**SUMMARY**

Invertebrate model systems are powerful tools for studying human disease owing to their genetic tractability and ease of screening. We conducted a mosaic genetic screen of lethal mutations on the *Drosophila* X chromosome to identify genes required for the development, function, and maintenance of the nervous system. We identified 165 genes, most of whose function has not been studied in vivo. In parallel, we investigated rare variant alleles in 1,929 human exomes from families with unsolved Mendelian disease. Genes that are essential in flies and have multiple human homologs were found to be likely to be associated with human diseases. Merging the human data sets with the fly genes allowed us to identify disease-associated mutations in six families and to provide insights into microcephaly associated with brain dysgenesis. This bidirectional synergism between fly genetics and human genomics facilitates the functional annotation of evolutionarily conserved genes involved in human health.

**INTRODUCTION**

Unbiased genetic chemical mutagenesis screens in flies have led to the discovery of the vast majority of genes in developmental signaling pathways (Nüsslein-Volhard and Wieschaus, 1980). Most genes important to these pathways have now been shown to function as oncogenes or tumor suppressors...
In the fly screen, we assessed the number and size of phenotypic consequences makes it challenging to tie a specific variant/gene to a given disease phenotype. Yet, these rare variants have a strong contribution to disease (Lupski et al., 2011). The interpretation of such genome-wide variation is hindered by our lack of understanding of gene function for the majority of annotated genes in the human genome.

We identified mutations in 165 genes, most of which have not been characterized previously in vivo. We provide data that suggest this gene set can be utilized as a resource to study numerous disease-causing genes. In addition, we present data that there is a fundamental difference between ethyl methane-sulfonate (EMS) screens and RNAi screens. Moreover, we show that fly genes with more than one homolog are much more likely to be associated with human genetic disorders. Finally, we demonstrate that merging data sets—genes identified in the fly screen and rare variant alleles in the human homologs in families with Mendelian disease—can assist in human disease gene discovery and provide biological insights into disease mechanisms.

RESULTS

A Mosaic Genetic Screen on the X Chromosome

To isolate mutations in essential genes that are required for proper development, function, and maintenance of the *Drosophila* nervous system, we performed an F3 adult mosaic screen on an isogenic (iso) y w FRT19A X chromosome (Figure 1 and Figures S1 and S2 available online). We mutagenized males using a low concentration of ethyl methane-sulfonate (EMS), established 31,530 mutagenized stocks, and identified 5,857 stocks that carry recessive lethal mutations. To identify a broad spectrum of mutations and isolate genes that affect multiple biological processes, we screened for numerous phenotypes that affect the nervous system. We also screened for seemingly unrelated phenotypes, such as wing and pigmentation defects. Genes that affect wing veins and notching have been shown to play roles in critical pathways that affect numerous organs, including the nervous system. To assess phenotypes in the tissues of interest, we induced mitotic clones in the thorax and wing with Ultrabithorax-flippase (Ubx-FLP) (Jafar-Nejad et al., 2005) and in the eye with eyeless-flippase (ey-FLP) (Newsome et al., 2000). We did not pursue mutations that caused cell lethality or showed no/minor phenotypes (Figure 1A). While these genes are clearly important, they are difficult to study and these mutants were not kept. We selected 2,083 lethal lines with interesting phenotypes for further characterization (Figures 1A and 1B).

In the Ubx-FLP screen, we assessed the number and size of mechanosensory organs (bristles) on the fly cuticle to identify genes required for neural development (Figures 1C and 1D and S2A–S2C) (Chang et al., 2014). We also screened for alterations in the color of bristles and cuticle to permit identification of genes involved in dopamine synthesis, secretion, metabolism, or mela-nization (Yamamoto and Seto, 2014) (Figure S2D). In addition, we selected mutations that affect wing morphogenesis to isolate genes that regulate core signaling pathways, including Notch, Wnt, Hedgehog, and BMP/TGF-β (Bier, 2005) (Figures S2E–S2J). Indeed, these pathways have been implicated in synaptic plasticity and neuronal maintenance in both fly and vertebrate nervous systems. In the ey-FLP screen, we assessed morphological defects in the eye and head to isolate genes involved in neuronal patterning, specification, and differentiation (Figures S2K–S2O). Moreover, we screened for mutations that cause glossy eye patches (Figure S2P) or mutations that cause a head overgrowth (Figures S2Q–S2S). Glossy eye phenotypes are associated with mitochondrial mutations (Liao et al., 2006), while head overgrowth is linked to genes in Hippo signaling, TOR signaling, intracellular trafficking, and cell polarity/adhesion, and these pathways are implicated in disorders such as autism, intellectual disability, and neurodegenerative diseases (Emoto, 2012; Saksena and Emr, 2009).

To isolate mutations that affect neuronal development, function, and maintenance in the visual system, we recorded electroretinograms (ERGs) in mutant eye clones in 3- to 4-week-old flies (Figures 1E–1I). By analyzing the on and off transients of ERGs
A

B

Notum
(Ubx-FLP)

Wing
(Ubx-FLP)

Eye/Head
(cl(1);ey-FLP)

Morphological Defects

<table>
<thead>
<tr>
<th>Defect</th>
<th>Morphology</th>
<th>ERG + Morphology</th>
<th>ERG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balding</td>
<td>56</td>
<td>230</td>
<td></td>
</tr>
<tr>
<td>Macrochaetes Loss</td>
<td>51</td>
<td>197</td>
<td></td>
</tr>
<tr>
<td>Macrochaetes Gain</td>
<td>12</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>Short Bristles</td>
<td>84</td>
<td>256</td>
<td></td>
</tr>
<tr>
<td>Depigmentation</td>
<td>8</td>
<td>140</td>
<td></td>
</tr>
<tr>
<td>No Eclosion</td>
<td>23</td>
<td>265</td>
<td></td>
</tr>
<tr>
<td>Notching</td>
<td>25</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>Vein Gain/Loss</td>
<td>29</td>
<td>145</td>
<td></td>
</tr>
<tr>
<td>Ectopic Wing Margin</td>
<td>10</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Bristles on Wing Blade</td>
<td>9</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Wing Blister</td>
<td>34</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Crinkled Wings</td>
<td>84</td>
<td>233</td>
<td></td>
</tr>
<tr>
<td>Glossy Eyes</td>
<td>38</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>Ectopic Eyes</td>
<td>17</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>Head Overgrowth</td>
<td>10</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>Head Cuticle Overgrowth</td>
<td>10</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>Rough Eyes</td>
<td>61</td>
<td>551</td>
<td></td>
</tr>
<tr>
<td>Small Heads</td>
<td>23</td>
<td>276</td>
<td></td>
</tr>
<tr>
<td>No Eclosion</td>
<td>0</td>
<td>150</td>
<td></td>
</tr>
<tr>
<td>ERG Defect</td>
<td>0</td>
<td>150</td>
<td></td>
</tr>
</tbody>
</table>

C

D

E

F

G

H

I

J

K

L

M

(legend on next page)
(Figure 1H), one can assess photoreceptor synaptic activity and axon guidance. A loss or reduction in the amplitude of depolarization (Figure 1G) is typically associated with genes that play a role in phototransduction, loss of which typically causes retinal degeneration (Wang and Montell, 2007). To identify mutations that cause a progressive demise of neurons, we screened young and old animals for ERG defects (Figures 1F and 1I). Ultrastructural defects in the photoreceptor terminals of young and old flies were also examined in some mutants with strong ERG phenotypes (Figures 1J–1M). Based on both the morphology screen and the ERG screen, we attempted to map 1,918 mutations (Figures S1 and S3).

**Mutation Identification**

On the X chromosome, complementation testing requires a genomic duplication on another chromosome to rescue male lethality. We selected 21 large (~0.5 Mb to ~2 Mb) duplications that cover ~95% of the X chromosome (Cook et al., 2010), crossed them into the mutant backgrounds, and rescued the lethality of 1,385 mutations (Figure S3). This permitted mapping of the lethality to 26 cytological intervals of the X chromosome. Complementation tests between mutants with similar phenotypes rescued by the same duplication allowed us to establish complementation groups. We grouped 450 mutations into 109 multiple allele complementation groups. The remaining 935 mutant strains include single alleles and a large number of mutations not yet assigned to complementation groups. To map the genes, we first performed deficiency mapping and Sanger sequencing. This allowed identification of the locus for 63 complementation groups. For the remaining groups and single alleles, we performed WGS (Haelterman et al., 2014) and rescued the phenotypes with molecularly defined ~80 kb [Bac-man] duplications (Venken et al., 2010). By using both approaches, we were able to map 614 mutations to 165 genes, including 81 loci that have not been characterized in vivo (Tables 1 and S1) and are predicted to be involved in many diverse processes based on gene ontology analysis (Figures S2T and S2U).

