The Berkeley Drosophila Genome Project Gene Disruption Project: Single *P*-Element Insertions Mutating 25% of Vital Drosophila Genes

Allan C. Spradling,* Dianne Stern,* Amy Beaton,† E. Jay Rhem,† Todd Laverty,† Nicole Mozden,* Sima Misra† and Gerald M. Rubin†

**Department of Embryology, Howard Hughes Medical Institute Research Laboratories, Carnegie Institution of Washington, Baltimore, Maryland 21210 and †Department of Molecular and Cellular Biology, Howard Hughes Medical Institute Research Laboratories, University of California, Berkeley, California 94720*

Manuscript received January 29, 1999
Accepted for publication April 26, 1999

ABSTRACT

A fundamental goal of genetics and functional genomics is to identify and mutate every gene in model organisms such as *Drosophila melanogaster*. The Berkeley Drosophila Genome Project (BDGP) gene disruption project generates single *P*-element insertion strains that each mutate unique genomic open reading frames. Such strains strongly facilitate further genetic and molecular studies of the disrupted loci, but it has remained unclear if *P* elements can be used to mutate all Drosophila genes. We now report that the primary collection has grown to contain 1045 strains that disrupt more than 25% of the estimated 3600 Drosophila genes that are essential for adult viability. Of these *P* insertions, 67% have been verified by genetic tests to cause the associated recessive mutant phenotypes, and the validity of most of the remaining lines is predicted on statistical grounds. Sequences flanking >920 insertions have been determined to exactly position them in the genome and to identify 376 potentially affected transcripts from collections of EST sequences. Strains in the BDGP collection are available from the Bloomington Stock Center and have already assisted the research community in characterizing >250 Drosophila genes. The likely identity of 131 additional genes in the collection is reported here. Our results show that Drosophila genes have a wide range of sensitivity to inactivation by *P* elements, and provide a rationale for greatly expanding the BDGP primary collection based entirely on insertion site sequencing. We predict that this approach can bring >85% of all Drosophila open reading frames under experimental control.

THE nucleotide sequences of several complex eukaryotic genomes, including those of *Caenorhabditis elegans*, *Drosophila melanogaster*, *Arabidopsis thaliana*, *Mus musculus*, and *Homo sapiens*, are virtually complete or scheduled for completion during the next several years (Collins *et al.* 1998; Meinke *et al.* 1998). Large-scale sequencing of human and model organism genomes, cDNAs, and expressed sequence tags (ESTs) is identifying tens of thousands of genes about which little is known. Obtaining mutations in these loci on chromosomes free of additional lesions is essential for their functions to be deduced using model organisms. However, mutations in particular open reading frames must still usually be obtained in piecemeal fashion, by producing specifically tailored gene knockouts or by identifying the desired strains within large, randomly mutagenized collections. Both approaches remain slow and uncertain. These problems could be circumvented by constructing complete mutation libraries, whose strains each disrupt single distinct genes. Genome-wide collections of gene knockouts would provide a vital resource for gene-based approaches to biological research.

Insertional mutagenesis provides a highly advantageous strategy for constructing mutations in advance throughout entire genomes, because it simplifies the problem of determining which genes have been disrupted. Insertional screens at low multiplicity have been carried out in bacteria (Kleckner *et al.* 1977), yeast (Burns *et al.* 1994; Garraway *et al.* 1997), Arabidopsis (Sundaresan *et al.* 1995; Bhatt *et al.* 1996; Smith *et al.* 1996), *C. elegans* (Plasterk 1993; Korswagen *et al.* 1996), Drosophila (Cooley *et al.* 1988; Rørth 1996), zebrafish (Allende *et al.* 1996; Gaiano *et al.* 1996), and mice (Jaenisch 1988; Gossler *et al.* 1989; Wurst *et al.* 1995; Zambrowicz *et al.* 1998). However, converting the products of raw screens into a complete mutation library is a challenging task. The site selectivity of the mutagenic element must be extremely broad to target all genes. High throughput methods must be developed and used to identify screen products that contain single insertions located within distinct genes, because strains bearing just one new insertion each are needed to assess gene function. Consequently, it remains uncertain

Corresponding author: Allan C. Spradling, Howard Hughes Medical Institute Research Laboratories, Department of Embryology, Carnegie Institution of Washington, 115 W. University Pkwy., Baltimore, MD 21210. E-mail: spradling@mail1.ciwemb.edu

whether it is possible in practice to construct a complete library of mutations using this approach.

