DOI: 10.1534/genetics.110.115121

Molecular Genetic Analysis of *Chd3* and Polytene Chromosome Region 76B-D in *Drosophila melanogaster*

Monica T. Cooper, Alexander W. Conant and James A. Kennison[1]

Program in Genomics of Differentiation, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health (NIH), Bethesda, Maryland 20892-2785

Manuscript received February 3, 2010
Accepted for publication April 20, 2010

ABSTRACT

The *Drosophila melanogaster Chd3* gene encodes a member of the CHD group of SNF2/RAD54 ATPases. CHD proteins are conserved from yeast to man and many are subunits of chromatin-remodeling complexes that facilitate transcription. Drosophila CHD3 proteins are not found in protein complexes, but as monomers that remodel chromatin *in vitro*. CHD3 colocalize with elongating RNA polymerase II on salivary gland polytene chromosomes. Since the role of *Chd3* in development was unknown, we isolated and characterized the essential genes within the 640-kb region of the third chromosome (polytene chromosome region 76B-D) that includes *Chd3*. We recovered mutations in 24 genes that are essential for zygotic viability. We found that transposon-insertion mutants for 46% of the essential genes are included in the Drosophila Gene Disruption Project collection. None of the essential genes that we identified are in a 200-kb region that includes *Chd3*. We generated a deletion of *Chd3* by targeted gene replacement. This deletion had no effect on either viability or fertility.

THE SNF2/RAD54 family proteins are components of ATP-dependent chromatin-remodeling complexes involved in transcription and/or DNA repair (Eisen *et al.* 1995; Pollard and Peterson 1998). Within the SNF2/RAD54 family, the CHD (chromodomain-helicase-DNA-binding) group all contain CHROMO (chromatin organization modifier) domains, which interact with methylated lysines in the N-terminal tails of histones (Jacobs and Khorasanizadeh 2002; Nielsen *et al.* 2002). The CHD group can be further divided into at least three classes on the basis of the presence or absence of two other protein domains. The CHD1 class proteins have only the SNF2/RAD54 ATPase domain and CHROMO domains. The sole CHD group protein in *Saccharomyces cerevisiae* is a CHD1 class protein (*Sc* CHD1) that is not essential for viability (Woodage *et al.* 1997). In addition to the SNF2/RAD54 and CHROMO domains, the KISMET class proteins also have a BRK (brahma and kismet) domain, while the MI-2 (or CHD3/CHD4) class proteins have one or more PHD (plant homeo domain) zinc fingers near the amino terminus. *Drosophila melanogaster* has a single CHD1 class gene, *Chd1*. While zygotically expressed *Chd1* is not required for viability, maternally encoded *Chd1* gene products are required for early embryogenesis (*i.e.*, *Chd1* mutants are female sterile) (Konev *et al.* 2007; McDaniel *et al.* 2008). The sole KISMET class gene in *D. melanogaster*, *kismet*, is required for zygotic viability and the function of homeotic genes (Daubresse *et al.* 1999). *Mi-2* and *Chd3*, the two MI-2 class genes in *D. melanogaster*, are both within the *Df(3L)kto2* region of the genome. *Mi-2* is required for zygotic viability (Kehle *et al.* 1998). The expression of *Chd3* has been shown to be developmentally regulated (Murawska *et al.* 2008), and CHD3 proteins colocalize with elongating RNA polymerase II on Drosophila salivary gland polytene chromosomes (Murawska *et al.* 2008). What is the role of *Chd3* in development? As part of a genetic investigation of the chromosomal region that spans polytene chromosomal bands 76B1 to 76D5, we have identified 24 essential genes around *Chd3*. Since *Chd3* was not among the essential genes that we identified, we generated a targeted gene replacement allele of *Chd3*. Deletion of the majority of the open reading frame of *Chd3* had no effects on either viability or fertility.

MATERIALS AND METHODS

Flies were raised on a yeast/cornmeal/molasses/Tegosept medium at 25°. All mutations and chromosome aberrations are described in Lindsley and Zimm (1992) or FlyBase (http://flybase.org/) except *Df(3L)A23*, *Df(3L)XS917*, and *Df(3L)kto22*. *Df(3L)A23* and *Df(3L)XS917* were generated in the laboratories of David Stein and Gerry Rubin, respectively and provided by Jürg Müller. We recovered *Df(3L)kto22* as a flanking deletion from *P{lacW}trc*^{S066917a}. Putative deletions were detected by a change in eye color after crossing to the balancer *TMS* (which carries a transposon that expresses the P transposase). Because the *P{lacW}* transposon is still present in *Df(3L)kto22*, we used inverse PCR (Huang *et al.* 2000) to

Supporting information is available online at http://www.genetics.org/cgi/content/full/genetics.110.115121/DC1.

[1]*Corresponding author:* 6 Center Drive, Bldg. 6B, Room 3B331, NIH, Bethesda, MD 20892-2785. E-mail: Jim_Kennison@nih.gov

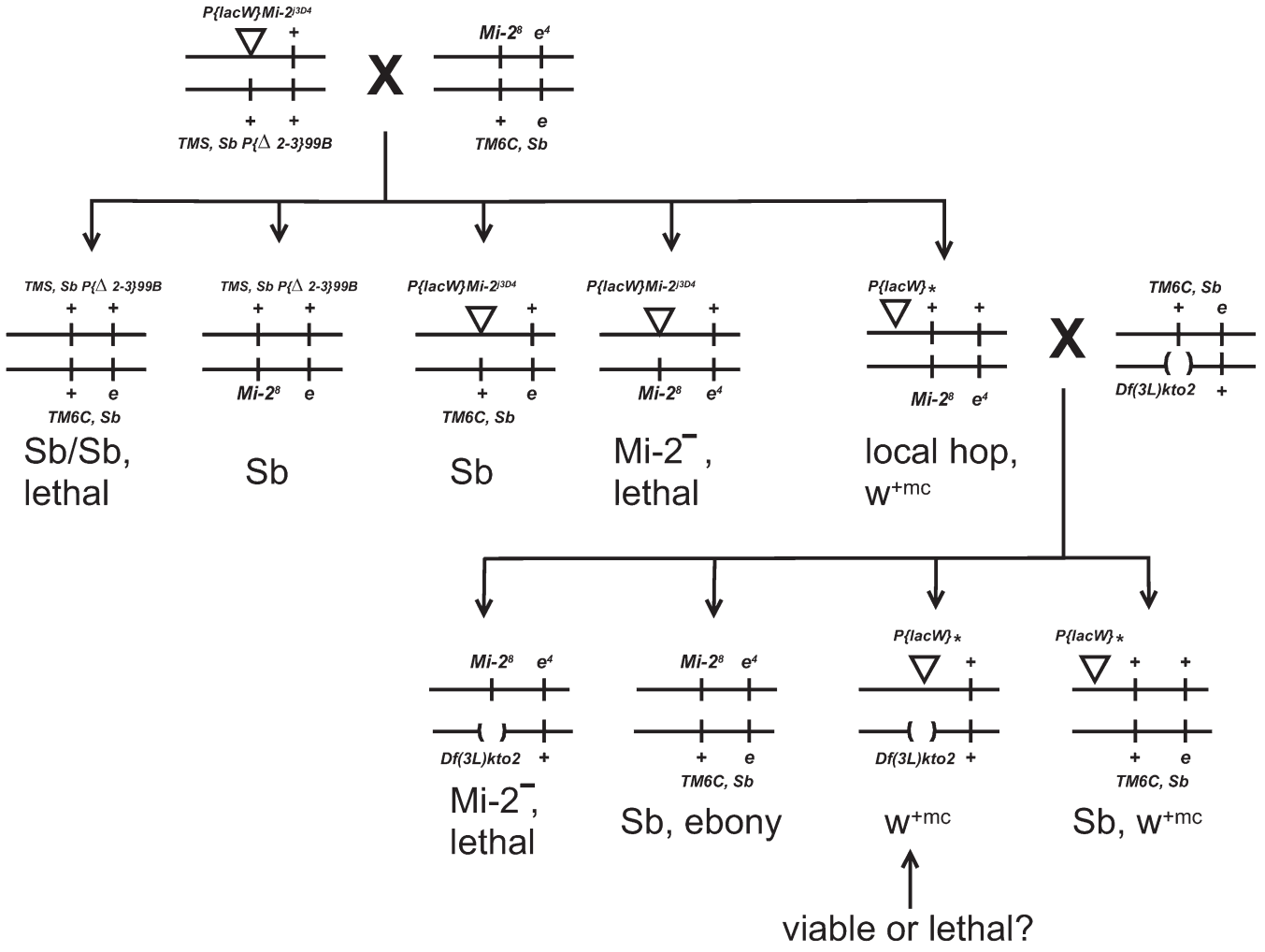

Figure 1.—Crossing scheme for local *P*-element transposition. Transpositions of the *P*-element insertion *P{lacW}Mi-2*^{j3D4} were generated in flies heterozygous to the *TMS* balancer chromosome, which carries a nonmobilizable transgene that produces the *P* transposase. Local transpositions, labeled *P{lacW}**, were recognized as non-*Sb* flies that had the mini-white marker of the *P{lacW}* vector and failed to complement *Df(3L)kto2* for viability. *TM6C, Sb* is the balancer chromosome; *In(3LR)TM6C, cu Sb e ca.*

determine the molecular endpoints of the deletion, which spans from the site of insertion of *P{lacW}trc*^{S066917a} to midway between the 3′ ends of *Papss* and *Rab8*.