**Chemical Mutagenesis versus RNAi Screens**

Two of the phenotypes that we screened, bristle development and depigmentation, allow a direct comparison between this screen and a genome-wide RNAi screen (Mummyery-Widmer et al., 2009). This RNAi screen covered ~80% of all X chromosome protein coding genes. Interestingly, only 14% of the genes we identified in the bristle screen were also isolated in the RNAi screen (Figures 2A and 2B). Similarly, only 18% of the genes that we identified from the pigmentation screen were also identified in the RNAi screen (Figures 2C and 2D). Conversely, we did not identify the vast majority of genes that were identified by RNAi. In addition, a comparison of our gene list and those of two RNAi screens for wing margin (Saj et al., 2010) and eye morphological defects (Oortveld et al., 2013), show that these screens also identified very different sets of genes (Figures 2E and 2F). In summary, chemical screens identify a distinctive set of genes when compared to RNAi-based screens.

**Links to Human Diseases Based on Online Mendelian Inheritance in Man**

We next sought to determine if the 165 genes we identified in flies could enhance the understanding of human disease associated genes. Strikingly, 93% (153) of the fly genes isolated have homologs in humans (Tables 1 and S1; Figure 3A). This is a strong enrichment ($\chi^2 = 129$, $p < 0.001$) for evolutionarily conserved genes between humans and flies when compared to the whole fly genome as only 48% of all fly genes have human homologs (Figure 3B). Moreover, the human homologs of 31% (48/153) of the identified fly genes have been associated with human disease in Online Mendelian Inheritance in Man (OMIM), 79% (38/48) of which exhibit neurological signs and symptoms (Figure 3A; Table S1). Of the genes that are conserved but not yet associated with Mendelian diseases with neurological symptoms, 65 genes have potential relationships to neurologic diseases (Figure 3A; Table S2). Therefore, the essential genes that we identified in this screen are highly conserved and many of their homologs have already been implicated in human disorders, showing that the screening strategy is effective.

Data analysis revealed a striking difference in the number of genes associated with disease depending on the number of human homologs for each fly gene. Fly genes that have a single human homolog have many fewer disease genes represented in the OMIM database than those that have more than one homolog. There is a 2-fold enrichment ($\chi^2 = 10.7$, $p < 0.001$) of fly genes with more than one human homolog associated with diseases in the OMIM database compared to fly genes that have single homologs.

---

**Figure 1. Summary of the Drosophila X Chromosome Screen**

(A and B) Pie chart (A) and bar graph (B) of phenotypes scored in the screen. The numbers represent mutations in each phenotypic category. Note that one strain may show more than one phenotype in (B).

(C and D) Examples of phenotypes observed in the notum. (C) Clones induced in a wild-type background, clone borders are marked by a white dotted line, (D) example of bristle loss in mutant clones (white arrows) (see Extended Experimental Procedures).

(E–I) Examples of ERG traces from mutant clones in the eye. A typical ERG has an on transient (blue arrows), depolarization (orange line) and an off transient (blue arrow head). ERGs were recorded in young (1- to 3-day-old) and old (3- to 4-week-old) flies for each genotype. (E) ERG of young or aged flies that show no obvious (E–I) Examples of ERG traces from mutant clones in the eye. A typical ERG has an on transient (blue arrows), depolarization (orange line) and an off transient (blue arrow head). ERGs were recorded in young (1- to 3-day-old) and old (3- to 4-week-old) flies for each genotype. (E) ERG of young or aged flies that show no obvious difference. (F) ERGs showing amplitude reduction in aged flies. (G) ERGs showing amplitude and on- and off-transient reduction in both young and aged mutants. (H) ERGs showing no or very small on transient in both young and aged flies. (I) ERGs showing on and off transients that are either absent or very small in aged flies carrying mutant clones in eye.

(J–M) Ultrastructural analysis using transmission electron microscopy (TEM) on young (2-day-old) and aged (3-week-old) mosaic flies. Red arrowheads indicate the rhabdomeres. (J) Young wild-type control eye: regular array of ommatidial structures with seven rhabdomeres surrounded by pigment (glia) cells. (K) Young mutant rhabdomeres showing intact structures. (L) Aged control eye tissue with intact rhabdomeres. (M) Aged mutant eye tissue with a strong degeneration of rhabdomeres.

See also Figures S1, S2, S3.
Table 1. List of 165 Fly Genes and 259 Corresponding Human Homologs Identified from the Screen

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Aats-his</td>
<td>HARS*, HARS2*</td>
<td>COQ7</td>
<td>COQ7</td>
<td>para</td>
<td>SCN1A*, SCN2A*, SCN3A,</td>
</tr>
<tr>
<td>AP-1γ</td>
<td>AP1G1, AP1G2, AP4E1*</td>
<td>Crag</td>
<td>DENND4A, DENND4B,</td>
<td>SCN4A*,</td>
<td></td>
</tr>
<tr>
<td>ari-1</td>
<td>ARIH1</td>
<td></td>
<td>DENND4C</td>
<td>SCN5A*,</td>
<td></td>
</tr>
<tr>
<td>arm</td>
<td>CTNNB1*</td>
<td>Cyp4d2</td>
<td>CYP4V2*</td>
<td>SCN6A*,</td>
<td></td>
</tr>
<tr>
<td>Arp2</td>
<td>ACTR2</td>
<td>DAAM</td>
<td>DAAM1, DAAM2</td>
<td>SCN7A*,</td>
<td></td>
</tr>
<tr>
<td>ATP7</td>
<td>ATP7A*, ATP7B*</td>
<td>Cg1</td>
<td>DENND4A, DENND4B,</td>
<td>SCN8A*,</td>
<td></td>
</tr>
<tr>
<td>baz</td>
<td>PARD3, PARD3B</td>
<td>Dl1c</td>
<td>DENND4C</td>
<td>SCN9A*,</td>
<td></td>
</tr>
<tr>
<td>ben</td>
<td>UBE2N</td>
<td>dorp</td>
<td>VPS18</td>
<td>SCN10A*,</td>
<td></td>
</tr>
<tr>
<td>br</td>
<td></td>
<td>dsh</td>
<td>jong</td>
<td>SCN11A*</td>
<td></td>
</tr>
<tr>
<td>Brms1</td>
<td>BRMS1, BRMS1L</td>
<td>Dsor1</td>
<td>MAP2K1*, MAP2K2*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cac</td>
<td>CACNA1A*, CACNA1B</td>
<td>dwg</td>
<td>MZF1, ZSCAN22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cap</td>
<td>SMC3*</td>
<td>egh</td>
<td>Pro4V4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>car</td>
<td>VPS33A</td>
<td>elf2B+</td>
<td>EFL2B5*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDC45L</td>
<td>CDC45L</td>
<td>elav</td>
<td>ELAV1, ELAV2,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cdk7</td>
<td>Cdk7</td>
<td>ELAV3, ELAV4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG11409</td>
<td></td>
<td>ewg</td>
<td>NRF1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG11417</td>
<td>ESF1</td>
<td>fh</td>
<td>FXN*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG11418</td>
<td>MTPAP*</td>
<td>flh</td>
<td>FLII</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG12125</td>
<td>FAM73A, FAM73B</td>
<td>flw</td>
<td>PPP1CB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG12991</td>
<td></td>
<td>fs(l)1h</td>
<td>BRD2, BRD3, BRD4, BRDT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG13385</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG14442</td>
<td>ZNF821</td>
<td>Fur2</td>
<td>PCSK5, PCSK6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG14786</td>
<td>LRPPRC*</td>
<td>Glp-bp</td>
<td>SRPR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG15208</td>
<td>C21or2</td>
<td>hfw</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG15896</td>
<td>KIAA0391</td>
<td>Hic</td>
<td>DDX56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG1597</td>
<td>MOGS*</td>
<td>hop</td>
<td>JAK1, JAK2*, JAK3*, TYK2*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG1703</td>
<td>ABCF1</td>
<td>hr4</td>
<td>NRP6A1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG1749</td>
<td>UBA5</td>
<td>Hsc70-3</td>
<td>HSPA5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG17776</td>
<td></td>
<td>Inx2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG17829</td>
<td>HINFP</td>
<td>kdn</td>
<td>CS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG18624</td>
<td>NDUFB1</td>
<td>l(1)1Bi</td>
<td>MYBBP1A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG20025</td>
<td>NRD1</td>
<td>l(1)G0156</td>
<td>IDH3A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG2918</td>
<td>HYOU1</td>
<td>l(1)G0222</td>
<td>ANKLE2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG3011</td>
<td>SHMT1, SHMT2</td>
<td>l(1)G0230</td>
<td>ATP5D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG3149</td>
<td>RFT1*</td>
<td>l(1)G0255</td>
<td>FH*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG32649</td>
<td>ADCK3*, ADCK4</td>
<td>l(1)G0334</td>
<td>PDHA1*, PDHA2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG32694</td>
<td></td>
<td>Marf</td>
<td>MFN1, MFN2*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG32795</td>
<td>TEMEM120A, TEMEM120B</td>
<td>Mcm6</td>
<td>MCM6*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG34001</td>
<td>ZSWIM8</td>
<td>mew</td>
<td>ITGA3*, ITGA6*, ITGA7*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG3446</td>
<td>NDUFA13*</td>
<td>mRNA-cap</td>
<td>RINGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG3704</td>
<td>GPN1</td>
<td>mRpl38</td>
<td>MRPL38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG3857</td>
<td>SMG9</td>
<td>mRps25</td>
<td>MRPS25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG4078</td>
<td>RTEL1*</td>
<td>mRps30</td>
<td>MRPS30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG4165</td>
<td>USP16</td>
<td>mst</td>
<td>MSTO1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG42337</td>
<td>PLA2G3, PROCA1</td>
<td>mus101</td>
<td>TOPBP1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG42593</td>
<td>UBR3</td>
<td>mxc</td>
<td>NPAT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued on next page)
only one human homolog, 47% versus 22% (Figure 3C). This prompted us to assess if the bias is conserved for all fly genes. We found that a similar bias holds throughout the genome. Fly genes with more than one human homolog are more likely to be associated with diseases in the OMIM database than those with a single homolog, 40% versus 20% (χ² = 386, p < 0.001) (Figure 3D and Extended Experimental Procedures). Indeed, 57 fly genes with more than one human homolog account for 100 diseases in the OMIM database (1.75 diseases per fly gene), an 8-fold enrichment when compared to fly genes with a single homolog (0.22 diseases per fly gene) (Figure 3E). This enrichment is not simply due to an absolute increase in the total number of human homologs because evolutionarily conserved genes that have more than one homolog are three times more enriched for OMIM diseases, 0.62 versus 0.22 diseases per human gene (Figure 3E). The difference between 1.75 and 0.62 is due to the number of homologs. Indeed, there are on average ~3 human homologs for every fly gene that has more than one human homolog (data not shown). These data suggest that evolutionary gene duplication with divergence and further specialization of gene function may allow tolerance of mutation and viability versus lethality.