Among model multicellular eukaryotes, insertional mutagenesis has been used for genetic analysis and functional genomics most extensively in Drosophila. Low-multiplicity mutageneses using engineered *P* elements have been carried out frequently (Cooley *et al.* 1988; Bellen *et al.* 1989; Bier *et al.* 1989; Berg and Spradling 1991; Karpen and Spradling 1992; Gaul *et al.* 1992; Török *et al.* 1993; Chang *et al.* 1993; Erdelyi *et al.* 1995; Rørth 1996; Deak *et al.* 1997; Sozen *et al.* 1997; Rørth *et al.* 1998). While the raw strain collections produced in such studies are highly redundant and contain lines with multiple mutations, they provide ideal starting material for constructing a genome-wide mutation library. In 1993, the Berkeley Drosophila Genome Project (BDGP) gathered ~3900 lines associated with mutant phenotypes (mostly lethality) from seven existing raw collections and began to construct a gene disruption library for use by the Drosophila research community (Spradling *et al.* 1995). The Drosophila genome is thought to contain ~3600 vital genes (Miklos and Rubin 1996), so the project had the potential to encompass a substantial fraction of genes that can mutate to a phenotype recognizable in large laboratory screens (primarily lethality). In contrast, the total number of genes is thought to be about three times larger (Miklos and Rubin 1996). The analysis of these collections also promised to indicate the feasibility of eventually using *P*-element insertional mutagenesis to disrupt all Drosophila genes.

The seven initial collections have now been analyzed, and we document here a library of strains that disrupts ~1000 different genes. A total of 450 of the strains mutate genes that have been described previously in Drosophila or represent novel loci defined by homology to well-studied genes in other organisms. These associations were made with the assistance of researchers throughout the Drosophila community who have used the collection to help characterize >250 genes, and through the efforts of the BDGP project, including 138 new gene-mutation links reported here. Another 135 disrupted genes are associated only with EST sequences that predict novel proteins or products related to proteins of unknown function in other organisms. An additional 138 lines are inserted within sequenced regions containing candidate open reading frames (ORFs). Thus, >700 of the 1045 mutant strains already link mutant phenotypes with specific open reading frames, and the remaining lines only await completion of the genomic DNA sequence. The ~1000 genes already represented in the library constitute ~25% of all Drosophila genes readily defined by mutation, and specify more gene-mutation links than are currently available in other model multicellular eukaryotes. Most important, on the basis of this work we have developed a new program

TABLE 1

Mutational saturation

Times hit (n)	Chromosome *3* genes in Df[a]	$N = 483$, $\lambda = 0.558$[b]	613 $\lambda = 0.241$, 115 $\lambda = 1.51$[c]	$N = 111$, $P = 0.0714$[d]	Chromosome *2* genes in Df[a]	$N = 577$, $\lambda = 0.600$[b]	680 $\lambda = 0.331$, 110 $\lambda = 2.08$[c]	$N = 202$, $P = 0.0374$[d]
0		276	507	0.0		317	502	0.0
1	154	154	154	0.0	190	190	190	0.17
2	43	43	43.0	0.17	57	57	56.5	0.64
3	16	8.0	15.7	0.49	19	11.4	23.6	1.6
4	6	1.1	5.6	1.01	17	1.7	11.0	2.9
5	5	0.1	1.7	1.7	5	0.2	4.5	4.3
6–12	14	0	0.52	13.6	27	0	2.1	25
>12	4	0	0.0	0.94	5	0	0.0	1.1

[a] The number of genes located within available deficiencies on the indicated chromosome that were hit one, two, three, etc., times by verified single insertions within the strains analyzed. The deficiencies remove ~60% of chromosomes *2* and *3*.