Males were fed ethylmethane sulfonate (EMS) as described (Lewis and Bacher 1968; Kennison 1983). The mutagenized males were mated to virgin females and discarded after 4 days; the inseminated females were returned to new cultures for subsequent brooding. Mutagenized males were homozygous for either an unmarked chromosome from the *iso-1* strain (Brizuela *et al.* 1994) or a third chromosome carrying *red*^{1} and *e*^{4}. Both third chromosomes were isogenized prior to mutagenesis. We recovered mutations that failed to complement *Df(3L)kto2* from three different experiments. Following the nomenclature suggested by Lindsley and Zimm (1992), we named these mutations *l(3)76BD*. In the first experiment, single sons with a mutagenized *red*^{1} *e*^{4} third chromosome were mated to three to five *Df(3L)kto2/TM6C* virgins. If few or no progeny carrying both the mutagenized third chromosome and *Df(3L)kto2* were recovered, a balanced stock of *l(3)76BD red*^{1} *e*^{4}*/TM6C* was selected and retested. From 5576 fertile sons, we recovered 124 *l(3)76BD* mutations. In the second and third experiments, balanced lines with a mutagenized third chromosome and *TM6B* were first generated. Only those lines in which few or no flies homozygous for the mutagenized third chromosome survived were subsequently tested by crossing to *Df(3L)kto2/TM6C* virgins. In the second experiment, from ~200 EMS-treated *iso-1* third chromosome lines, we recovered four *l(3)76BD* mutations. In the third experiment, from 1938 EMS-treated *red*^{1} *e*^{4} third chromosome lines (which we estimated from the Poisson distribution to contain 3342 lethal mutations on the third chromosome), we recovered 67 *l(3)76BD* mutations.

We used another approach to try to tag additional genes in this region with *P*-element insertions. During *P*-element transposition, it has been reported that new insertion sites are more frequent than expected near the original insertion site (Tower *et al.* 1993; Zhang and Spradling 1993, 1994; Tower and Kurapati 1994; Matsubayashi and Yamamoto 1998). This is referred to as local *P*-element transposition. We hoped that this would be useful in isolating additional lethal *P*-insertion mutations in this region of the genome. We should point out that this is different from the generation of flanking deletions by exposure of a *P*-transposon insertion to *P*-element transposase, which is also sometimes referred to as local *P*-element transposition. To minimize the number of flanking deletions recovered, we started with *P{lacW}Mi-2*^{j3D4} *Su(Tpl)*^{j3D4}, which is located in the first introns of both the *Mi-2* and *Su(Tpl)* genes. This insertion mutant completely fails to complement

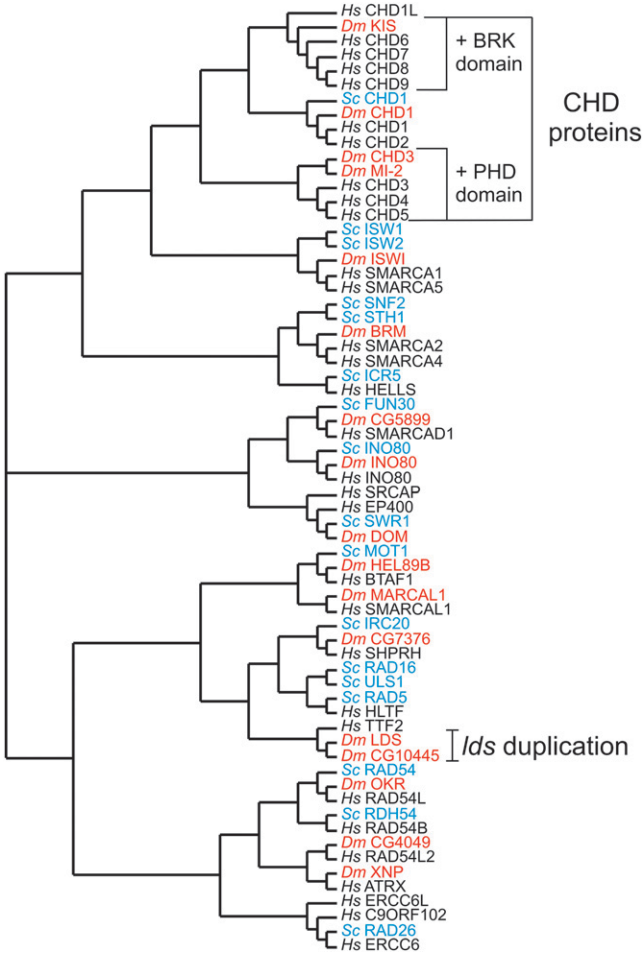

FIGURE 2.—Evolutionary relationships of SNF2/RAD54 family proteins. A cladogram of the SNF2/RAD54 family proteins from *Homo sapiens* (31 *Hs* proteins in black), *Drosophila melanogaster* (17 *Dm* proteins in red), and *Saccharomyces cerevisiae* (17 *Sc* proteins in blue). The CHD group is indicated by the large bracket on the right, with the KISMET class (+ BRK domain) and MI-2 class (+ PHD domain) indicated by the smaller brackets within the CHD group. The *lds* and *CG10445* genes in *D. melanogaster* are also indicated on the right as the *lds* duplication.

Mi-2 alleles, but complements *Su(Tpl)* alleles well (at least 80% of the expected flies eclose). We selected new *P{lacW}* insertions that complement *Mi-2*. The crossing scheme is shown in Figure 1.

To construct the vector for targeted gene replacement, we started with the "ends-out" **arm-GFP** vector constructed for the *mre11* gene replacement (MIN *et al.* 2005). For the downstream flanking DNA fragment, we used the primers 5′-TAC GGCACTAGTCGACGAGTGTAATTGACTCC-3′ and 5′-CATC CACAATGATAGTCTAGGTAGC-3′ to PCR amplify a genomic fragment. We cut the PCR product with *Spe*I and cloned the resulting 3.4-kb fragment into the *Spe*I site in the **arm-GFP** vector. For the upstream flanking DNA, we used the primers 5′-TGACTCGGTACCGACTGTTGACGGTCTTGCCGCTG-3′ and 5′-CGTAGCTGGTCAGCATTACG-3′ to PCR amplify a genomic fragment. We cut the PCR product with *Kpn*I and *Hin*dIII and cloned the resulting 2.2-kb fragment into the *Kpn*I *Hin*dIII sites in the **arm-GFP** vector. A *Kpn*I–*Not*I fragment with both flanking DNA fragments and the *arm-GFP* fragment was cloned into pW30 (GONG and GOLIC 2003). This construct was co-injected with *P* transposase into *white* mutant embryos and stable transformants were selected using the mini-*white* marker in pW30. The targeted replacement was generated and selected as a GFP-positive embryo as described (MIN *et al.* 2005). The targeted replacement was verified by the failure to amplify by PCR *Chd3* genomic sequences in *Chd3′* homozygotes, using the primers CHD3-2 (5′-GGCGATTGGATCTGCCCGCG-3′), CHD3-5 (5′-TCTCTGTGGACGCAGCCTTTC-3′), CHD3-6 (5′-CGGATGTATTGAAGAGCATGC-3′), CHD3-7 (5′-ATCAC AGGGTCCTTTTATTC-3′), CHD3-9 (5′-CGAGGAGCATGAT GACTACAA-3′), CHD3-10 (5′-CTCTTCGTTTTGTCATTT GTC-3′), CHD3-11 (5′-GTACCCAATGGCGGTTATTCC-3′), CHD3-12 (5′-CCCAAGACTCCTTCTCTTCA-3′), and CHD3-19 (5′-CTATGATTGATAACCCGCTGG-3′).

For the evolutionary comparisons, phylograms and cladograms were generated using the ClusalW2, TreeView32, and TreeViewX programs. Only the 11 Drosophila species that can be searched with the BLAT program on the UCSC Genome Browser website (http://genome.ucsc.edu/; KENT 2002) were examined for most of the species comparisons. The two exceptions were *MI-2/CHD3* and *lodestar/CG10445*, for which NCBI BLAST (http://blast.ncbi.nlm.nih.gov/) was used to identify and characterize the *D. willistoni* genes. For the determination of evolutionarily conserved amino acids, we used the EvoPrinter version 1.1 (http://evoprinter.ninds.nih.gov/index11.html; ODENWALD *et al.* 2005). Since the program will only compare a maximum of nine species, we used *D. melanogaster* and the eight most distantly related Drosophila species (*D. erecta*, *D. yakuba*, *D. annanassae*, *D. pseudoobscura*, *D. persimilis*, *D. virilis*, *D. mojavensis*, and *D. grimshawi*) with available BLAT files. Only those amino acid residues that were identical in all nine Drosophila species were counted as evolutionarily conserved. The sequences used to construct the cladogram of the OAK cluster genes are in supporting informaton, File S1.

The following primers were used for the PCR analysis of the 31.4-kb tandem duplication that includes *Gyc76C*: primer 1, 5′-TGGCAACGAACTCTAGGGACT-3′; primer 2, 5′-GCAGA GGAACTACTTAGCAGTC-3′; primer 3, 5′-ACCAACCACCTC CAACTAGAG-3′; and primer 4, 5′-GGGCACGTTTTTAGATG AGCTTC-3′.

RESULTS

A cladogram of the SNF2/RAD54 family proteins from *D. melanogaster*, *S. cerevisiae*, and *Homo sapiens* is shown in Figure 2. There are four CHD proteins in *D. melanogaster*. The MI-2 class of CHD proteins is distinguished by the presence of one or more PHD domains. The two Mi-2 class genes in *D. melanogaster*, *Mi-2* and *Chd3*, are both within the chromosomal region deleted in *Df(3L)kto2*. The ~640 kb of genomic DNA deleted by *Df(3L)kto2* is shown in Figure 3, with the genes annotated by the Drosophila Genome Project. We first present our efforts to identify all of the essential genes within *Df(3L)kto2* and then our efforts to generate and characterize a *Chd3* mutant.

After EMS mutagenesis, we recovered 195 *l(3)76BD* mutations that failed to complement *Df(3L)kto2*. These mutations define 24 complementation groups, which we provisionally designated *l(3)76BDa* through *l(3)76BDx*. These complementation groups and the number of alleles that we recovered for each are shown in Table 1.