Since all of the mutations we isolated cause homozygous lethality, we analyzed the correlation between lethality, the number of human homologs, and their links to OMIM diseases for the entire fly genome. The number of essential genes in Drosophila has been estimated to be approximately 5,000 (Benos et al., 2001). Currently only ~2,000 essential genes in FlyBase have transposable elements or EMS/X-ray-induced mutations (Marygold et al., 2013), representing about 40% of all essential fly genes. The proportion of essential genes varies with evolutionary conservation: an estimated 11% of the genes that do not have human homologs are essential, whereas 38% of the genes that have a single human homolog are essential (χ² = 354, p < 0.001) (Figure 3F). Finally, an estimated 61% of the fly genes with more than one human homolog are essential. These data show that fly genes that have more than one human homolog are more likely to cause lethality when mutated. Finally, human homologs of essential genes in Drosophila are more likely to be associated with human genetic diseases (χ² = 88, p < 0.001) (Figure 3G). Therefore, we conclude that genes that are essential in flies and have multiple human homologs are the most likely to be associated with human diseases, potentially due to gene duplication and redundancy.

Combining Fly and Human Mutant Screen Data Sets to Identify Disease Genes

We next utilized the fly gene data set uncovered from the forward genetic screen in combination with a human exome data set to identify new human disease genes. We undertook a systematic search of all the variants in the human homologs of the genes identified from the Drosophila screen within WES data generated from undiagnosed cases of Mendelian diseases. This included 1,929 individuals in the Baylor-Hopkins Centers for Mendelian Genomics (BHCMG) (Figure 4).

BHCMG uses next-generation sequencing to discover the genetic basis of as many Mendelian diseases as possible (Bamshad et al., 2012). The study population includes singleton cases with sporadic disease, single families, and when possible, larger cohorts of affected individuals with a range of rare Mendelian phenotypes. A wide range of disorders are under investigation (http://www.mendelian.org/). In general, patients are recruited when a Mendelian disease seems highly likely and all reasonable efforts at a molecular diagnosis have failed. Due to the rare nature of the phenotypes, information from other patients or additional biological information from model organisms is required to fulfill the burden of proof for gene/disease association in such cases. For this reason, our Drosophila resource of mutant genes was integrated with our human exome variant and Mendelian phenotype (Hamosh et al., 2013) databases, and the combination approach was used to solve some of the cases.

We analyzed 237 out of the 259 (Table 1) homologs of fly genes identified through the X chromosome screen as they were validated at the time of analysis. We included all 237 genes,
regardless of whether they were previously identified to be associated with Mendelian diseases in OMIM, to avoid any bias. We filtered out variants reported as having greater than 1% allele frequency in databases of control individuals (See Extended Experimental Procedures). Under the assumption of a recessive model data set, we included all variants that met these criteria and were homozygous or had two heterozygous variants affecting the same gene. The latter set was not tested for cis or trans orientation of the variants prior to analysis. A dominant model included heterozygous variants. These were filtered even more stringently for allele frequency such that only variants that had not been observed in the control data sets were studied (Table S3).

To explore potential associations with disease, we prioritized variants for segregation analysis within families (Figure 4). We performed Sanger sequencing or explored segregation in families for 64 variants in 24 genes within 34 individuals in the recessive data set and found that 15 variants in 8 genes within 10 individuals fulfilled Mendelian expectations for recessive inheritance. Likewise, for the dominant data set, we explored the segregation for 158 variants in 85 genes within 99 individuals. We found 22 variants in 15 genes within 21 individuals that fulfilled Mendelian expectations of a dominantly inherited disorder in the family under investigation. Interestingly, 22/31 individuals in which the variant met Mendelian expectations had a neurological disease.

As a proof-of-principle, we report six patients/families with mutations in three genes. In addition, we identified 25 other individuals in which the variant in the homolog of the fly gene met Mendelian expectation. Some of these individuals were found to have candidate variants in multiple genes, some had too few living relatives for further study, and for others, studies are ongoing. Therefore, a systematic search of variants within the genes identified in the Drosophila screen was able to identify and prioritize a subset of variants with Mendelian inheritance in families that could be studied. Among these, we found examples of known disease genes (DNM2), a novel disease association to a known disease gene (CRX), and novel candidate genes for disease (ANKLE2).

**DNM2 and Charcot-Marie-Tooth Neuropathy**

Examination of a homolog of Drosophila shibire (shi), the gene that encodes Dynamin, led to a molecular diagnosis for two individuals with heterozygous mutations in DNM2 (Figures S4A and S4B). Both patients were diagnosed with a distal symmetric polyneuropathy consistent with Charcot-Marie-Tooth disease (CMT) (See Extended Results). Mutations in DNM2 are associated with CMT Type 2M (OMIM 606482), an axonal form primarily affecting neurons (Figure S4C). Patient 1, the proband in Figure S4A, presented at age 12 with hand tremor, calf cramps, lower limb paresthesias, and difficulty with heel walking. She is a member of a family with three generations of neuropathy (Figure S4A). Patient 2, the proband in Figure 4B, currently 88 years old, presented at age 40 with lower extremity weakness. His nerve conduction studies showed low amplitudes and borderline...
slowed velocities. He carries an E341K mutation in DNM2 (Figure S4D). In addition to DNM2, WES revealed a variant in another CMT gene, LRSAM1 in this patient (Figure S4B). Interestingly, dominant as well as recessive mutations in LRSAM1 can cause CMT2P (OMIM 614436). Hence, either one or a combination of both genes may cause CMT in this family. While some clinical features of the probands made diagnosis difficult, the phenotypes of these cases were indeed consistent with CMT type 2.

CRX and Bull’s Eye Maculopathy

Examination of one of the human homologs of Drosophila ocelliless (oc, CRX in humans) led to the identification of three cases of bull’s eye maculopathy associated with dominant CRX alleles. oc encodes a homeobox transcription factor that regulates photoreceptor development (Vandendries et al., 1996). Identifying cases of bull’s eye maculopathy, a late-onset slowly progressive retinal disorder, with CRX alleles was surprising because CRX is typically associated with much more severe childhood vision loss seen in dominant cone-rod dystrophy, Leber congenital amaurosis, and autosomal dominant retinitis pigmentosa (OMIM 120970, 613829).

The three cases of bull’s eye maculopathy included two individuals with no family history of retinal disease (patients 3 and 4) and one multigenerational pedigree (patient 5 [S150X]) (Figure 5A). The affected individuals in the family of patient 5 developed symptoms at age 50 (range 28–63 years), and three family members with the S150X mutation had minimal symptoms at initial evaluation between the age of 55–60. Despite having near normal vision, ophthalmologic exam in the retina of these individuals revealed advanced bull’s eye maculopathy with foveal sparing explaining the modest effect on vision.

Patient 5 exhibits retinal abnormalities (Figure 5B–B’), abnormal autofluorescence in the fundus (Figure 5C–C’), aberrant Optical Coherence Tomography (OCT, Figure 5D–D’) and electroretinograms (Figure 5E), all consistent with bull’s eye maculopathy. The three new alleles are all encoding predicted truncations of the OTX transcription factor domain (Figure 5F).

Conservation of fly genes and their links to OMIM diseases

Figure 3. Essential Fly Genes Associated with More Than One Human Homolog Are More Likely to be Linked to Human Diseases

(A) Classification of genes identified in the screen based on human homologs and associated diseases.