[b] The predicted values from a model of a single mutability class (N, number of genes; λ, Poisson parameter). The fit to the observed data is poor: for just $n = 1$–5, $\chi^2 = 270$, $P \ll 0.001$ (chromosome *3*) and $\chi^2 = 263$, $P \ll 0.001$ (chromosome *2*). After correcting for the contributions from hotspots according to columns 5 and 9, the corresponding values still fit poorly: $\chi^2 = 20$, $P \ll 0.001$ (chromosome *3*) and $\chi^2 = 36$, $P \ll 0.001$ (chromosome *2*).

[c] The number of genes hit one to five times predicted by the sum of two Poisson distributions that model the coldspot and warmspot classes. For both mutability classes the first number corresponds to the total number of genes and the second number to the Poisson parameter λ. The parameters were determined using an Excel spreadsheet that allowed the distributions to be modeled using numerical methods. Genes hit >5 times are strong hotspots that are not modeled by the two predicted mutability classes. The fits after correcting for contributions from hotspots according to columns 5 and 9, for $n = 1$–5 are quite good: $\chi^2 = 0.81$, $P \gg 0.05$ (chromosome *3*) and $\chi^2 = 3.8$, $P \gg 0.05$ (chromosome *2*).

[d] The estimated number of genes hit 1–5 times contributed by hotspot genes. Hotspot genes hit between 6 and 12 times were assumed to be a uniform class and modeled using a binominal distribution. There were 14 such loci on chromosome *3*, while on chromosome *2* there were 25 loci, after subtracting the 2 loci in this category contributed by warmspot genes.

TABLE 2

Screen summaries

Screen	Reference[a]	Element	Raw II	Localized II	Single II	In Df II	Group II	Primary II	Gene II
Spradling	1	Phsneo	2	2	2	2	1	1	1
Spradling	2	PZ[ry]	495	479	432	292	291	240	233
Rubin	3	PZ[ry]	50	47	42	34	11	7	6
Jan	4	PlacW	1	1	1	1	1	1	1
Scott	5	PlacW	34	33	33	19	15	12	9
Kiss	6	PlacW	2153	1611	1133	712	855	351	302
Laughton	7	PlacW							
Totals			2735	2173	1643	1060	1174	612	552

Screen	Percentage in Df verified	Screen efficiency	Raw III	Localized III	Single III	in Df III	Group III	Primary III	Gene III
Spradling	63	61	68	68	66	35	45	23	21
Spradling	85	78	609	597	560	361	331	257	248
Rubin	78	76	96	94	94	55	43	24	23
Jan	88	83	201	190	185	124	84	64	61
Scott	63	50	79	67	59	34	37	20	18
Kiss	58	31							
Laughton	48	39	125	116	88	48	74	45	32
Totals			1178	1132	1052	657	614	433	401

The results of analyzing lines from seven single *P*-element autosomal insertion screens are summarized, grouped by chromosome (*II* and *III*). *raw*, number of starting lines; Localized, number of lines whose *P* insertions were localized by *in situ* hybridization; single, number of lines with a single *P* insertion; in Df, number of lines whose insertion fell within the cytogenetic boundaries of an available valid deficiency; Group, number of different complementation groups; Primary, number of lines in the primary collection. Note: The distribution of lines to the primary collection from the different screens is essentially arbitrary and was determined by the relative time each screen was analyzed as indicated by the order in which the screens are listed in the table. Genes, estimated number of different genes disrupted; % in Df verified, fraction of lines in valid deficiencies that were verified; Screen efficiency, the fraction of raw lines in a screen that are verifiable single insertions. Efficiency was calculated using percentage verified in Df to estimate the fraction of single insert lines outside of Dfs that would have been verified. The formula is eff = [verified + (% in Df verified) [not in Df and single insert]]/raw.

[a] 1, Cooley *et al.* (1988); 2, Karpen and Spradling (1992); 3, Gaul *et al.* (1992); 4, Bier *et al.* (1989); 5, M. Scott and M. Fuller, personal communication; 6, Török *et al.* (1993); 7, Chang *et al.* (1993).

to disrupt the remaining genes while the Drosophila genome sequence is being completed and annotated (see Table 1).