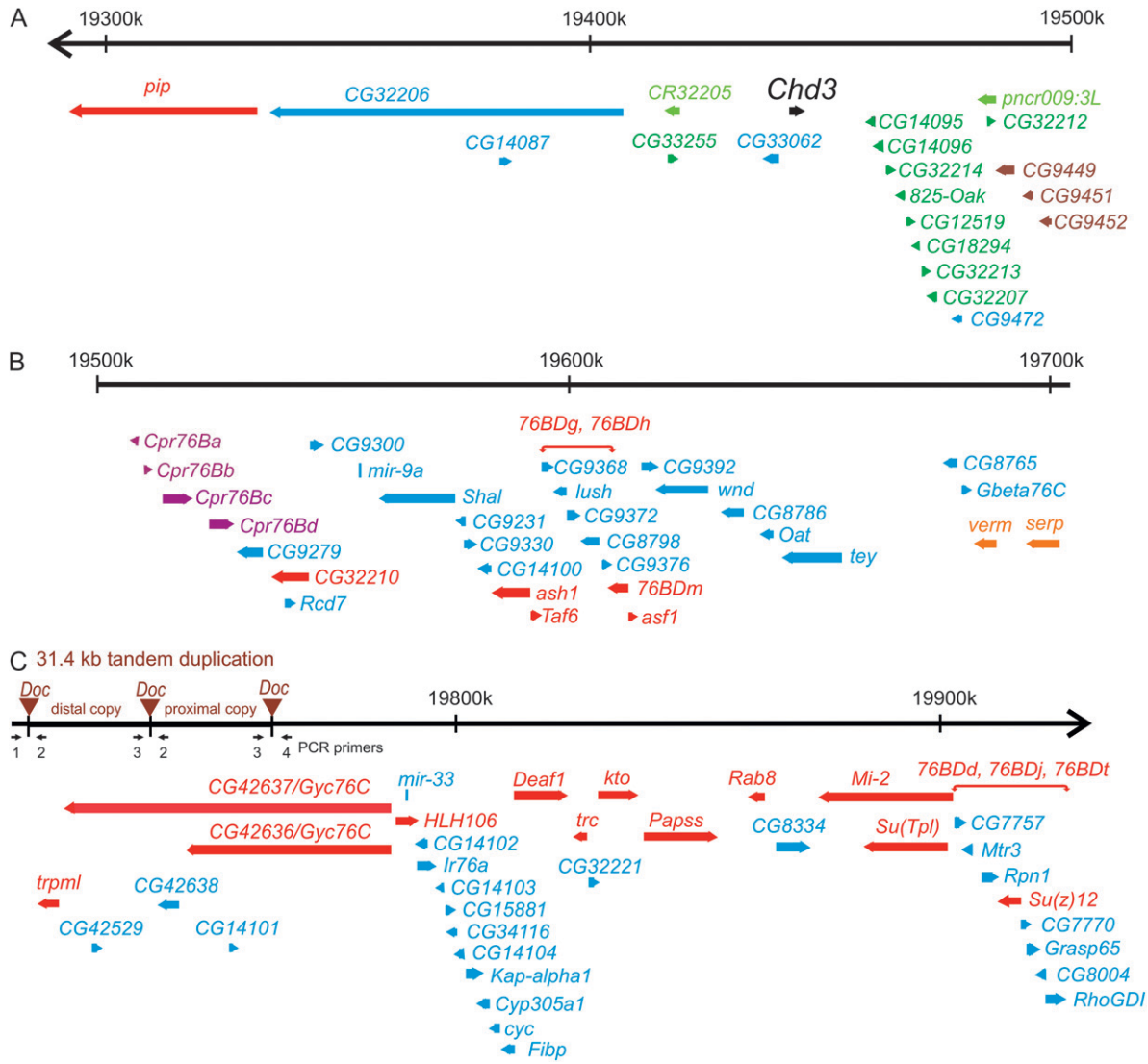

Figure 3.—Molecular map of the genomic region deleted in *Df(3L)kto2*. The ~640 kb of genomic DNA (from 3L: 19291k to 19927k, release 5.23) is broken into three parts (A, B, and C), and is represented by the horizontal black arrows at the top of each part. The annotated transcription units are represented by colored thick horizontal arrows. The *Chd3* transcription unit in 3A is black. The essential transcription units are red and orange. The clusters of transcription units encoding related proteins are dark green (the OAK cluster in 3A), light green (the two duplicated noncoding RNAs in 3A), brown (the acid phosphatase cluster in 3A), purple (the cuticular protein cluster in 3B), and orange (the essential genes *verm* and *serp* in 3B). All other transcription units are blue. The two regions that include the five essential genes (*76BDg*, *76BDh*, *76BDd*, *76BDj*, and *76BDt*) for which the transcription units have not been identified are indicated by red horizontal brackets above the candidate transcription units. The putative 31.4-kb tandem duplication (distal copy and proximal copy) flanked by *Doc* transposable elements in the *iso-1* strain is shown on the genomic DNA at the left of 3C, with the *Doc* elements represented by inverted brown triangles. The locations of the PCR primers (1, 2, 3, and 4) that were used to try to verify the presence of the tandem duplication are indicated below the genomic DNA in 3C.

We recovered an average of 8.1 alleles per complementation group, with only one complementation group [*l(3)76BDp*] represented by a single allele. To further localize our complementation groups, we also crossed representatives of each complementation group to the 26 chromosomal deletions in Figure 4 that overlap *Df(3L)kto2* to differing extents. Thirteen of the deletions (those indicated by the red bars in Figure 4) have molecularly defined breakpoints, which are useful in integrating the genetic and molecular maps.

We also tested all of the putative lethal transposon insertion mutants in this 640-kb region that the Drosophila Gene Disruption Project has made available from the Bloomington Drosophila Stock Center. These transposon insertion mutants are listed in Table 2, and include *P*- (P), piggyBac- (PBac), and Minos- (Mi) transposable element insertions. Four of the transposon insertion mutants complement *Df(3L)kto2* for viability and fertility, indicating that the lethality of the insertion chromosome is not due to disruption of the associated

TABLE 1

Complementation groups represented by the mutations within the *Df(3L)kto2* region of the genome

Complementation group	No. of alleles	Identified gene	No. of core amino acids[a]	No. evolutionarily conserved[b]
l(3)76BDa	9	*trc*	459	440
l(3)76BDb	17	*ash1*	2217	968
l(3)76BDc	20	*kto*	2531	1598
l(3)76BDd	6			
l(3)76BDe	16	*Mi-2*	1982	1536
l(3)76BDf	4	*Taf6*	606	428
l(3)76BDg	2			
l(3)76BDh	3			
l(3)76BDi	15	*trpml*	652	468
l(3)76BDj	8			
l(3)76BDk	16	*serp*	541	490
l(3)76BDl	15	*Gyc76C*	1525	1054
l(3)76BDm	13	*CG8793*	1319	866
l(3)76BDn	8	*Papss*	615	533
l(3)76BDo	9	*Su(z)12*	805	202
l(3)76BDp	1	*pipe*	95	89
l(3)76BDq	2	*asf1*	218	175
l(3)76BDr	4	*CG32210*	1747	781
l(3)76BDs	5	*Su(Tpl)*	1059	280
l(3)76BDt	3			
l(3)76BDu	3	*Rab8*	207	200
l(3)76BDv	2	*Deaf1*	573	353
l(3)76BDw	8	*HLH106*	1113	493
l(3)76BDx	6	*verm*	489	389

[a] Core amino acids are those present in all protein isoforms for that gene.

[b] Evolutionarily conserved amino acids are those core amino acids conserved among nine Drosophila species. Because of sequence gaps, we did not include in the analysis the *D. yakuba pipe* gene or the *D. persimilis serp*, *Mi-2*, and *Gyc76C* genes.

gene. The remaining 19 transposon insertion mutants fail to complement one or more of our complementation groups. The complementation groups that failed to complement each transposon insertion mutant are shown in Table 2. For all of the transposon insertion mutants from the Drosophila Gene Disruption Project, the location from the deletion mapping coincides with the location of the transposon insertion. Eleven of our 24 complementation groups (46%) were represented by transposon insertion mutants from the Drosophila Gene Disruption Project. This is only slightly higher than the 40% frequency reported by the Gene Disruption Project for the entire autosomal collection 6 years ago (Bellen *et al.* 2004). We used this information to assign 9 of our complementation groups in Table 1 to molecularly identified genes. Because of the overlapping nature of the *Mi-2* and *Su(Tpl)* transcription units, the locations of the transposon insertion mutants were insufficient to comfortably assign the *l(3)76BDe* and *l(3)76BDs* complementation groups.

Since more than half of the essential genes in this region were not represented in the Drosophila Gene Disruption Project collection, we wanted to isolate additional *P*-transposon insertion mutants within this region of the genome to correlate more of our genetic complementation groups with molecularly identified genes. We chose to use local *P* transposition to isolate additional lethal transposon mutations within *Df(3L)kto2*. We started with a *P{lacW}* insertion mutant of *Mi-2*. Among ~68,000 progeny, we recovered and tested 475 putative local *P*-element transpositions (see materials and methods). Twenty-five of the local transpositions failed to complement *Df(3L)kto2* for viability. By crossing to representative deletions and mutations, we found that in all 25 putative local transpositions, the failure to complement *Df(3L)kto2* mapped to *Su(Tpl)*. The majority (22/25) of the new insertion chromosomes also showed variegation of the *mini-white* gene in *P{lacW}*. This variegation is suppressed when heterozygous to *Su(z)12* alleles, suggesting that the variegation is due to repeat-induced silencing (Dorer and Henikoff 1994). We characterized 18 of the local transpositions by inverse PCR. The original *P{lacW}* insertion appeared to be present in all 18. We characterized six of these chromosomes in more detail. We could not detect any changes in the flanking DNA on either side of the original *P{lacW}* insertion and inverse PCR gave multiple bands for most of the local transpositions. Sequencing of the PCR products showed that one product from each transposition was usually from the original insertion, and the additional bands were from insertions of multiple *P{lacW}* elements into the original *P{lacW}* insertion. Repeat-induced silencing caused by the insertion of multiple *P{lacW}* elements appears to restore at least some function for the *Mi-2* gene, while reducing function of the *Su(Tpl)* gene. The *Su(Tpl)* transcription unit is entirely within the first large intron of the *Mi-2* transcription unit (and transcribed from the same DNA strand). Both genes are ubiquitously expressed during embryogenesis. The ability of repeat-induced silencing of a transposon insertion to alter function of both genes in opposite ways is intriguing. However, the important point for this study is the failure of local transposition to identify any new insertion mutants in the surrounding genes in the region. As in the original studies (Tower *et al.* 1993; Zhang and Spradling 1993), the vast majority of local transpositions are within the original *P*-transposon insertion. While these data do not help to identify additional transcription units with essential functions within this genomic region, they do suggest that local transposition does not occur at a sufficiently high frequency to make it a useful practical approach to tagging new genes.