(B) Classification of the whole fly genome according to the same criteria as in (A).

(C and D) Relationship between the number of human homologs per fly gene and their association with human diseases for genes identified in the screen (C) and the whole fly genome (D).

(E) The number of human homologs per fly gene and their enrichment in OMIM associated human diseases.

(F) Relationship between the number of human homologs per fly gene and lethality in flies.

(G) Relationship between genes associated with lethality in flies and OMIM associated human diseases.

See also Table S2.
variants in \textit{ANKLE2} in a family with apparent recessive microcephaly (Figures 6B and 6C). The proband, patient 6, has an extreme small head circumference, a low sloping forehead, ptosis, small jaw, multiple hyper- and hypopigmented macules over all areas of his body, and spastic quadriplegia (Figure 6D–6H; Extended Results, “Clinical Case Histories”). During his first year of life, he had unexplained anemia, and glaucoma. At 3 years, he had onset of seizures, and at 5.5 years, his weight was 10.7 kg (−4 SD), length 83.8 cm (−6 SD) and fronto-occipital circumference 38.2 cm (−9 SD).

Brain MRI in the newborn period demonstrated a low forehead, several scalp ruggae, and mildly enlarged extra-axial space with communication between the posterior lateral ventricles and the mesial extra-axial space. Other brain abnormalities included a simplified gyral pattern, mildly thickened cortex, small frontal horns of the lateral ventricles with mildly enlarged posterior horns of the lateral ventricles, and agenesis of the corpus callosum. The brainstem and cerebellum appeared relatively normal (Figures 6G and 6H). A younger sister born a year later had severe microcephaly, spasticity, and similar hyper- and hypopigmented macules over all areas of her body. She died 24 hr after delivery from cardiac failure associated with poor contractility, although the basis for this was not known.

WES data of the proband, his affected sister, and both parents revealed four candidate genes that meet Mendelian expectation and are expressed in the CNS (Table S4). Table S4 shows the variants with their scores from four predictions programs (Liu et al., 2011). \textit{ANKLE2} was prioritized as a good candidate. To assess if \textit{dAnkle2} is involved in CNS development, we examined the brains of \textit{Drosophila} mutant larvae. Brain size in early third instar larval stages is similar to that of controls (Figure S5A). However, later in third larval stage, the brain becomes progressively smaller than control larvae (Figure S5A and Figures 6I and J). To confirm that \textit{dAnkle2} is an ortholog of human \textit{ANKLE2}, we ubiquitously expressed human \textit{ANKLE2} in mutant flies and observed rescue of lethality and the small brain phenotype (Figures 6K–6L). These data indicate that \textit{ANKLE2} is implicated in CNS development and its molecular function is evolutionarily conserved.

To explore the cause of the small brain phenotype in \textit{dAnkle2} mutants, we assessed defects in processes which can cause small brain phenotypes: mitosis, asymmetric cell division, and apoptosis (Rujano et al., 2013). The number of neuroblasts, marked by Miranda (Ceron et al., 2001) is severely reduced in late third instar brain lobes (Figures 6M–6O and S5B and S5C). In the few neuroblasts that undergo division, the spindles are properly oriented toward the polarity axis (Figures S5D and S5E). In addition, centriole duplication, impaired in many primary human microcephaly syndromes (Kaindl et al., 2010), is not affected in \textit{dAnkle2} mutants (Figures S5F and S5G). Hence, loss of \textit{dAnkle2} causes a severe reduction in neuroblast number but does not seem to affect asymmetric division or centriole number.

To assess proliferation in the CNS, we induced mitotic clones of \textit{dAnkle2} in the brain and labeled them with Bromodeoxyuridine (BrdU)(Figures 6P–6R). As shown in Figure 6R, BrdU incorporation is strongly reduced in mutant clones when
Figure 5. Mutations in CRX Cause Bull’s Eye Maculopathy

(A) Pedigree of the family of patient 5 (red arrow) with multiple individuals with bull’s eye maculopathy. The S150X mutation in CRX was identified in eight family members. DNA was not available for family members for whom screening results are not indicated.

(B–D) Clinical phenotypes of patient 5. (B–B’) Fundus photography show fine granularity in the outer retina and speckled glistening deposits arranged in a ring around the macula. Peripheral fundi appear unaffected. (C–C’) Autofluorescence images reveal a bull’s eye phenotype with hypofluorescent macula surrounded by a hyperautofluorescent ring, suggesting a continuously atrophic macular area. (D–D’) Optical coherence tomography shows central loss of the outer nuclear layer, ellipsoid line, external limiting membrane, and retinal pigment epithelium atrophy corresponding to area of hypoautofluorescence in (C–C’).

(E) ERG of the proband: Electroretinographic traces showed implicit time delay and amplitude reduction in both scotopic and especially photopic responses in keeping with generalized cone-rod dysfunction.

(F) Structure of CRX protein and mutations in patients 3–5.

(G) ERG on fly oc mutants. Blue arrows indicate on transient in ERG. On transients are lost in 7-day-old flies. The orange line indicates the amplitude of ERG.
compared to wild-type clones, indicating that cell proliferation is severely impaired. In addition, the mutant clones (Figure 6Q) that contain a neuroblast and its progeny, the ganglion mother cells and neurons, contain many fewer cells than wild-type clones (Figure 6P). Finally, we observe a dramatic increase in apoptotic cells marked by TUNEL in the dAnkle2 mutant brain lobes (Figure 6Q).
(Figures 6S, 6T, and 6V). This cell death is rescued by the expression of the human cDNA encoding ANKLE2 (Figures 6U and 6V). Therefore, defects in proliferation and excessive apoptosis are both contributing to the loss of CNS cells in dAnkle2.

**DISCUSSION**

Here we describe the generation of a large set of chemically induced lethal mutations on the Drosophila X chromosome that were screened for predominantly neurological phenotypes in adult mosaic flies. The mutations were assigned to complementation groups, mapped, and sequenced to associate as many genes as possible with specific phenotypes. We identified and rescued the lethality associated with mutations in 165 genes using a variety of mapping and sequencing methods. These mutations are available through the Bloomington Drosophila Stock Center and provide a valuable resource to study the function of human genes in Drosophila especially since 93% of the genes are evolutionarily conserved in human.

This mutant collection contains 21 genes associated with human diseases for which no mutations were previously available. The fly mutants thus enable the study of the basic molecular mechanism of 26 human diseases, including Leigh syndrome (CG14786/LRPPRC, l(1)G0334/PDHA1, and sicily/NDUFAF6), congenital disorders of glycosylation (CG1597/MOGS, and CG3149/RFT1), Usher syndrome (Aats-his/HARS), Friedreich ataxia (frh/FXN), and amyotrophic lateral sclerosis (ubqn/UBQLN2). Based on the gene list from the Drosophila screen, we explored a database of 1,929 human exomes from a Mendelian disease resource of patients with rare diseases. We examined the personal genomes for rare variants of the fly homologs and prioritized a subset of human rare variant alleles for segregation analysis. We report six families with distinct diseases in which the variants segregate and are likely responsible for causing the associated Mendelian disease.

The approach described here provides a valuable resource to study the function of many disease genes in different tissues. We propose that the screen strategy be expanded to the autosomes, and a number of guiding principles should be considered based on this study. First, the use of low concentrations of EMS is important as it minimizes the number of second site lethal and visible mutations (Haefterman et al., 2014). Second, screening for lethal mutations has major advantages as 93% of the isolated genes that are essential for viability are conserved, whereas only 48% of all Drosophila genes have evolutionarily conserved human homologs. Third, the isolation of lethal mutations also greatly facilitates genetic mapping. Fourth, screening for many different phenotypes casts a broader net and permits isolation of mutations in many different genes, a strategy that is also used in mice (White et al., 2013). Fifth, analyzing different phenotypes revealed that mutations in the majority of the genes cause more than one phenotype, consistent with extensive pleiotropy.

Comparison of the gene list identified from our EMS screen and several RNAi screens have shown that these approaches reveal very distinct sets of genes. There are multiple reasons that may lead to this difference. For example, since our screen was aimed at identifying mutations that cause lethality, we have not screened for genes that are nonessential. Thus, a number of genes that are nonessential but cause morphological defects are missed in our screen. On the other hand, RNAi may not be efficient or cause off-targeting effects (Green et al., 2014; Mohr, 2014). Regardless of the methods that are being used, rescue experiments and independent validation are critical to determine that the phenotype one observes is due to loss of the gene of interest when performing a genetic screen.

It is interesting to note that from our screen, essential fly genes with two or more homologs in humans have a significantly higher likelihood of being associated with Mendelian diseases than those that only have a single human homolog (Figure 3). This suggests that gene duplications of essential genes and subsequent evolutionary divergence may lead to genes that are partially redundant and more likely to be disease associated. Hence, when analyzing human exomes, it would seem more productive to start with homologs of evolutionarily conserved essential Drosophila genes that have two or more human homologs. In addition to these relationships to Mendelian traits, 17% (26/153) of the fly genes that have human homologs have been identified in GWAS (genome-wide association studies) for neurological disorders (Table S5). Hence, the collection of mutations described here may permit us to study genes for complex traits.