MATERIALS AND METHODS

Drosophila strains: Flies were grown on standard corn meal/agar media (Ashburner 1990) at 22°. Approximately 3900 lethal, semilethal, sterile, semisterile, or visible lines (Table 2) were collected from seven *P*-element screens (Cooley *et al.* 1988; Bier *et al.* 1989; Gaul *et al.* 1992; Karpen and Spradling 1992; Chang *et al.* 1993; Török *et al.* 1993; M. Scott and M. Fuller, unpublished results) as described (Spradling *et al.* 1995). Three different *P*-element vectors were used: PZ[ry] (Mlodzik *et al.* 1990), Plac W (Bier *et al.* 1989), or Puc-hsneo (Steller and Pirrotta 1986). About 40% of the starting lines were marked with *rosy*$^+$ and 60% with *white*$^+$. The Gaul *et al.* (1992) collection was stained for enhancer trap patterns in third instar larval eye-antenna imaginal discs, and only lines showing expression were analyzed further. Lines from the Török *et al.* (1993) screen that share the first three numbers in their designator (see *Nomenclature*) were obtained from the same parents and may derive from premeiotic clusters. When two or more such lines were found to contain insertions at the same polytene site, only one was retained and the other(s) was treated as a duplicate(s). Many lines containing multiple insertions from this screen were discarded prior to localization because they exhibited a diagnostic strong eye coloration.

Deficiency strains were obtained from the Bloomington Stock Center and from many individual laboratories. The deficiencies used are listed in Table 3.

Strain names: BDGP strain names start with a prefix that indicates the chromosome and phenotypic effect of their single *P*-element insertion. For example, third chromosome strain names begin with either "l(3)" (lethal or strong semilethal), "fs(3)" (female sterile or strong semisterile), "ms(3)" (male sterile or strong semisterile), "v(3)" (visible), or "n(3)" (no obvious phenotype). Semilethal and semisterile mutations were utilized only if they were strong enough to score in complementation tests. Only the effect of the *P* insertion, not of any secondary mutations on the same chromosome, whether present initially or acquired later, is indicated by the prefix. The phenotypic prefix is followed by a unique designator to distinguish individual lines and to preserve the