In another approach to correlate transcription units with our complementation groups, we crossed our new mutants to mutants in previously identified genes. We assigned 4 of the complementation groups [*l(3)76BDb*, *l(3)76BDc*, *l(3)76BDi*, and *l(3)76BDp*] to previously

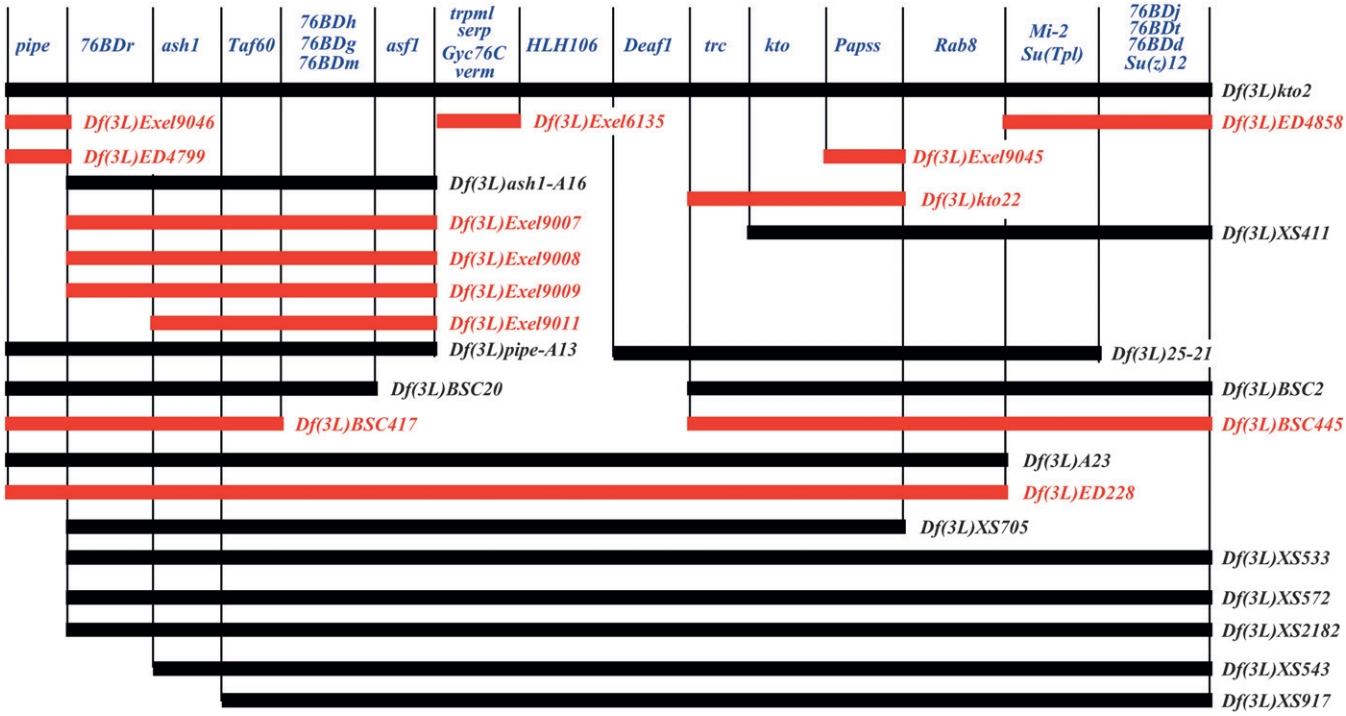

Figure 4.—Complementation of essential genes with a set of deletions that overlap *Df(3L)kto2*. The essential genes are listed at the top of the map in blue, with a horizontal bar for each deletion indicating which genes failed to complement that deletion. Not shown is the failure of *Df(3L)25-21* to complement *76BDm*, because the parental chromosome in which *Df(3L)25-21* was induced carried an allele of *76BDm* prior to generation of the deletion. Deletions with molecularly characterized breakpoints are indicated by red bars. Deletions with no molecular information are indicated by black bars.

identified genes by their failure to complement the mutations *ash1*^{6} (Tripoulas *et al.* 1994), *kto*^{1} (Kennison and Tamkun 1988; Treisman 2001), *trpml*^{1} (Venkatachalam *et al.* 2008), and *pip*^{C14} (Zhu *et al.* 2005), respectively. In addition, mutant alleles from six of our complementation groups were sequenced in earlier reports by ourselves and others [*l(3)76BDa* (Geng *et al.* 2000), *l(3)76BDe* (Kehle *et al.* 1998), *l(3)76BDo* (Birve *et al.* 2001), *l(3)76BDq* (Moshkin *et al.* 2002), *l(3)76BDs* (Eissenberg *et al.* 2002), and *l(3)76BDv* (Veraksa *et al.* 2002)] and assigned to the transcription units listed in Table 1. Finally, *HLH106* has been reported to be essential (Horn *et al.* 2003), but we were unable to obtain the original allele to test for complementation with our mutants. Our data from the deficiency mapping suggested that *l(3)76BDw* was the best candidate. We sequenced the *HLH106* gene from three alleles of our *l(3)76BDw* complementation group. All three alleles were isolated on the *red*^{1} *e*^{4} chromosome, which differs from the *iso-1* sequence at amino acid 224 (F in iso-1 and L in the *red*^{1} *e*^{4} marked chromosome). In addition, all three mutant alleles have single GC to AT bp transitions that introduce stop codons and prematurely truncate the HLH106 protein. *HLH106* alleles *1*, *2*, and *3* change amino acid residues Q715, W508, and W719, respectively, to stop codons. Using all of this information, we have been able to assign 19 of our 24 complementation groups to identified genes.

Chd3 is in a 200-kb region in which we have not identified any lethal complementation groups (Figure 3A). The CHD3 protein is 892 amino acids in length compared to 1982 amino acids for MI-2. Since we recovered 16 alleles of *Mi-2*, we might have expected to recover at least seven alleles of *Chd3*, if gene mutability is related to the number of amino acids in the encoded proteins. However, it was possible that *Chd3* is essential for zygotic viability and that we had failed to isolate mutant alleles by chance. It was also possible that *Chd3* is not required for zygotic viability, but is required for fertility of males and/or females. We decided to use ends-out, or replacement, gene targeting techniques (Gong and Golic 2003; Min *et al.* 2005) to make a deletion of most of the *Chd3* open reading frame. The scheme for generating the targeted deletion is shown in Figure 5. Although the *CG33062* gene model has been withdrawn by the FlyBase Genome Annotators (http://flybase.org/reports/FBgn0053062.html), at the time that we designed our targeted deletion, the close proximity of the *CG33062* and *Chd3* transcription start sites was a concern. Another factor that influenced the design of our deletion was the presence of several *Hin*dIII restriction sites in close proximity (*Hin*dIII is one of the restriction sites in the vector into which DNA flanking the targeted deletion is inserted). To avoid possible effects on transcription of *CG33062*, we decided to leave the *Chd3* promoter intact and to replace

TABLE 2

Transposon insertion mutants identified by the Drosophila Gene Disruption Project and maintained by the Bloomington Drosophila Stock Center

Transposon insertion mutants	Complementation groups affected
P{wHy}CG32210^{DG04109}	*l(3)76BDr*
Mi{ET1}ash1^{MB03235}	*l(3)76BDb*
PBac{WH}Taf6^{f06930}	*l(3)76BDf*
P{EP}CG8793^{G14537}	*l(3)76BDm*
P{lacW}l(3)76BDm^{L3809}	*l(3)76BDm*
PBac{5PHw^{+}}Gyc76C^{A377}	*l(3)76BDl*
P{SUPor-P}Gyc76C^{KG02500}	*l(3)76BDl*
PBac{RB}serp^{e02821}	*l(3)76BDk*
P{SUPor-P}verm^{KG07819}	*l(3)76BDx*
P{EPgy2}Papss^{EY08999}	*l(3)76BDn*
PBac{5HPw^{+}}Papss^{B412}	*l(3)76BDn*
PBac{5HPw^{+}}Rab8^{B229}	*l(3)76BDu*
P{lacW}l(3)L1243^{L1243}	*l(3)76BDe*
P{lacW}Mi-2^{j3D4} Su(Tpl)^{j3D4}	*l(3)76BDe*
P{PZ}Mi-2^{01058}	*l(3)76BDe*
P{wHy}Mi-2^{DG14402} Su(Tpl)^{DG14402}	*l(3)76BDe*
P{EPgy2} Su(Tpl)^{EY08696} Mi-2^{EY08696}	*l(3)76BDe*
P{EPgy2}Mi-2^{EY08138} Su(Tpl)^{EY08138}	*l(3)76BDs*
P{wHy}Su(Tpl)^{DG12505} Mi-^{2DG12505}	*l(3)76BDs*
P{EPgy2}EY03157	Viable and fertile
P{EPg}HP30622	Viable and fertile
PBac{WH}Shal^{f00495}	Viable and fertile
PBac{3HPy^{+}}C260	Viable and fertile

1898 bp of the *Chd3* transcription unit (and an additional 1175 bp downstream) with a transgene that expresses the GFP protein from an *armadillo* gene promoter (Min *et al.* 2005). Our mutant allele, which we call *Chd3^{1}*, lacks the final 529 amino acids (~60% of the protein), including ~80% of the SNF2 domain. We verified the presence of the deletion in the *Chd3^{1}* mutant chromosome by the failure to amplify by PCR *Chd3* genomic sequences in the *Chd3^{1}* homozygotes (shown in Figure 5B). We examined transcript levels of the remaining 5′ region of *Chd3* in *Chd3^{1}* homozygotes using quantitative RT–PCR. Surprisingly, the *Chd3^{1}* mutant transcripts are present at higher levels than *Chd3* transcripts in wild type (compared to *RpL32* transcript levels) in both embryos and first instar larvae (data not shown). *Chd3^{1}* homozygotes and hemizygotes show no reduction in viability, nor any discernible effects on fertility of either males or females. Thus, we conclude that *Chd3* is dispensable for viability and fertility. Although we have not noted any phenotypes associated with loss of *Chd3* function, we have not carefully examined the *Chd3^{1}* mutant flies for changes in behavior, learning, or responses to environmental challenges.