We uncovered a genetic basis in a few cases for which the gene was previously known. For example, the study of DNMT revealed previously studied phenotypes associated with mutations in the gene (CMT, Figure S4). In another case we observed that mutations in a gene caused unexpected phenotypes. Indeed, we identified three families with bull’s eye maculopathy, a condition that is much milder and with a later age of onset than conditions typically associated with CRX truncations such as Leber congenital amaurosis (leading to blindness before a year of life) and cone rod dystrophy (a condition with onset in the first or second decade). Interestingly, other truncating alleles have been reported both N- and C-terminally to the OTX transcription factor domain in patients with these severe phenotypes. Therefore, while CRX mutations can produce variable phenotypes (Huang et al., 2012), bull’s eye maculopathy has not been associated with deleterious CRX variants. Our data suggest that some symptoms may manifest at older ages, and the phenotypic spectrum of CRX mutations includes late-onset mild retinopathy.

We identified deleterious alleles in ANKLE2 in two individuals in a family affected by severe microcephaly. In flies, we observed severe defects in neuroblast proliferation and excessive apoptosis in the third instar larval brain of dAnkle2 mutants. This knowledge, combined with the observation that expression of human ANKLE2 in dAnkle2 mutants rescues lethality, brain size, and apoptosis, provide strong evidence that ANKLE2 is responsible for the microcephaly in the family. Moreover, ANKLE2 has been shown to physically and genetically interact with VRK1 in C. elegans and vertebrates (Asencio et al., 2012), and loss of fly VRK1 (also known as balchen (ball) or nhk-1 in flies) also causes a small brain phenotype in third instar larvae (Cullen et al., 2005). It is therefore interesting to note that mutations in VRK1 also cause microcephaly in patients (Figure S5H) (Gonzaga-Jauregui et al., 2013).
The pattern of brain abnormalities and microcephaly in our patient with ANKLE2 mutations is somewhat similar to patients with autosomal recessive CLP1 mutations. CLP1 encodes an RNA kinase involved in tRNA splicing (Karaca et al., 2014; Schaffer et al., 2014). The Clp1 homozygous kinase-dead mouse exhibits microcephaly that worsens with age due to apoptosis. Hence, apoptosis may be a common denominator in these forms of microcephaly.

Phenotypic information of Drosophila mutants allows researchers to understand the potential in vivo function of their human homologs. The cases of oc/CRX and dAnkle2/ANKLE2 are examples in which some direct phenotypic comparisons are possible between the fly mutant and human conditions. However, one of the major drawbacks of comparing phenotypes in different species is that a comparison between different tissues and organs is not always obvious. How do we relate wing vein defects or a rough eye with the phenotypes observed in human genetic diseases? Numerous strategies have been outlined by Lehner (Lehner, 2013) and one of the most compelling strategies is based on orthologous phenotypes or phenologs (McGary et al., 2010). Genes tend to work in evolutionarily conserved pathways, allowing the direct transfer from genotype-phenotype relations between species. For example, mutations in a subset of genes that function in mitochondrial quality control cause a high incidence of muscle mitochondrial defects in adult flies and Parkinson disease (PD) in humans (Jaiswal et al., 2012), suggesting that new genes that affect muscle mitochondria in adult flies are good candidates for PD. Indeed, it may well be that phenotypic similarities between fly and man will be the exception rather than the rule. Regardless, we provide evidence that the use of unbiased screens in the fly and the resulting genetic resources will provide opportunities to prioritize human exome variants and to explore the underlying function of these and many other disease-causing genes in vivo.

EXPERIMENTAL PROCEDURES

Fly Strains
The strains used in this study including the mutations and duplications and deletion strains used for mapping are described in Flybase (Marygold et al., 2013) (see also Extended Experimental Procedures).

Isogenization and Mutagenesis
Isogenization of y w FRT19A chromosome was performed using standard genetic crosses. Mutagenesis was performed by feeding isogenized y w FRT19A iso males with sucrose solution containing a low concentration (7.5–10 mM) of EMS as described (Boekel, 2008). After recovery from mutagenesis, these males were mated en masse with Df(1)A27/FM7c Kr > GFP virgin females for 3 days. In the F1 generation, y w mut FRT19A/FM7c Kr > GFP virgin females were crossed with FM7c Kr > GFP males to establish independent balanced stocks. A total of 5,859 lines carried lethal mutations and the remaining stocks were discarded.

Complementation and Mapping
Lines that exhibited a strong morphological and/or ERG phenotype were subjected initially to duplication mapping. Subsequently, lines that were rescued by the same duplication and exhibit similar phenotypes were crossed inter se to establish complementation groups based on lethality. Complementation groups were further fine mapped using deficiencies that cover the region of interest.

Gene Identification
When a complementation group was mapped to a small region (~30–300 kb, varies depending on available resources), we searched for publically available lethal mutations that map to the same region using FlyBase (Marygold et al., 2013). We performed complementation tests using >1 mutant allele when possible. For complementation groups that complemented all available lethal mutations in the region, we performed Sanger sequencing using standard methods. To expedite gene identification we also used illumina-based whole-genome sequencing technology (Haeffterman et al., 2014).

Ethics Statement
Informed consent was obtained prior to participation from all subjects or parents of recruited subjects under an Institutional Review Board approved protocol at BCM.

Study Subjects
The analysis of 1,929 exomes from BHCMG described was performed in a database from the WES of over 160 separate phenotypic cohorts. The sequencing data included family-based studies in which both affected and unaffected family members were sequenced, single individuals with unique phenotypes, as well as larger cohorts of up to 50–60 cases with the same phenotype. Selection of subjects was performed by a phenotypic review committee based on the likelihood of the Mendelian inheritance for the disease phenotype.

Whole-Exome Capture, Sequencing and Data Analysis
All of the subjects enrolled in the BHCMG underwent WES using methods previously described (Lupski et al., 2013)(Extended Experimental Procedures). Produced sequence reads were mapped and aligned to the GRCh37 (hg19) human genome reference assembly using the HGSC Mercury analysis pipeline (http://www.tinyurl.com/HGSC-Mercury). Variants were determined and called using the Atlas2 suite to produce a variant call file (VCF). High-quality variants were annotated using an in-house developed suite of annotation tools (Bainbridge et al., 2011a).

ANKLE2 Construct and Transgenesis
Human ANKLE2 cDNA was cloned into pUASTattB (Bischof et al., 2007) tagged vectors (N-terminal FLAG) using In-Fusion HD Cloning Kit (Clontech) and vector was linearized with NotI and XhoI. The construct was inserted in VK33 (Venken et al., 2006).

ACCESSION NUMBERS

The dbGAP accession number for the data reported in this paper is phs000711.v1.p1. Additional details are available at http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_ids=phs000711.v1.p1. Biosample IDs for patient 1 (BAB3655), patient 2 (BAB3659), patient 6 (LR06-300a1), and patient 6 family data (LR06-300a2, LR06-300f, LR06-300m) are in Table S6 and data for these individuals is available at this link: http://www.ncbi.nlm.nih.gov/sra?db=sra&DbFrom=bioproject&Cmd=Link&LinkName=bioproject_sra&LinkReadableName=SRA&ordinalpos=1&IdsFromResult=237879

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Results, Extended Experimental Procedures, five figures and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2014.09.002.

AUTHOR CONTRIBUTIONS

REFERENCES


EXTENDED RESULTS

Clinical Case Histories

Patient 1- G358R Variant in DNM2
Patient 1, the proband in Figure S4A, is a 14 year old female who presented with hand tremor, calf cramps, paresthesias of the lower limbs, and difficulty with heel walking at age 12. Patient 1 is a member of a Turkish family with three generations of neuropathy. Her first neurological exam showed distal weakness of all limbs with more prominent weakness in the lower limbs. Nerve conduction studies showed low amplitude and velocities of the median nerve were normal (39 m/s). A sural nerve biopsy revealed rare onion bulbs. The patient was noted by WES to have a heterozygous disease-causing mutation in DNM2 (DNM2:NM_001190716:exon8:c. G1072A:p.G358R). The heterozygous G358R mutation in the DNM2 gene cosegregated with the CMT phenotype in the family in all six affected individuals who were genotyped.

Patient 2- E341K Variant in DNM2
Patient 2, the proband in Figure S4B, is a now 88 year old male who developed weakness in lower extremities starting at age 40 years. His mother who is deceased was also reportedly affected but had not been formally evaluated. The patient remains ambulatory. He has no living relatives. The patient has an E341K mutation in DNM2 (DNM2:NM_001190716:exon8:c.G1021A:p.E341K). In addition to the DNM2, WES revealed a variation in LRSAM1 (c.G334A, p.E112K), a novel mutation altering an amino acid in the fourth leucine-rich repeat region of the protein. Both mutations were confirmed in the proband by Sanger sequencing but no other relatives were alive.