TABLE 3

Deficiency stocks

1. *Df(2L)net-PMF*	21A1;21B7-8
2. *Df(2L)al*	21B8-C1;21C8
3. *Df(2L)ast2*[a]	21D1-2;22B2-3
4. *Df(2L)dp-79*[b]	22A2-3;22D5-22E1
5. *Df(2L)DTD2*	22D5;23B1-2
6. *Df(2L)C144*	22F3-4;23C3-5
7. *Df(2L)lS17*	23C1-2;23E1-2
8. *In(2LR)DTD16[L]DTD42[R]*	23C;23E3-6
9. *Df(2L)ed-dp*	24C3;25A2
10. *Df(2L)ed1*	24A3;24D4
11. *Df(2L)tkv2*	25D2-4;25E1
12. *Df(2L)c1-h3*	25D2-3;26B2-5
13. *Df(2L)GpdhA*	25E1;26A8-9
14. *Df(2L)E110*	25F3-26A1;26D3-11
15. *Df(2L)Dwee-&dgr;5*	27A1-2;28A1-6
16. *Df(2L)spd[j2]*	27C1;28A1
17. *Df(2L)Dwee[wo5]*	27C2-3;27C4-5
18. *Df(2L)l-H*	27C7-9;28B3-4
19. *Df(2L)spd*	27E1-8;28C1-6
21. *Df(2L)XE-2750*	28B2;28D3
23. *Df(2L)TE29Aa-11*	28E4-7;29B2-29C1
24. *Df(2L)N22-14*	29C1-2;30C8-9;30D1-2;31A1-2
25. *Df(2L)N22-5*	29C3-5;30C8-9
26. *Df(2L)30A-C*	30A3-6;30C5
27. *Df(2L)l2*	31B1-5;32A1-2
28. *Df(2L)Prl*	32F1-3;33F1-2
29. *Df(2L)prd1.7*	33B2,3-34A1,2
30. *Df(2L)b84h50*	34D4;35C1-2
31. *In(2L)75c*	35A2;35D4-D7
32. *Df(2L)A48*	35B3;35D5-7
33. *Df(2L)r10*	35D1-2;36A7
34. *Df(2L)cact-255rv64*	35F6-12;36D1-3
35. *Df(2L)H20*	36A8-9;36F1
36. *Df(2L)VA18*	36D1;37C2-5
37. *Df(2L)TW50*	36E4-36F1;38A6-7
38. *Df(2L)TW161*	38A6;40A4-40B1
39. *Df(2L)TW84*	37F5-38A1;39D3-E1
40. *Df(2L)TW65*	38A1;39F1
41. *Df(2R)M41A4*	h44-46;42A2
42. *Df(2R)nap1*	41D2-41E1;42B1-3
43. *Df(2R)nap2*	41F4-9;43A1
44. *Df(2R)cn88b*	42A2-19;42E1-7
45. *Df(2R)nap16*	42A1-2;44B1-44C1
46. *Df(2R)42*	42C3;42D2
47. *In(2R)pk78s*	42C7;43F8;59F5-8
48. *Df(2R)pk78k*	42E3-43C3
49. *Df(2R)cn9*	42E1-7;44C1-2
50. *Df(2R)ST1*	43B3-4-43E18
51. *Df(2R)CA53*	43E6;44B6
52. *Df(2R)H3C1*	43F1-9;44D3-8
53. *Df(2R)44CE*	44C1-2;44D2-4
54. *Df(2R)Np3*	44D2-E1;45B8-C1
55. *Df(2R)Np1*	44F2-3;45C5-6
56. *Df(2R)Np4*	44F8-9;45C1
57. *Df(2R)Np5*	44F10;45D9-45E1
58. *In(2R)G63[L]w45-73n[R]*	45B1;45D1[b]
59. *Df(2R)X1*	46C1-2;47A1
60. *Df(2R)X3*	46C1-2;46E1-2
61. *Df(2R)12*	46E1-F11;47A13-B14
62. *Df(2R)Stan2*	46F1;47B9
63. *Df(2R)E3363*	47A3;47E1->
64. *Df(2R)en-SFX31*	48A1;48B5
65. *In(2R)vg[W]*	48A1-2;49D1-7
66. *Df(2R)vg135*	49A1-13;49E1-2
67. *Df(2R)vg-B*	49D3-4;50A2
68. *Df(2R)CX1*	49D1;50D1
69. *Df(2R)50C-101*	50C12-D1;50D1-7
70. *Df(2R)50C-38*	50C20-23;50D4-D7
71. *Df(2R)trix*	51A2;51B6
72. *Df(2R)03072*	51A5;51C1
73. *Df(2R)lp1*	51C3;52F8-9
74. *Df(2R)lp4*	51F13;52F8-9
75. *Df(2R)lp5*	52A13-14;52F10-11
76. *Df(2R)KL32*	52D1;52E2-5
78. *Df(2R)Pcl7B*	54E8-F1;55B9-C1
79. *Df(2R)RM2-1*	54F2;56A1
80. *Df(2R)PC4*	55A1;55F1-3
81. *Df(2R)PC29*	55C1-2;56B1-2
82. *Df(2R)P34*	55E6-55F3;56C1-11
83. *Df(2R)017*	56F5;56F15
84. *Df(2R)exu1*	57A2;57B1
85. *Df(2R)Pu-D17*	57B5;58B12
86. *Df(2R)XE3030*	57C2;58C1-7
87. *Df(2R)02311*	58D2;58E1
88. *Df(2R)59AD*	59A1-A3;59D1-D4
89. *In(2R)bw[VDe2L]Px[Kr]*	59D6,E1-60C,D
90. *Df(2R)or-BR11*	59F6;60B1->
91. *Df(2R)bw-S46*	59D8;60A8-16
92. *In(2LR)Px[4]*	60C5-C6;60D1
93. *Df(2R)Px2*	60C6;60D11
94. *Df(2R)M60E*	60E5-9;60E11
95. *Df(2R)Kr10*	60F1;F5