There are several features of the 76B-D genomic region that we believe merit discussion. The first feature is the presence of a 31.4 -kb tandem duplication (present at the left of Figure 3C) that involves three transcription units. Two of the transcription units are completely duplicated (*trpml/CG42638* and *CG42529/CG14101*). Only the coding exons of the third gene (*CG42636/CG42637*, also known as *Gyc76C*) are duplicated, with several 5′ noncoding exons proximal to, and outside of, the tandem duplication. The duplication is flanked by *Doc* transposable elements. Tandem duplications appear to be generated during meiosis at appreciable frequencies (Green 1959, 1961; Gelbart and Chovnick 1979). One mechanism for the generation of tandem duplications is the unequal exchange between different insertions of the same transposable element (Goldberg *et al.* 1983). In the present example, unequal exchange between the *Doc* transposable elements could have generated the duplication in the sequenced *iso-1* strain. We isolated EMS-induced mutants for both *trpml* and *Gyc76C* in the *red^{1} e^{4}* strain, which would be unlikely if the coding regions of both genes are duplicated in this strain. It is possible that the duplication is present only in the *iso-1* strain. To test for the presence of this tandem duplication, we designed primers flanking each of the *Doc* transposable elements in the genomic sequence and used them for PCR analysis. Each of these primers was ~200 bp from the ends of the *Doc* elements. The locations of these primers are shown in Figure 3C. We confirmed the presence of the distal and proximal *Doc* insertions in the *iso-1* strain. These *Doc* insertions were not present in any of the other strains that we examined (Oregon R, Canton-S, and the *red^{1} e^{4}* strain in which our mutants were induced). We were unable to confirm the presence of the middle *Doc* element that forms the junction between the two duplicated segments in our copy of the *iso-1* strain, even though the primers that should amplify this middle element are two of the same primers that were used to amplify the distal and proximal *Doc* elements under the same PCR conditions. Thus, we believe that this duplication arose after we constructed the *iso-1* strain in 1986, but before the BAC and WGS genomic libraries were made in 1998 and 1999, respectively (Celniker *et al.* 2002). Another polymorphic tandem duplication of >30 kb was found by the Drosophila Genome Project in the *iso-1* strain (Celniker *et al.* 2002). This duplication is flanked by *hobo* and *cruiser* transposable elements. Polymorphic duplications are common in human populations and may be a significant cause of human disease (Iafrate *et al.* 2004; Sebat *et al.* 2004; De Smith *et al.* 2008; Wain *et al.* 2009; Stankiewicz and Lupski 2010).

There are also several genes in this 640-kb region of the genome that appear to have arisen by tandem gene duplication before the divergence of the Drosophila genus. Approximately 40 kb proximal to *Chd3* there is a cluster of three adjacent genes (*CG9449*, *CG9451*, and *CG9452*), which encode predicted acid phosphatases of ~400 amino acids each. These three putative acid phosphatases are more related to each other (~50% identical over the entire length) than to any other proteins in

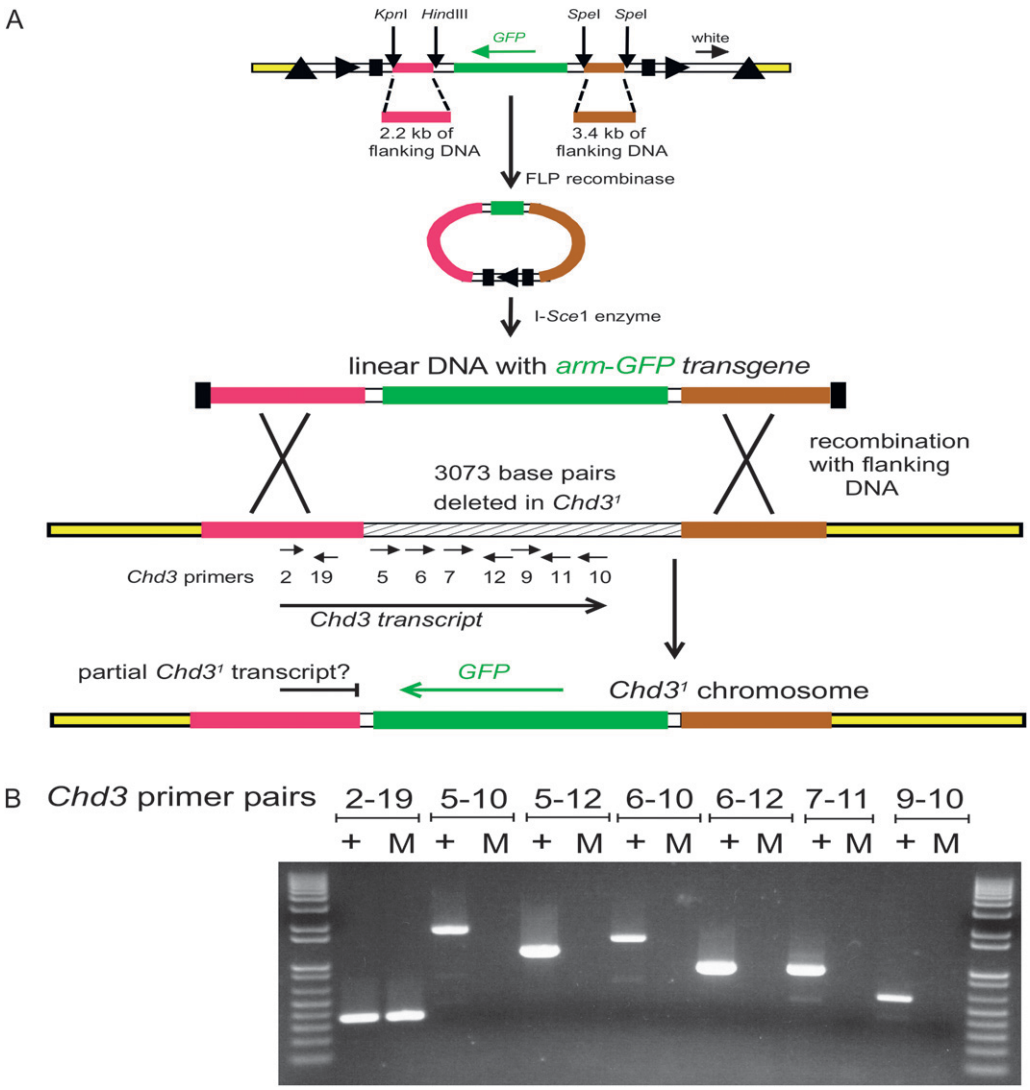

FIGURE 5.—Scheme for generating the *Chd3*^{1} allele. (A) The *P*-element transposon for generating the *Chd3*^{1} deletion allele is shown at the top, integrated into genomic DNA at a random location. The *P*-element terminal repeats are indicated by upward pointed black triangles. The yeast FRT elements are indicated by the rightward pointing black triangles. The I-*Sce*I restriction sites are indicated by the vertical black bars. The *arm-GFP* transgene is indicated by a green bar, with the green arrow above showing the direction of transcription. The flanking DNA distal to *Chd3* is indicated by the pink bars. The flanking DNA proximal to *Chd3* is indicated by the brown bars. Flanking genomic DNA is indicated by the yellow bars. Flies carrying the *P*-element transposon are crossed to flies carrying transgenes that express the FLP recombinase and the I-*Sce*I restriction enzyme from *Hsp70* promoters. After two 37° heat shocks (during the first and second instars) recombination between the FRT sites and cutting of the I-*Sce*I restriction sites generates the linear fragment shown in the middle of the figure. Recombination between this linear DNA fragment and the endogenous *Chd3* gene replaces 3073 bp of the *Chd3* gene (indicated by the hatched bar) with the *arm-GFP* transgene, generating the *Chd3*^{1} allele shown at the bottom. (B) Agarose gel showing PCR products from genomic DNA of Canton S wild type (+) and *Chd3*^{1} mutants (M) using various combinations of *Chd3* primers (locations shown in A). Primers 2 and 19 are from the 5′ region of *Chd3* that was not deleted. The remaining primers are within the targeted deletion region. As expected, the primer pair 2–19 amplified bands of the expected size from both wild type and *Chd3*^{1}. The primer pairs from the targeted deletion region amplified bands of the expected sizes from wild type, but failed to amplify bands from *Chd3*^{1}.

the genome. This cluster is conserved among all of the Drosophila species that we examined; however, we were unable to identify a homolog in the more distantly related dipteran, the mosquito *Anopheles gambiae*. Just proximal to the acid phosphatase cluster is another cluster of genes related to cuticular proteins. Only the first two genes (*Cpr76Ba* and *Cpr76Bb*) appear to have arisen by tandem gene duplication. Both predicted proteins are ~200 amino acids in length and are 44% identical. The other two genes, (*Cpr76Bc* and *Cpr76Bd*) are both predicted to encode much larger proteins with only limited homology to each other and to the *Cpr76Ba* and *Cpr76Bb* proteins. While all four of these genes are conserved in all of the sequenced Drosophila species that we examined, *Cpr76Bb* is the only annotated gene within *Df(3L)kto2* for which no cDNA or EST has been recovered. In the mosquito *A. gambiae*, this cluster has expanded to include at least nine genes related to *Cpr76Ba* and *Cpr76Bb*. Finally, the most proximal pair of adjacent genes within the *Df(3L)kto2* genomic region, which appear to have arisen by gene duplication, *vermiform* (*verm*) and *serpentine* (*serp*), both encode predicted proteins of just over 540 amino acids that are 57% identical along their entire lengths. Both genes are present in all of the Drosophila species examined, as well as in the mosquito *A. gambiae*. Both *verm* and *serp* are essential for zygotic viability.

