015) Plug-and-play genetic access to Drosophila cell types using exchangeable exon cassettes. *Cell Rep.* 10, 1410–1421
14. Venken, K.J. *et al.* (2011) MiMIC: a highly versatile transposon insertion resource for engineering *Drosophila melanogaster* genes. *Nat. Methods* 8, 737–743
15. Chao, Y.H. *et al.* (2016) Missense variants in the middle domain of DNM1L in cases of infantile encephalopathy alter peroxisomes and mitochondria when assayed in Drosophila. *Hum. Mol. Genet.* 25, 1846–1856
16. Harel, T. *et al.* (2016) Recurrent de novo and biallelic variation of ATAD3A, encoding a mitochondrial membrane protein, results in distinct neurological syndromes. *Am. J. Hum. Genet.* 99, 831–845
17. Chao, H.T. *et al.* (2016) A syndromic neurodevelopmental disorder caused by de novo variants in EBF3. *Am. J. Hum. Genet.* 100, 128–137
18. Jakobsdottir, J. *et al.* (2016) Rare functional variant in TM2D3 is associated with late-onset Alzheimer's disease. *PLoS Genet.* 12, e1006327
19. Chow, C.Y. (2015) Bringing genetic background into focus. *Nat. Rev. Genet.* 17, 63–64
20. Gasch, A.P. *et al.* (2016) The power of natural variation for model organism biology. *Trends Genet.* 32, 147–154
21. Mackay, T.F. *et al.* (2012) The *Drosophila melanogaster* Genetic Reference Panel. *Nature* 482, 173–178
22. Long, A.D. *et al.* (2014) Dissecting complex traits using the Drosophila Synthetic Population Resource. *Trends Genet.* 30, 488–495
23. Chow, C.Y. *et al.* (2013) Large neurological component to genetic differences underlying biased sperm use in Drosophila. *Genetics* 193, 177–185
24. Hunter, C.M. *et al.* (2016) The genetic architecture of natural variation in recombination rate in *Drosophila melanogaster*. *PLoS Genet.* 12, e1005951
25. Unckless, R.L. *et al.* (2015) A genome-wide association study for nutritional indices in Drosophila. *G3 (Bethesda)* 5, 417–425
26. Huang, W. *et al.* (2014) Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. *Genome Res.* 24, 1193–1208
27. Harbison, S.T. *et al.* (2013) Genome-wide association study of sleep in *Drosophila melanogaster*. *BMC Genom.* 14, 281
28. Chow, C.Y. *et al.* (2013) Using natural variation in Drosophila to discover previously unknown endoplasmic reticulum stress genes. *Proc. Natl. Acad. Sci. U. S. A.* 110, 9013–9018
29. Unckless, R.L. *et al.* (2015) The complex contributions of genetics and nutrition to immunity in *Drosophila melanogaster*. *PLoS Genet.* 11, e1005030
30. Zhou, S. *et al.* (2016) The genetic basis for variation in sensitivity to lead toxicity in *Drosophila melanogaster*. *Environ. Health Perspect.* 124, 1062–1070
31. Howick, V.M. and Lazzaro, B.P. (2017) The genetic architecture of defense as resistance to and tolerance of bacterial infection in *Drosophila melanogaster*. *Mol. Ecol.* 26, 1533–1546
32. He, B.Z. *et al.* (2014) Effect of genetic variation in a Drosophila model of diabetes-associated misfolded human proinsulin. *Genetics* 196, 557–567
33. Park, S.Y. *et al.* (2014) Genetic complexity in a Drosophila model of diabetes-associated misfolded human proinsulin. *Genetics* 196, 539–555
34. Chow, C.Y. *et al.* (2015) Candidate genetic modifiers of retinitis pigmentosa identified by exploiting natural variation in Drosophila. *Hum. Mol. Genet.* 25, 651–659
35. Mikhail, F.M. (2014) Copy number variations and human genetic disease. *Curr. Opin. Pediatr.* 26, 646–652
36. Lindsley, D.L. *et al.* (1992) *The Genome of Drosophila melanogaster,* Academic Press
37. Grossman, T.R. *et al.* (2011) Over-expression of DSCAM and COL6A2 cooperatively generates congenital heart defects. *PLoS Genet.* 7, e1002344