(continued)

original names of the lines. Designators for lines from Cooley *et al.* (1988) take the form "*neo*" and a 1–3 digit number ("*neo63*"); from Karpen and Spradling (1992), lines retain their original names ("06253"); from Bier *et al.* (1989), the letter "j" precedes the original name (*i.e.*, "*5C2*" becomes "*j5C2*"); from Gaul *et al.* (1992), the letter "r" is contained within the original name (*i.e.*, "*rJ713*"); from Török *et al.* (1993), the letter "k" precedes the original name and the slash is omitted (*i.e.*, "*133/45*" becomes "*k13345*"); from the Scott and Fuller screen, the letter "s" precedes the original name (*i.e.*, "*1629*" becomes "*s1629*"); and for regular names from Chang *et al.* (1993), the L is moved to the start, the R omitted, and a zero added after the number (*i.e.*, "*534RL*" becomes "*L5340*"). While the phenotypic prefix may rarely be changed to reflect new information about the effect of the *P* insertion, the designator is invariant. Thus, l(2)06253 and n(2)06253

TABLE 3
(Continued)

1. *Df(3L)emc24*	61C3-4;61E1-3	35. *Df(3R)2-2*	81F4-5;83A1-9
2. *Df(3L)Ar14-8*	61C4;62A8	36. *Df(3R)Dfd13*	83E3;84B1
3. *Df(3L)bab-PG*	61D3-E1;61F5-8	37. *Df(3R)Scr*	84A1;84B2
4. *Df(3L)R-G5*	62A10-B1;62C4-D1	38. *Df(3R)Antp17*	84A6;84D13-14
5. *Df(3L)Aprt32*	62B1;62E3	39. *Df(3R)p712*	84D4-6;85B6
6. *Df(3L)R-G7*	62B9;62E7	40. *Df(3R)by10*	85D8;85E10-13
7. *Df(3L)HR370*	63A1;63D1	41. *Df(3R)by62*	85D11-13;85F16
8. *Df(3L)1227*	63C1-2;63F1-2	42. *Df(3R)M-Kx1*	86C1;87B5
9. *Df(3L)GN24*	63F5-7;64C13-15	43. *Df(3R)kar-D1*	87A7;87D3-4
10. *Df(3L)ZN47*	64C1-10;65C1-5	44. *Df(3R)ry615*	87B12;87E11
11. *Df(3L)pb1-X1*	65F3;66B10	45. *Df(3R)red3l*	87F15;88C2
12. *Df(3L)66C-G28*	66B8-9;66C9-10	46. *Df(3R)red1*	88A2;88D3-4
13. *Df(3L)h-i22*	66D10;66E4-F5	47. *Df(3R)sbd105*	89A1;89B9-89B10
14. *Df(3L)AC1*	67A2;67D13	48. *Df(3R)bxd100*	89B6;89E2
15. *Df(3L)1xd6*	67E1-2;68C1-2	49. *Df(3R)C4*	89E3-4;90A1-7
16. *Df(3L)vin2*	67F2;68D6	50. *Df(3R)P14*	90C2;91B1-2
17. *Df(3L)vin5*	68A2;69A1	51. *Df(3R)Cha7*	90F1-90F4;91F5
18. *Df(3L)vin7*	68C8-11;69B4-5	52. *Df(3R)D1-BX12*	91F1-2;92D3-6
19. *Df(3L)Ly*	70A2-3;70A5-6	53. *Df(3R)H-B79*	92B3;92F13
20. *Df(3L)fz-GF3b*	70C2;70D6	54. *Df(3R)e-R1*	93B6;93D2
21. *Df(3L)fz-CAL5*	70C2-C6;70E1-2	55. *Df(3R)e-N19*	93B2-13;94A3-12
22. *Df(3L)fz-GS1a*	70D2;70E4-5	56. *Df(3R)e-BS2*	93C3;93F14
23. *Df(3L)fz-M21*	70D3;71E4-5	57. *Df(3R)hhE23*	94A1-16;94D1-4
24. *Df(3L)st-f13*	72C1;73A4	58. *Df(3R)crb87-4*	95D1-2;96A2
25. *Df(3L)st4*	72D10;73C1	59. *Df(3R)crb87-5*	95F7;96A18
26. *Df(3L)st-j7*	73A2;73B1-2	60. *Df(3R)XTA1*	96A17-21;96D1-2
27. *Df(3L)81k19*	73A3;74F1-4	61. *Df(3R)Espl10*	96F5-7;97B1
28. *Df(3L)W10*	75A6-7;75C1	62. *Df(3R)Tl-P*	97A1-10;98A1-A2
29. *Df(3L)Cat*	75B8;75F1	63. *Df(3R)3450*	98E3;99A6
30. *Df(3L)W4*	75B10;75C5-6	64. *Df(3R)01215*	99A6;99C1
31. *Df(3L)kto2*	76B1-2;76D5	65. *Df(3R)tll-e*	100A2-100C2-3
32. *Df(3L)rdgC-Ci2*	77A1;77D1	66. *Df(3R)awd-KRB*	100C6-7;100D3-4
33. *Df(3L)ri-79c*	77C1;77F1-5	67. *Df(3R)04661*	100D2;100F5
34. *Df(3L)Pc-MK*	78A3;78D1-2		