Surrounding *Chd3*, there is an unusual cluster of related genes that we call the OAK cluster (after the *825-Oak* gene). The *D. melanogaster* OAK cluster genes are

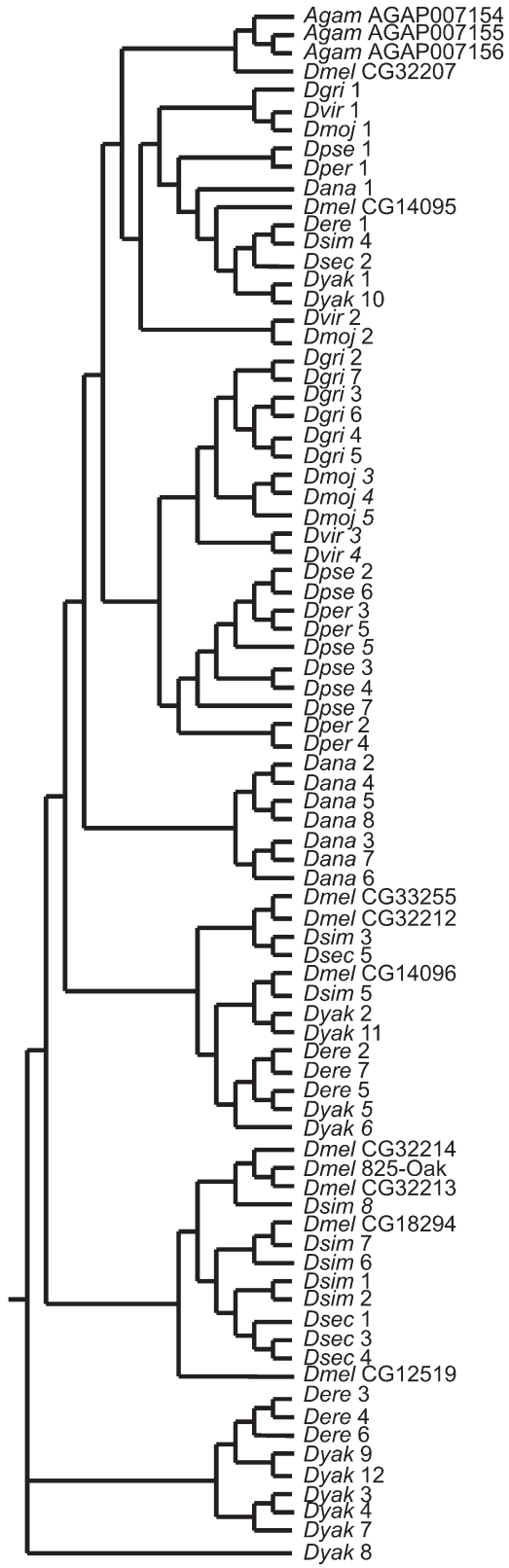

Figure 6.—Cladogram of the OAK cluster proteins. Only the proteins from *A. gambiae* (*Agam*) and *D. melanogaster* (*Dmel*) are encoded by annotated genes. The proteins from *D. grimshawi* (*Dgri*), *D. mojavensis* (*Dmoj*), *D. virilis* (*Dvir*), *D. pseudoobscura* (*Dpse*), *D. persimilis* (*Dper*), *D. ananassae* (*Dana*), *D. yakuba* (*Dyak*), *D. erecta* (*Dere*), *D. sechellia* (*Dsec*), and *D. simulans* (*Dsim*) are derived from open reading frames in the genomic regions between *CG32206* and *CG9449* in each species, or from unlocalized contigs (*Dper* 5, *Dsec* 4, *Dsec* 5, *Dyak* 8, *Dyak* 9, *Dyak* 10, *Dyak* 11, and *Dyak* 12). The proteins from each species are numbered in order on the chromosome beginning with the open reading frame that is closest to *CG32206*.

the dark green transcription units in Figure 3A. There are 3 OAK cluster genes in *A. gambiae* and between 4 and 12 OAK cluster genes in each of the other Drosophila species examined. The 3 OAK cluster genes in *A. gambiae* are adjacent to each other. Most of the OAK cluster genes in Drosophila are between *CG32206* and *CG9449* (both of which are conserved between all 11 Drosophila species examined). The remainder of the OAK cluster genes (1 in *D. persimilis*, 2 in *D. sechellia*, and 5 in *D. yakuba*) are on small unlocalized contigs and could also be from the genomic regions between *CG32206* and *CG9449*. The sequence comparisons of the OAK cluster proteins (the cladogram in Figure 6) suggest either that many of the duplications occurred after the divergence of each species or that there is concerted evolution actively constraining sequence divergence within a particular species. For example, 7 of the 8 genes in *D. ananassae* are more related to each other than to any other genes in the Drosophila genus. Similarly, all 6 of the *D. grimshawi* genes are more closely related to each other than to any of the other Drosophila genes. Another odd feature of the OAK cluster is that the distal and proximal genes in *D. melanogaster* overlap duplicated noncoding RNAs (*CR32205* and *pncr009:3L*, the light green transcription units in Figure 3A) that are transcribed from the opposite DNA strand. The *CR32205* and *pncr009:3L* RNA sequences are ~75% identical over their entire lengths. In addition to the OAK cluster in 76B, there are more distantly related clusters in the same chromosome arm (Muller element D) of all Drosophila species. In *D. melanogaster*, these include a cluster of 3 genes (*CG13679*, *CG13678*, and *CG13674*) in 66B12-66C1 and a cluster of 4 genes (*CG13051*, *CG13066*, *CG13067*, and *CG13069*) in 72D10.

DISCUSSION

D. melanogaster has two MI-2 type genes, but only one (*Mi-2*) appears to be essential for viability. *Chd3*, the second MI-2 type gene, is most highly expressed maternally (Murawska *et al.* 2008), but the maternal expression does not appear to be required for oogenesis or for the survival of the resulting progeny. Both MI-2 and CHD3 proteins are recruited to the same sites on polytene chromosomes, sites that also bind elongating RNA polymerase II (Murawska *et al.* 2008). At least one MI-2 type gene can be found in all multicellular eukaryotes, including *A. gambiae* and all Drosophila species examined. Only in the *melanogaster* subgroup, is there a second MI-2 type gene, *Chd3* (Figure 7A). *Mi-2* and *Chd3*