DNM2 Case Summary
Two cases of CMT type 2 were found to have DNM2 mutations. Their phenotypes were consistent with those reported for CMT associated with DNM2 mutations. DNM2 is associated with CMT type 2 and dominant intermediate CMT (Züchner et al., 2005). In patient 1, the G358R mutation has been previously associated with CMT type 2 (Gallardo et al., 2008). For patient 2, a mutation was found in the same domain, but another mutation in another CMT locus (LRSAM1) was also noted. Either one or a combination of both genes may cause CMT in this family.

Patient 3- Y221fs Variant in CRX
Patient 3 is an individual of European descent with visual symptoms at age 61 years. Noted to have bull’s eye maculopathy. No other affected relatives. A CRX frameshift allele was noted (CRX:NM_000554:exon4:c.661 delT:p.Y221fs).

Patient 4- D219fs Variant in CRX
Patient 4 is an individual of Asian Indian descent who presented with visual symptoms at age 26 years. Noted to have bull’s eye maculopathy. No other affected relatives. A CRX frameshift allele was noted (CRX:NM_000554:exon4:c.657 delC:p.D219fs).

Patient 5- S150X Variant in CRX
Patient 5 is the proband in Figure 5A, who presented with visual symptoms at age 43 years. Patient 5 is a member of a large family with Spanish heritage (from the Dominican Republic) with a dominant mode of inheritance. A CRX nonsense allele was noted in the proband (CRX:NM_000554:exon4:c.C449G:p.S150X). In this family, the S150X mutation segregates with the phenotype in the seven affected individuals tested, with ages of onset that range from 28 years to 63 years. The unaffected mother also carried the S150X variant consistent with incomplete penetrance.

CRX Summary
Three individuals with bull’s eye maculopathy were found to have truncating CRX alleles. CRX has been associated with a range of early-onset retinal phenotypes including cone-rod dystrophy, (Kitiratschky et al., 2008), Leber's congenital amaurosis (Freund et al., 1998), and retinitis pigmentosa (Huang et al., 2012). While parents carrying the same alleles have been noted to be without visual impairment, presumably in early adulthood (Freund et al., 1998; Silva et al., 2000), the late onset bull’s eye maculopathy has never been noted in association with CRX.

Patient 6- Compound Heterozygous Variants in ANKLE2
Patient 6 (patient LR06-300a1 in Dobyns database) is a boy of Mexican descent with a birth weight of 2.67 kg (3rd percentile) and a very small head circumference. Examination demonstrated severe microcephaly with low sloping forehead, ptosis, small jaw, multiple hyper- and hypopigmented macules over all areas of his body, and spastic quadriplegia. During his first year of life, he had unexplained anemia, glaucoma, and surgery for ptosis and undescended testes. At 3 years, he had onset of seizures consisting of multiple staring episodes with a few episodes of facial twitching. When evaluated at 5.5 years (Figures 6C–6G), his weight was 10.7 kg (~4 standard deviations, SD), length 83.8 cm (~6 SD) and Fronto-occipito circumference (FOC) of 38.2 cm (~9 SD). He was awake and had good eye contact, symmetric movements, but severe spastic quadriplegia, adducted thumbs and flexion contractures at the knees. He had severe microcephaly with low sloping forehead, normal ears, bilateral ptosis, telecanthus, open mouth with drooling, prominent vertebral bodies in the midthoracic region, and unchanging hyper- and hypopigmented macules.

Brain MRI in the newborn period demonstrated a low forehead, several scalp ruggae, and mildly enlarged extra-axial space with a wide open communication between the posterior lateral ventricles and the mesial extra-axial space. Other changes included a markedly simplified gyral pattern, mildly thickened cortex, small frontal horns of the lateral ventricles with mildly enlarged posterior horns of the lateral ventricles, and agenesis of the corpus callosum. The brainstem and cerebellum appeared relatively normal.
A younger sister born a year later had severe microcephaly, spasticity, and similar hyper- and hypopigmented macules over all areas of her body. She died in the first few weeks of life from cardiac failure associated with poor contractility, although the basis for this was not known.

Whole-exome sequencing was performed on the proband, his affected sister, and both parents. Homozygous and compound heterozygous variants were prioritized based on segregation in the family and then by expression in the nervous system. This led to four candidate genes which met Mendelian expectation and were expressed in the CNS. Table S4 shows the variants with their scores and predictions from the phylop, SIFT, Polyphen2, likelihood ratio test (LRT), and MutationTaster algorithms on dbNSFP (Liu et al., 2011). The ANKLE2 variants noted in the proband (ANKLE2:NM_015114:exon11:c.C2344T:p.Q782X; and NM_015114: exon10:c.C1717G:p.L573V) were prioritized for further study.

EXTENDED EXPERIMENTAL PROCEDURES

Fly Strains
We used the following *Drosophila melanogaster* strains in this study.

**Mutagenesis and Phenotypic Analysis**
y w P[neoFRT]19A
Df(1)JA27/FM7c Kr-GAL4, UAS-GFP
w sn P[neoFRT]19A; Ubx-FLP (Yamamoto et al., 2012)
Tub-Gal80 hsFLP FRT19A; Act-Gal4, UAS-GFP/CyO
cl Ubi-GFP FRT19A/ Dp(1;Y)y+ v+; Ubx-FLP (ll chr)
c(1) P[neoFRT]19A/ Dp(1;Y)y+ v+; ey-FLP (Call et al., 2007) (gift from Drs. John Olson and Utpal Banerjee, UCLA)
c(1) is a recessive cell lethal mutation that is caused by a P-element transposon insertion inRPIL215, the major subunit of RNA polymerase II). P[neoFRT]19A and Kr-GAL4, UAS-GFP are abbreviated as FRT19A and Kr>GFP respectively.

**Duplication Mapping**
Df(1)1srw, N(II) ras*/tw/ Dp(1;1)y+ 67 g19.1/C(1)DX, y1 f1 (Dp901),
Dp(1)fR, y1/1 f1 dor/ Dp761), Df(1)64c18, g1 sqD/Lp(1;2)yw/C(1)DX, y1 w1 f1 (Dp936)
Df(1)hdc81, w118/C(1)DX, y1 f1; Dp(1;2)4FRdup/+ (Dp5594)
Df(1)JC70/Dp(1;Y)ydX5, y1/C(1)M5 (Dp5279)
Df(1)1ct-J4, In(1)1d49, f1/C(1)DX, y1 w1 f1; Dp(1;3)sn1at/Y+ (Dp948)
winscp/Dp(1;Y)BSC174/C(1)DX, y1 w1 f1 (Dp(1;Y)BSC174) (Gift from Dr. Kevin Cook, Indiana University)
Dp(1;Y)y619, y1 B9/w1 oc9/C(1)DX, y1 f1 (Dp5678)
y1 neo22, v1 f1/Dp(1;Y)FF1, y1/C(1)DX, y1 w1 f1 (Dp5292)
Df(1;v-L15, y1/C(1)DX, y1 w1 f1; Dp(1;2)v/+75d/+ (Dp929)
Df(1;v-N48, f1/Dp(1;Y)yv+/C(1)DX, y1 f1 (Dp3560)
Dp(1;Y)BSC1, y1 w1/C(1)DX, y1 w1 f1 (Dp5277)
Dp(1;Y)BSC31, y1 P(3;RS5+3.3')BSC27, B9/1Df(1)ED2665,
w118 P(3;RS5+3.3')ED2665/C(1)RA, In(1)sc7, In(1)sc8, l(1)1Ac1, sc7 sc8 (Dp33250)
Dp(1;Y)BSC223, y1 P(3;RS5+3.3')BSC16, B9/1Df(1)ED7344
w118 P(3;RS5+3.3')ED7344/C(1)RA, In(1)sc7, In(1)sc8, l(1)1Ac1, sc7 sc8 (Dp33244)
Df(1;f1119, f1/C(1)DX, y1 w1 f1; Dp(1;4)y1+ (Dp5273)
Dp(1;Y)W73, y11B1, f1, B9/C(1)DX, y1 f1v y1 bazEH171 (Dp1537)
Df(1)osUE69/C(1)DX, y1 f1/Dp(1;Y)W39, y1+ (Dp1538)
Dp(1;Y)BSC129, y1 P(3;RS5+3.3')BSC22, B9/1Df(1)ED7441
w118 P(3;RS5+3.3')ED7441/C(1)RA, In(1)sc7, In(1)sc8, l(1)1Ac1, sc7 sc8 (Dp30450)
Df(1)R20, y1/C(1)DX, y1 w1 f1/Dp(1;Y)yv+mal+ (Dp3033)
B9/C(1)DX, y1 w1 f1

**Deficiency Mapping and Complementation Test**
Lines that carry deficiencies or lethal mutations in specific regions of interest were identified using Cytosearch (http://flybase.org/static_pages/cytosearch/cytosearch15.html) in FlyBase (Marygold et al., 2013) and publically available lines were obtained from BDSC. Information on the specific lines used for mapping of each complementation group can be obtained upon request.