The names and cytogenetic breakpoints of the deficiency chromosomes used for genetic verification as summarized in Tables 4 and 5.

[a] This stock was assumed to be *Df(2L)ast2* rather than *Df(2L)ast1* as originally labeled, based on its complementation behavior; see FBab0001693.

[b] Breakpoints based on this study; no information available from FlyBase.

refer to a single BDGP strain, whose *P* insertion was initially thought to cause lethality, but was subsequently shown to cause no obvious phenotype. Because phenotypic prefixes can change, it is wise to search Internet databases using the designator. Periodically updated information on the BDGP strains can be obtained by searching the BDGP website at http://www.fruitfly.org/p_disrupt/, or from FlyBase (the Drosophila database project) at http://flybase.bio.indiana.edu/transposons/fbinsquery.hform.

Gene names: Symbols for Drosophila gene names are as given by FlyBase. For potentially novel loci defined only by a BDGP insertion strain, the name of the primary strain constitutes the provisional gene name, in accordance with FlyBase rules. Allele names for all the mutations are represented using the designator as the allele superscript. For example, because strain l(2)k10325 is part of the complementation group whose primary strain is l(2)03350 defining a new gene, its mutation is designated *l(2)03350*^{k10325}. The *P*-element mutation in strain l(2)s4771 that is allelic to *kismet* (*kis*) is designated *kis*^{s4771}. Again, because it is the designator that is presented in allele tables, it is wise to search FlyBase with the wild-carded designator.

Localization of inserts by *in situ* hybridization: *P* elements were localized by *in situ* hybridization to polytene chromosomes as described previously (Spradling *et al.* 1995); see also http://www.fruitfly.org/methods. Digitized images of these localizations are available at http://www.fruitfly.org/p_disrupt/. A few lines were localized by others; these were assumed to be less accurate and are given only to a polytene lettered section, rather than a range of specific bands. To reduce the number of *in situ* localizations, many alleles of seven known hotspots were removed from the Török *et al.* (1993) collection by complementing each starting strain with the following tester loci: 1(2)07815 (*kis*), l(2)01209 (*vkg*), l(2)04208 (*Eif4A*), l(2)02657 (*wg*), l(2)00255 (*bun*), l(2)00642 (*lola*), and l(2)03505 (*mam*). The insertion(s) in lines that

failed to complement were not localized, and they are not included in the tabulation of hotspot allele numbers. Consequently, the allele numbers for these loci are lower than would have otherwise been the case.