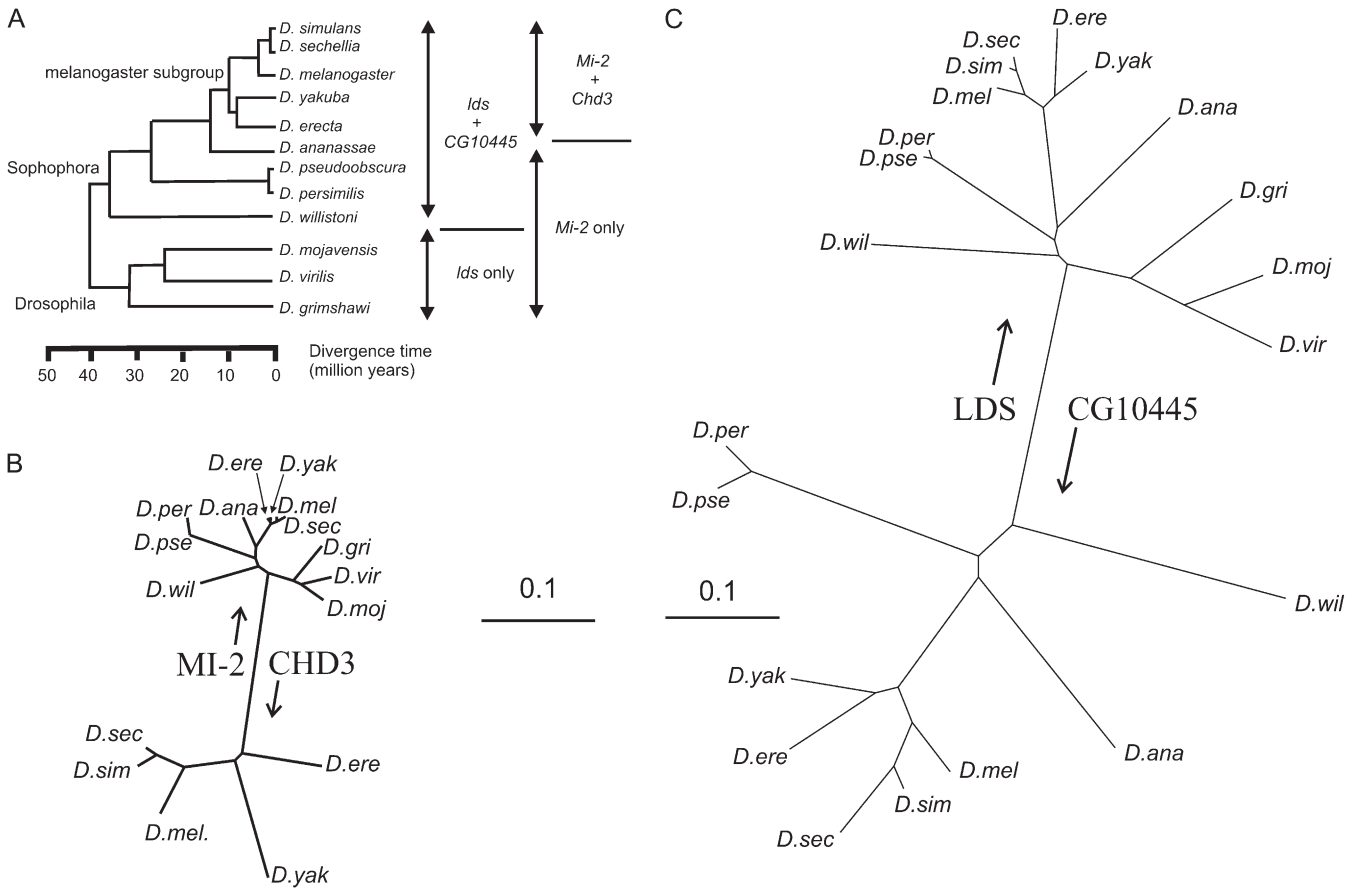

Figure 7.—Evolution of the *Mi-2/Chd3* and *lds/CG10445* gene pairs in the Drosophila lineage. (A) The phylogeny of the 12 sequenced Drosophila species. *Mi-2* and *lds* are present in all species, while *Chd3* is present only in the *melanogaster* subgroup and *CG10445* is present only in the Sophophora. (B) The radial phylogram for the MI-2 and CHD3 proteins in Drosophila. The *D. simulans* MI-2 protein was not included in the analysis because there is a large sequencing gap in the central half of the open reading frame. (C) The radial phylogram for the LDS and CG10445 proteins in Drosophila. The phylograms in B and C are both at the same scale. The difference in sizes between the phylograms is due primarily to the slower divergence of the *Mi-2* genes.

can be easily distinguished, as the MI-2 protein is almost twice the size of the CHD3 protein. *Mi-2* has also been more highly conserved during evolution of the genus Drosophila (Figure 7B).

It has been suggested that *Chd3* originated as a duplication in the *melanogaster* subgroup from the integration of a truncated, reverse-transcribed *Mi-2* mRNA, based on the lack of introns in *Chd3* (Murawska *et al.* 2008). We would point out that the homology between *Mi-2* and *Chd3* in the *melanogaster* subgroup is entirely within a single exon of the *Mi-2* genes. Thus, regardless of the mechanism of gene duplication, *Chd3* would be expected to lack introns. The phylogram of MI-2 and CHD3 proteins in the genus Drosophila (Figure 7B) does not fit the proposal that *Chd3* arose only in the *melanogaster* subgroup, but suggests that the original duplication occurred much earlier and might have been present in all of the Sophophora, with subsequent deletion in some species. *Chd3* is within the OAK cluster of related genes. Unequal recombination between OAK cluster genes that flank *Chd3* would delete *Chd3* (Green 1959, 1961; Goldberg *et al.* 1983). Unequal recombination would also help to explain the concerted evolution of the OAK cluster genes (Williams *et al.* 1989). The transcription unit just distal to *Chd3*, *CG33062*, is found in all of the Sophophora, but not in the subgenus Drosophila, suggesting that it might have arisen during the same event that generated *Chd3*. Another of the SNF2/RAD54 related genes shown in Figure 2, the *lodestar* (*lds*) gene, also appears to be duplicated in the Sophophora (Figure 7, A and C). The duplicated gene, *CG10445*, is adjacent to *lds* in all species except *D. willistoni*, with *lds* and *CG10445* divergently transcribed.

While there are large differences in gene mutability in *D. melanogaster* (Lefevre and Watkins 1986), the basis has not been clear. There are several factors that could contribute, such as differing lengths of transcription units, differing lengths of proteins, and differing degrees in the ability of proteins to tolerate single amino acid changes and still retain function. We have identified the transcription units for 19 of our essential complementation groups, and since the mutations in Table 1 were all selected in the same sets of experiments, we can use the numbers of alleles for each of the 19 comple-

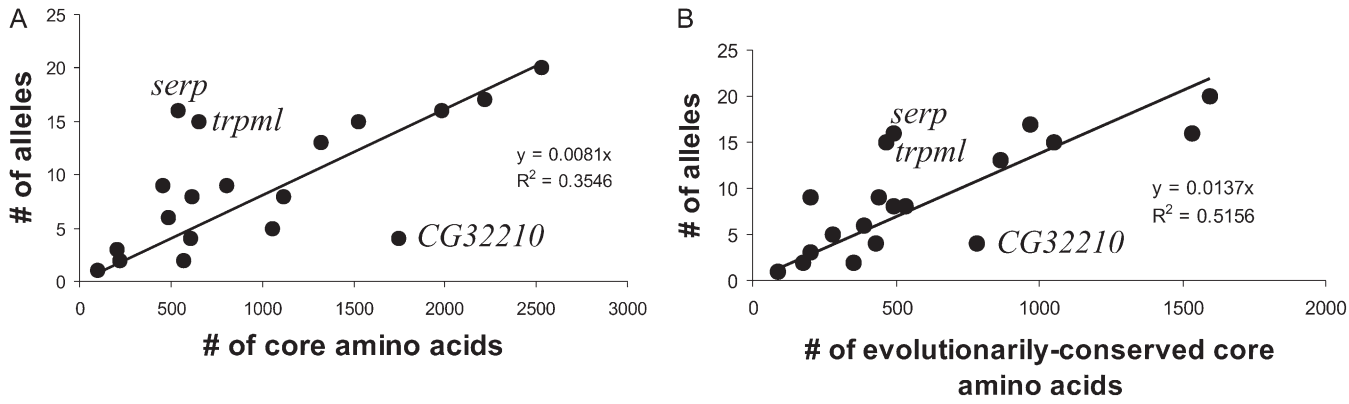

FIGURE 8.—Correlations between the numbers of alleles recovered after EMS mutagenesis and the numbers of amino acids. (A) The correlation between the numbers of alleles and the number of core amino acids (the amino acid residues present in all protein isoforms encoded by a given gene). (B) The correlation between the numbers of alleles and the evolutionarily conserved core amino acids (core amino acid residues identical in nine Drosophila species). In both cases, the genes with the poorest correlation are *serp*, *trpml*, and *CG32210*. R^2 is the coefficient of determination.

mentation groups to assess the role of various parameters in causing differences in gene mutability. Of the parameters above, we find that the best correlations are between mutability and length of protein products. Figure 8 shows comparisons of EMS mutability with both the numbers of core amino acids (those present in all protein isoforms for a given gene) and with the numbers of evolutionarily conserved core amino acids. With the exceptions of *serp* and *trpml*, which are more mutable than expected, and *CG32210*, which is less mutable than expected, the correlations are striking. From the linear regressions, we would estimate that we have isolated an average of one mutation for every 123 core amino acids, or one mutation for every 73 evolutionarily conserved core amino acids.

Another important conclusion from the work presented here is that the Drosophila Gene Disruption Project has been very successful in producing mutants for almost half of the transcription units, but that additional methods will be necessary to generate mutants for the other half of the transcription units. The more traditional approach that we have presented here, *i.e.*, to saturate specific regions of the genome with radiation or chemically induced mutations, can identify the essential functions, but is not very effective in identifying mutations in nonessential genes. While targeted gene replacement (RONG and GOLIC 2000) can be used for any gene regardless of function, it is currently not efficient enough for large-scale production of mutants. KOUNDAKJIAN *et al.* (2004) described a very powerful approach that could be used to identify both essential and nonessential genes. They produced a large number of stocks with highly mutagenized autosomes. Even heterozygotes from such a collection could be screened for mutations in specific genes using high-throughput methods, such as TILLING (targeting induced local lesions in genomes) (MCCALLUM *et al.* 2000). Another approach is the generation of stocks with transgenes for conditional RNA interference (DIETZL *et al.* 2007; NI *et al.* 2008). These conditional mutants are useful for characterizing the functions of both essential and nonessential genes. While tremendous progress has been made in the past 100 years, much work still remains before we completely understand the functions encoded by the *D. melanogaster* genome.

We thank David Stein, Jürg Müller, and Gerry Rubin for sharing unpublished information and mutant strains. We also thank Craig Montell and the Bloomington Drosophila Stock Center for providing mutant strains. Finally, we are grateful to Judy Kassis for helpful comments on both the experiments and the manuscript. This research was supported by the Intramural Research Program of the National Institutes of Health, National Institute of Child Health and Human Development.

LITERATURE CITED

BELLEN, H. J., R. W. LEVIS, G. LIAO, J. W. CARLSON, G. TSANG *et al.*, 2004 The BDGP gene disruption project: single transposon insertions associated with 40% of Drosophila genes. Genetics **167:** 761–781.

BIRVE, A., A. K. SENGUPTA, D. BEUCHLE, J. LARSSON, J. A. KENNISON *et al.*, 2001 *Su(z)12*, a novel *Drosophila* Polycomb group gene that is conserved in vertebrates and plants. Development **128:** 3371–3379.

BRIZUELA, B. J., L. ELFRING, J. BALLARD, J. W. TAMKUN and J. A. KENNISON, 1994 Genetic analysis of the *brahma* gene of *Drosophila melanogaster* and polytene chromosome subdivisions 72AB. Genetics **137:** 803–813.

CELNIKER, S. E., D. A. WHEELER, B. KRONMILLER, J. W. CARLSON, A. HALPERN *et al.*, 2002 Finishing a whole-genome shotgun: release 3 of the *Drosophila melanogaster* euchromatic genome sequence. Genome Biol. **3:** 0079.1–0079.14.

DAUBRESSE, G., R. DEURING, L. MOORE, O. PAPOULAS, I. ZAKRAJSEK *et al.*, 1999 The *Drosophila kismet* gene is related to chromatin-remodeling factors and is required for both segmentation and segment identity. Development **126:** 1175–1187.