**Evaluation of Isogenized y w FRT19A Lines**
Isogenization of y w FRT19A chromosome was performed using standard genetic crosses. We established 10 independent lines and selected one line (line F1) as the starter line for mutagenesis. We examined the external structure under light microscope to confirm
Phenotypic Analysis of Morphological Defects in Mutant Clones

To induce homozygous mutant clones of recessive lethal mutations obtained, we collected virgin females from each y w mut* FRT19A/ FM7c K > GFP strain and crossed them with two different FLP lines. To generate clones in the thorax and wing, virgin females were crossed with w sn FRT19A; Ubx-FLP males and we screened y w mut* FRT19A/ w sn FRT19A; Ubx-FLP/+ progeny for morphological defects (Figures 1 and S2). Homozygous mutant tissues were marked by y' sn' bristles, heterozygous tissues were marked by y' sn bristles, and homozygous wild-type bristles were marked by y+ sn' bristles. Since homozygous mutant and wild-type cells are progeny of the same mitotic division, the size of the homozygous mutant clones relative to homozygous wild-type clones should be similar if the mutation does not affect cell division or cell survival. Flies that comprise mostly homozygous wild-type and heterozygous tissue were annotated as “cell lethal.” To generate clones in the eye and head, virgin females were crossed with cl(1)* FRT19A/ Dp(1;Y)y+ v+; ey-FLP males and we screened y w mut* FRT19A/ cl(1)* FRT19A; ey-FLP/+ progeny for morphological defects. Homozygous mutant clones were marked by w' and heterozygous clones were marked by w”. Homozygous wild-type cells were eliminated by the recessive cell lethal mutation (cl(1)*) to give the mutant clones a growth advantage. Morphological defects were documented and recorded in a database that is publicly accessible (http://flypush.imgen.bcm.tmc.edu/ bellenx screendata/mutantsandphenotypes.xlsx).

ERG Analysis of Mutant Clones

y w mut* FRT19A/ FM7c K > GFP virgins were crossed with cl(1)* FRT19A/ Dp(1;Y)y+ v+; ey-FLP males to obtain y w mut* FRT19A/ cl(1)* FRT19A; ey-FLP/+ flies. Flies were aged for 3–4 weeks at room temperature under a normal light-dark cycle and then ERG was recorded. ERG recordings were performed as described earlier (Xiong et al., 2012).

Duplication Rescue and Rough Mapping Using Large Duplications

Lines that exhibited a strong morphological and/or ERG phenotype were subjected to duplication mapping. Virgin females from the mutant lines were crossed to males carrying different X chromosome duplications (Cook et al., 2010). Progenies were scored to determine whether the duplication rescued the lethality of the mutation. The duplication mapping was performed in 3 rounds. Round 1: Dp901, Dp936, Dp5279, Dp5678, Dp5292, Dp3560, Dp5596, Dp1537, Dp1538, Dp3033. Round 2: Dp761, Dp5594, Dp948, Dp8-28-8A, Dp929, Dp5459, Dp26276, Dp5273. Round 3: Dp33250, Dp33244, Dp30450. Rescued males were crossed to a stock that carries a compound X chromosome (C(1)DX) or to the original mutant stock to establish stocks that stably produce rescued male flies. For Dp5459, this was not possible due to technical reasons.

Complementation Testing

Lines that were rescued by the same duplication and exhibit similar phenotypes were crossed inter se to establish complementation groups based on lethality. We did not perform complementation tests for mutations rescued by Dp5594, Dp948, Dp929, and Dp5273 since the X chromosome duplication did not possess any useful visible markers. In addition, we did not perform complementation tests for mutants rescued by Dp5459 since we were not able to obtain lines that stably produce rescued male flies. In cases where mutations with different phenotypes were fine mapped to similar regions, we performed complementation tests between these lines.

Fine Mapping Using Deletions and P[acman] Duplications

Complementation groups were further fine mapped using deficiencies that cover the region of interest. We selected ~5 deficiencies to further subdivide the rough mapped regions into smaller regions. Most of the deficiencies we selected were molecularly defined (Cook et al., 2012; Parks et al., 2004). Whenever a molecularly defined deficiency was not available, we selected cytologically mapped deficiencies to cover the gap regions (Lindsley and Zimm, 1992). Rescued males from the mutant lines were crossed to virgin females that carried X chromosome deficiencies. In addition, we occasionally used strains carrying BACs that cover a portion (~80 kb) of the X chromosome generated using the P[acman] technology (P[acman] mapping) (Venken et al., 2010). Females from mutant strains were crossed with males that carry the P[acman] duplication and we scored the rescue of lethality in the subsequent generation.

Gene Identification

When a complementation group was mapped to a small region (~30–300 kb, varies depending on available resources), we searched for publically available lethal mutations that map to the same region using FlyBase (Marygold et al., 2013). We performed complementation tests using >1 mutant allele when possible. For complementation groups that complemented all available lethal mutations in the region, we performed Sanger sequencing using standard methods. To expedite gene identification we also used Illumina-based whole-genome sequencing technology (Haelterman et al., 2014).
Gene Ontology Analysis

The molecular functions (MF) and biological processes (BF) annotated for each gene were retrieved using the online tool DAVID (the Database for Annotation, Visualization and Integrated Discovery) with Flybase ID as the identifier (Dennis et al., 2003). MF and BF terms for genes that are not annotated in DAVID were manually extracted through FlyBase. MF and BF were further classified manually. MF and BF associated with individual genes can be downloaded from the following website: http://flypush.imgen.bcm.tmc.edu/bellenxscreendata/go.xlsx

Identification of Human Homologs of Fly Genes and Their Association with OMIM Diseases

The human homologs of the fly genes were identified using HGNC Comparison of Orthology Predictions (HCOP) search tool (http://www.genenames.org/cgi-bin/hcop) (Wright et al., 2005). Once we assembled a human homolog list for the genes identified from our screen and for the whole fly genome based on data downloaded from FlyBase (Marygold et al., 2013), we searched human diseases that have been associated with each human homolog based on data downloaded from OMIM (http://www.omim.org). Estimation of the number of genes in the fly genome that are lethal versus viable was based on the following criteria. The number of essential loci in the fly genome has been repeatedly been estimated to be ~5,000 based on saturation mutagenesis experiments (Benos et al., 2001). Currently, 1,934 loci have been associated with a lethal mutation (excluding uncharacterized transposon insertions and RNAi-based phenotypes) according to FlyBase. This is ~40% of all essential loci based on the predicted total number of essential genes. The raw data we used to generate the graphs and tables in Figure 3 can be found in Table S3 (genes from the screen) or can be downloaded from the following website (for all genes in the fly genome):
http://flypush.imgen.bcm.tmc.edu/bellenxscreendata/wholeflygenome.xlsx

Imaging of Larval Brains

Larval brains (Figures 6I–6K) were dissected in PBS from similar sized late third instar larvae and fixed (3.7% formaldehyde in PBS) for 20 min and washed in PBS. DIC images of brains were taken by Zeiss microscope (Axio Imager-Z2) equipped with the AxioCam MRm digital camera. Images are acquired using image acquisition software Zen and processed by Adobe Photoshop.

Immunohistochemistry

For immunostaining of the fly PNS in the notum (Figure 6H), white fly pupae (0 hr after puparium formation) were aged at 25°C for 24–27 hr before dissection. For larval brain immunostainings, wandering third instar larvae brains were dissected, fixed in 3.7% formaldehyde in PBS for 20 min, and washed in PBS with 0.3% Triton X-100 (PBT). Fixed larvae were blocked in 1x PBS containing 5% normal goat serum and 0.3% Triton X-100 (PBTS) overnight at 4°C. Samples were incubated in secondary antibody diluted in PBTS overnight at 4°C. Samples are washed in PBT, incubated in secondary antibody diluted in PBT for two hrs, and then washed in PBT prior to mounting. Primary antibodies were used at the following dilutions: rat anti-Elav 1:500 (DSHB) (O’Neill et al., 1994), mouse anti-Cut 1:500 (DSHB) (Blochlinger et al., 1990), and chicken anti-GFP 1:1,000 (Abcam), rat anti-Mira 1:250 (Chabu and Doe, 2008), rabbit anti-phospho-Histone 3 (PH3) 1:1,000 (Upstate Biotechnology), mouse anti-α-tubulin 1:5000 (Sigma), and Cnn 1:100 (Heuer et al., 1995). Images were taken by confocal microscope (Zeiss 510 or Zeiss LSM 710) and processed using ImageJ or Imaris (Bitplane).

TUNEL Staining

Wandering 3rd instar larvae brains were dissected, fixed in 100 mM Pipes, 1 mM EGTA, 0.3% Triton X-100, and 1 mM MgSO4 containing 4% formaldehyde for 20 min, and blocked in 1x PBS containing 1% BSA and 0.3% Triton X-100 supplemented with 0.01 M glycine and 0.1% normal serum for 1hr. Fixed brains were treated with 20 µg/ml proteinase K for 2 min, rinsed 4x in 1x PBS containing 0.3% Triton X-100 (PBST), re-fixed for 20 min, rinsed 3x in PBST, equilibrated in TdT Equilibration Buffer (Calbiochem Fluorescein-FragEL Kit) for 30 min, and incubated with TdT enzyme and Fluorescein labeled dNTPs at 37°C for 2hrs. Brains images were acquired using confocal microscope (Zeiss LSM 710) and TUNEL positive cells were quantified using Imaris (Bitplane).

BrdU Incorporation

Tub-Gal80 hsFLP FRT19A; Act-Gal4, UAS-GFP/CyO was crossed to y, w, dAnkle2A FRT19A or y, w, FRT19A (control) to generate clones that are marked by GFP (MARCM (Lee and Luo, 1999), Figures 6P and 6Q). Embryos were collected for 24hrs, aged 12–24 hr, heat shocked at 37°C for 2 hr, and resulting 3rd instar larvae containing MARCM clones were shifted to blue food containing 1 mg/ml BrdU for 4hrs. Brains were dissected, fixed in 100 mM Pipes, 1 mM EGTA, 0.3% Triton X-100, and 1 mM MgSO4 containing 4% formaldehyde for 20 min, and blocked in 1x PBS containing 1% BSA and 0.3% Triton X-100 supplemented with 0.01 M glycine and 0.1% normal serum for 1hr. Fixed brains were treated with 2N HCl for 30 min and blocked in 1x PBS containing 1% BSA and 0.3% Triton X-100 with 0.1% normal serum for 1hr. Samples were incubated with mouse anti-BrdU (DSHB, 1:250), rat anti-Elav (DSHB, 1:250), and rabbit anti-GFP (Invitrogen, 1:1000) overnight at 4°C. Images were acquired with Zeiss LSM 710 or Apotome.2 and analyzed using Imaris (Bitplane).

Whole-Exome Sequencing

Briefly, 1 μg of genomic DNA in 100 μl volume was sheared into fragments of approximately 300–400 base pairs in a Covaris plate with E210 system (Covaris, Inc. Woburn, MA). Genomic DNA samples were constructed into Illumina paired-end precapture libraries.
according to the manufacturer’s protocol (Illumina Multiplexing Sample Prep Guide, 1005361_D) with modifications as described in the BCM-HGSC Illumina Barcoded Paired-End Capture Library Preparation protocol. Libraries were prepared using Beckman robotic workstations (Biorad NXP and FXp models). The complete protocol and oligonucleotide sequences are accessible from the HGSC website (https://hgsc.bcm.edu/sites/default/files/documents/Illumina_Barcoded_Paired-End_Capture_Library_Preparation.pdf). Four precapture libraries were pooled together (approximately 500 ng/sample, 2 μg per pool) and hybridized in solution to the HGSC CORE design (Bainbridge et al., 2011b) (52Mb, NimbleGen) according to the manufacturer’s protocol NimbleGen SeqCap EZ Exome Library SR User’s Guide (Version 2.2) with minor revisions. Captured DNA fragments were sequenced using paired end mode on an Illumina HiSeq 2000 platform (TruSeq SBS Kits, Part no. FC-401-3001) producing 9–10 Gb per sample and achieving an average of 90% of the targeted exome bases covered to a depth of 20× or greater.

Illumina sequence analysis was performed using the HGSC Mercury analysis pipeline (http://www.tinyurl.com/HGSC-Mercury/) that addresses all aspects of data processing and analyses from the initial sequence generation on the instrument to annotated variant calls (SNPs and intraread indels). This pipeline uses .bcl files to then generate sequence reads and base-call confidence values (qualities) using Illumina primary analysis software (CASA). Reads were mapped to the GRCh37 Human reference genome (http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/) using the Burrows-Wheeler aligner (BWA (Li and Durbin, 2009; http://bio-bwa.sourceforge.net/) producing a BAM (Li et al., 2009) (binary alignment/map) file. BAM postprocessing including in/del realignment and quality recalibration is done using a variety of tools (SAMtools, GATK, etc). Variants were determined using the Atlas2 (Challis et al., 2012) suite (Atlas-SNP and Atlas-indel) to call variants and produce a variant call file (VCF) (Danecek et al., 2011). Finally, annotation data are added to the vcf using a suite of annotation tools “Cassandra” (Bainbridge et al., 2011a).

**Sanger Confirmation**

Primers for Sanger confirmation for all variants reported were designed using Primer3 (Untergasser et al., 2012).

**Variant Analysis**

Variants were filtered out for having greater than 1% allele frequency in the 1000 Genomes Project (http://www.1000genomes.org), the Exome Variant Server of NHLBI GO Exome Sequencing Project (http://evs.gs.washington.edu/EVS/), or within the Atherosclerosis Risk in Communities Study (ARIC) (http://drupal.cscc.unc.edu/ARIC).


Figure S1. Flow Chart of the F3 Adult Mosaic Genetic Screen on the X Chromosome of Drosophila, Related to Figure 1
(A and B) The y w FRT19A chromosome was isogenized (A) and male flies were mutagenized (B).
(C and D) The mutagenized X-chromosomes were balanced with FM7c Kr > GFP balancer (C) and strains with X-linked recessive lethal mutations were kept (D).
(E) Mosaic flies with Ubx-FLP were used to screen mutant clones in wing and thorax and with ey-FLP were used to screen mutant clones in head and eye.
(F and G) We assessed morphological and ERGs defects in mosaic flies.
Figure S2. Phenotypic Screening of Morphological and Electrophysiological Defects in Mutant Clones, Related to Figure 1

(A–D) Examples of phenotypes observed in the fly notum. Homozygous wild-type bristles are marked by singed. Homozygous mutant bristles are marked by yellow (encircled by dotted lines). Heterozygous bristles are wild-type for these two markers. (A) Macrochaetae loss. (B) Short bristles. (C) Cell lethal. (D) Depigmentation.

(E–J) Examples of phenotypes observed in wings. The exact clonal boundaries are not obvious since yellow does not show a strong phenotype in the wing. (E) Notching. (F) Ectopic wing margin. (G) Vein loss (arrow) and gain (arrowhead). (H) Ectopic bristles on the wing blade. (I) Wing blistering. (J) Crinkled wings.

(K–S) Examples of phenotypes observed in eyes and heads. Homozygous wild-type cells are eliminated by a recessive cell lethal mutation. Homozygous mutant clones in the eyes are marked by white. Heterozygous clones appear red (white*). (K) Wild-type eye and head clones. (L) Rough eye. (M) Cell lethal. (N) Small eye. (O) Ectopic eye (black arrow). (P) Glossy eye. (Q) Ectopic antenna formation (two left antennae are marked by two black arrows) and overgrowth of the eye and head. (R) Noncell autonomous overgrowth of the eye (marked by a white arrow). (S) Overgrowth of the head cuticle (marked by a white arrow).

(T and U) Gene Ontology (GO) analysis based on (T) molecular functions and (U) biological processes.
Figure S3. Flow Chart of Mapping of X-Linked Recessive Lethal Mutants, Related to Figure 1.
Figure S4. Missense Mutations in DNM2 Associated with Charcot Marie Tooth Disease, Related to Figure 4

(A) Pedigree of the family of patient 1, a 14 year old (red arrow) who was diagnosed CMT neuropathy, demonstrating 13 individuals affected with neuropathy (black indicates clinical neuropathy). Six affected individuals were genotyped and all six carry the G358R allele. Two additional unaffected individuals did not carry the allele.

(B) Pedigree of the family of patient 2 with CMT neuropathy (black indicates clinical neuropathy). This individual (red arrow) was also found to be heterozygous for an E341K allele in DNM2 and a heterozygous variant in LRSAM1 (E112K).

(C) Sural nerve biopsy of control and patient 1 showing rare “onion bulb” structures (red arrows).

(D) Structure of DNM2 protein and position and nature of the mutations in patient 1 and 2.
Figure S5. dAnkle2 Regulates Brain Size, Related to Figure 6

(A) Quantification of control and dAnkle2A larval brain lobe volume from early, mid, and late third instar larvae. ‘**’ indicates p value < 0.01 and ‘*’ indicates p value < 0.05.

(B and C) Neuroblast cells in control (B) and dAnkle2A mutants (C) are marked by Miranda (Mira, green). PH3 (red) marks chromosomes in dividing cells. Quantification of the number of neuroblasts in these brains is shown in Figure 6O.

(D and E) Neuroblasts undergoing mitosis in control (D) and dAnkle2A larval brains (E). Mitotic spindles (α-Tub, red) are oriented toward the polarity axis both in control and dAnkle2A. Mira (green) marks the basal side of asymmetrically dividing neuroblast cells. Condensed chromosomes are marked by PH3 (blue).

(F and G) Centrioles in mitotic neuroblast in larval brains are marked by Cnn (red). Mitotic spindles are marked by α-Tub (green) and DNA is marked by DAPI (blue).

(H) A comparison of clinical features observed in patients carrying variants in VRK1 and ANKLE2.