DE SMITH, A. J., R. G. WALTERS, P. FROGUEL and A. I. BLAKEMORE, 2008 Human genes involved in copy number variation: mechanisms of origin, functional effects and implications for disease. Cytogenet. Genome Res. **123:** 17–26.

Dietzl, G., D. Chen, F. Schnorrer, K. C. Su, Y. Barinova *et al.*, 2007 A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. Nature **448:** 151–156.

Dorer, D. R., and S. Henikoff, 1994 Expansions of transgene repeats cause heterochromatin formation and gene silencing in Drosophila. Cell **77:** 993–1002.

Eisen, J. A., K. S. Sweder and P. C. Hanawalt, 1995 Evolution of the Snf2 family of proteins: subfamilies with distinct sequences and functions. Nucleic Acids Res. **23:** 2715–2723.

Eissenberg, J. C., J. Ma, M. A. Gerber, A. Christensen, J. A. Kennison *et al.*, 2002 dELL is an essential RNA polymerase II elongation factor with a general role in development. Proc. Natl. Acad. Sci. USA **99:** 9894–9899.

Gelbart, W. M., and A. Chovnick, 1979 Spontaneous unequal exchange in the rosy region of *Drosophila melanogaster*. Genetics **92:** 849–859.

Geng, W., B. He, M. Wang and P. N. Adler, 2000 The tricornered gene, which is required for the integrity of epidermal cell extensions, encodes the Drosophila nuclear DBF2-Related kinase. Genetics **156:** 1817–1828.

Goldberg, M. L., J.-Y. Sheen, W. J. Gehring and M. M. Green, 1983 Unequal crossing-over associated with asymmetrical synapsis between nomadic elements in the *Drosophila melanogaster* genome. Proc. Natl. Acad. Sci. USA **80:** 5017–5021.

Gong, W. J., and K. G. Golic, 2003 Ends-out, or replacement, gene targeting in Drosophila. Proc. Natl. Acad. Sci. USA **100:** 2556–2561.

Green, M. M., 1959 Non-homologous pairing and crossing over in Drosophila melanogaster. Genetics **44:** 1243–1256.

Green, M. M., 1961 Further data on non-homologous pairing and crossing over in Drosophila melanogaster. Genetics **46:** 1555–1560.

Horn, C., N. Offen, S. Nystedt, U. Hacker and E. A. Wimmer, 2003 piggyBac-based insertional mutagenesis and enhancer detection as a tool for functional insect genomics. Genetics **163:** 647–661.

Huang, A. M., E. J. Rehm and G. M. Rubin, 2000 Recovery of DNA sequences flanking P-element insertions: inverse PCR and plasmid rescue, pp. 429–437 in *Drosophila Protocols*, edited by W. A. Sullivan, M. Ashburner and R. S. Hawley, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Iafrate, A. J., L. Feuk, M. N. Rivera, M. L. Listewnik, P. K. Donahoe *et al.*, 2004 Detection of large-scale variation in the human genome. Nat. Genet. **36:** 949–951.

Jacobs, S. A., and S. Khorasanizadeh, 2002 Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. Science **295:** 2080–2083.

Kehle, J., D. Beuchle, S. Treuheit, B. Christen, J. A. Kennison *et al.*, 1998 dMi-2, a hunchback-interacting protein that functions in Polycomb repression. Science **282:** 1897–1900.

Kennison, J. A., 1983 Analysis of *Y*-linked mutations to male sterility in *Drosophila melanogaster*. Genetics **103:** 219–234.

Kennison, J. A., and J. W. Tamkun, 1988 Dosage-dependent modifiers of Polycomb and Antennapedia mutations in *Drosophila*. Proc. Natl. Acad. Sci. USA **85:** 8136–8140.

Kent, W. J., 2002 BLAT: the BLAST-like alignment tool. Genome Res. **12:** 656–664.

Konev, A. Y., M. Tribus, S. Y. Park, V. Podhraski, C. Y. Lim *et al.*, 2007 CHD1 motor protein is required for deposition of histone variant h3.3 into chromatin in vivo. Science **317:** 1087–1090.

Koundakjian, E. J., D. M. Cowan, R. W. Hardy and A. H. Becker, 2004 The Zuker collection: a resource for the analysis of autosomal gene function in *Drosophila melanogaster*. Genetics **167:** 203–206.

Lefevre, Jr., G., and W. Watkins, 1986 The question of the total gene number in *Drosophila melanogaster*. Genetics **113:** 869–895.

Lewis, E. B., and F. Bacher, 1968 Method of feeding ethyl methane sulfonate (EMS) to Drosophila males. Dros. Inf. Serv. **43:** 193.

Lindsley, D. L., and G. G. Zimm, 1992 *The Genome of Drosophila melanogaster*. Academic Press, San Diego.

Matsubayashi, H., and M. T. Yamamoto, 1998 Dissection of chromosome region 89A of *Drosophila melanogaster* by local transposition of P elements. Genes Genet. Syst. **73:** 95–103.

McCallum, C. M., L. Comai, E. A. Greene and S. Henikoff, 2000 Targeted screening for induced mutations. Nat. Biotechnol. **18:** 455–457.

McDaniel, I. E., J. M. Lee, M. S. Berger, C. K. Hanagami and J. A. Armstrong, 2008 Investigations of CHD1 function in transcription and development of Drosophila melanogaster. Genetics **178:** 583–587.

Min, G., X. Bi and Y. S. Rong, 2005 Targeted mutagenesis of *Drosophila atm* and *mre11* genes. Dros. Inf. Serv. **88:** 79–83.

Moshkin, Y. M., J. A. Armstrong, R. K. Maeda, J. W. Tamkun, C. P. Verrijzer *et al.*, 2002 Histone chaperone ASF1 cooperates with the Brahma chromatin-remodelling machinery. Genes Dev. **16:** 2621–2626.

Murawska, M., N. Kunert, J. van Vugt, G. Längst, E. Kremmer *et al.*, 2008 dCHD3, a novel ATP-dependent chromatin remodeler associated with sites of active transcription. Mol. Cell. Biol. **28:** 2745–2757.

Ni, J.-Q., M. Markstein, R. Binari, B. Pfeiffer, L.-P. Liu *et al.*, 2008 Vector and parameters for targeted transgenic RNA interference in *Drosophila melanogaster*. Nat. Methods **15:** 49–51.

Nielsen, P. R., D. Nietlispach, H. R. Mott, J. Callaghan, A. Bannister *et al.*, 2002 Stucture of the HP1 chromodomain bound to histone H3 methylated at lysine 9. Nature **416:** 103–107.

Odenwald, W. F., W. Rasband, A. Kuzin and T. Brody, 2005 EVOPRINTER, a multigenomic comparative tool for rapid identification of functionally important DNA. Proc. Natl. Acad. Sci. USA **102:** 14700–14705.

Pollard, K. J., and C. L. Peterson, 1998 Chromatin remodeling: A marriage between two families? Bioessays **20:** 771–780.

Rong, Y. S., and K. G. Golic, 2000 Gene targeting by homologous recombination in Drosophila. Science **288:** 2013–2018.

Sebat, J., B. Lakshmi, J. Troge, J. Alexander, J. Young *et al.*, 2004 Large-scale copy number polymorphism in the human genome. Science **305:** 525–528.

Stankiewicz, P., and J. R. Lupski, 2010 Structural variation in the human genome and its role in disease. Annu. Rev. Med. **61:** 437–455.

Tower, J., and R. Kurapati, 1994 Preferential transposition of a *Drosophila* P element to the corresponding region of the homologous chromosome. Mol. Gen. Genet. **244:** 484–490.

Tower, J., G. H. Karpen, N. Craig and A. C. Spradling, 1993 Preferential transposition of Drosophila *P* elements to nearby chromosomal sites. Genetics **133:** 347–359.

Treisman, J. E., 2001 Drosophila homologues of the transcriptional coactivation complex subunits TRAP240 and TRAP230 are required for identical processes in eye-antennal disc development. Development **128:** 603–615.

Tripoulas, N. A., E. Hersperger, D. La Jeunesse and A. Shearn, 1994 Molecular genetic analysis of the *Drosophila melanogaster* gene *absent, small or homeotic discs1* (*ash1*). Genetics **137:** 1027–1038.

Venkatachalam, K., A. A. Long, R. Elsaesser, D. Nikolaeva, K. Broadie *et al.*, 2008 Motor deficit in a *Drosophila* model of mucolipidosis type IV due to defective clearance of apoptotic cells. Cell **135:** 838–851.

Veraksa, A., X. Li, J. A. Kennison and W. McGinnis, 2002 DEAF-1 function is essential for the early embryonic development of *Drosophila*. Genesis **33:** 67–76.

Wain, L. V., J. A. L. Armour and M. D. Tobin, 2009 Genomic copy number variation, human health, and disease. Lancet **374:** 340–350.

Williams, S. M., J. A. Kennison, L. G. Robbins and C. Strobeck, 1989 Reciprocal recombination and the evolution of the ribosomal gene family of *Drosophila melanogaster*. Genetics **122:** 617–624.

Woodage, T., M. A. Basrai, A. D. Baxevanis, P. Hieter and F. S. Collins, 1997 Characterization of the CHD family of proteins. Proc. Natl. Acad. Sci. USA **94:** 11472–11477.

Zhang, P., and A. C. Spradling, 1993 Efficient and dispersed local *P* element transposition from Drosophila females. Genetics **133:** 361–373.

Zhang, P., and A. C. Spradling, 1994 Insertional mutagenesis of *Drosophila* heterochromatin with single *P* elements. Proc. Natl. Acad. Sci. USA **91:** 3539–3543.

Zhu, X. J., J. Sen, L. Stevens, J. S. Goltz and D. Stein, 2005 Drosophila Pipe protein activity in the ovary and the embryonic salivary gland does not require heparan sulfate glycosaminoglycans. Development **132:** 3813–3822.

Communicating editor: J. Tamkun