Complementation testing: Complementation crosses were carried out among single-insertion lines whose insertions were localized within six to eight polytene bands of each other. A two-stage strategy was used to limit the number of crosses and to minimize redundancy. Each line was first crossed to representative of any locus within range having multiple alleles. Lines failing to complement were identified as additional alleles and eliminated from further crosses. Lines not allelic to such local "hotspots" were subsequently crossed to representatives of the other complementation groups within the relevant zone. As soon as two complementation groups were joined, it was assumed that their behavior was uniform, and few additional crosses between the subgroups were carried out. Generally this strategy worked well. However, in a small number of cases, incomplete or inconsistent complementation behavior was observed due to localization errors larger than four to eight bands, to intergenic complementation, to semilethality, to inadvertent selection of a rearranged allele as the representative allele, to stock instability, or to errors in obtaining or recording complementation data. Problem complementation groups were reanalyzed on a case-by-case basis and the source of the contradiction resolved.

Verification: Strains from the primary collection were crossed to deficiencies (see Table 3) to verify that the *P* insertion caused the recessive phenotype. In 1717 single-insert strains, the cytogenetic locus of the *P* element clearly fell within the boundaries of existing deficiency (Df) chromosomes (Table 2). An uncertainty of four to six bands in the cytogenetic breakpoints was assumed, and the previous results of complementation tests with verified lines in the region were also considered (see Spradling *et al.* 1995). Complementation with deficiencies that unequivocally remove the *P* insertion site was taken as proof that the *P* element did not cause the associated phenotype. Failure to complement indicated that the strain was "verified." While lines with secondary mutations closely linked to the *P* insertion might be erroneously verified by this procedure, further molecular and genetic analyses suggest that the frequency of such errors is small. The results of the complementation and verification crosses are summarized in Tables 2, 4, and 5. The data are also available on the BDGP website (http://www.fruitfly.org/p_disrupt/).

The availability of DNA sequence information that can link insertion sites to nearby ESTs, transcripts, and predicted genes is expected to significantly change the way decisions to retain or discard lines are made. Except within the Adh region (Ashburner *et al.* 1999), we retained insertions only if they caused or were likely to cause a detectable mutant phenotype. However, in the future, as genomic sequences become more highly annotated, it will increasingly be possible to select strains solely on the basis of whether they are likely to disrupt a novel ORF, regardless of whether a recessive phenotype can be observed. In a few cases reported here, viable insertions reside near or within novel transcripts recognized by nucleotide sequence. The prefixes of these lines were changed to n(2) or n(3) to indicate the absence of a scorable phenotype. Only within the Adh region, where sequence annotation is now extensive (Ashburner *et al.* 1999), did a significant fraction of the retained lines lack strong phenotypes.

Flanking sequence determination: Flanking sequences from one or both ends of most *P*-element insertions in the primary collection were determined by one or both of two methods. Plasmids containing the 5′ *P* element and flanking genomic sequences were rescued from many strains. Prior to rescue, the line was expanded, and 40–100 adult flies were collected and frozen at −20°. The plasmid rescue procedure (based on Hamilton *et al.* 1991) entails macerating 30–40 flies in a grinding buffer, then one cycle of freeze-thaw, followed by a 20-min incubation at 70°. Subsequently, residual proteins and SDS were removed by addition of potassium acetate (KOAc) and incubation on ice for 30 min. The supernatant obtained after removal of particulate matter was ethanol precipitated to recover genomic DNA. Finally, the samples were treated with RNase A at 37° for 2 hr.

For plasmid rescue, a sample of genomic DNA equivalent to two to four flies was digested with an appropriate restriction enzyme (*e.g.*, *Xba*I for the PZ lines), then ligated at low DNA concentration to circularize the restriction fragments. Subsequently, DH10B cells were transformed by electroporation. The resulting colonies had acquired the circularized restriction fragment containing the selectable marker, the bacterial origin of replication, one *P*-element inverted repeat, and a variable amount of flanking genomic DNA. For each rescue, four to six transformants were screened by DNA miniprep and restriction digestion. In cases where at least three of the four (or five of the six) transformants exhibited identical patterns, a plasmid was chosen for sequencing that represented the major class. Occasionally, the appropriate plasmid was identified from a transformation experiment that yielded more than one plasmid form by *in situ* hybridization. These plasmids were sequenced directly using a primer designed to the *P*-element inverted repeat. The success rate in this procedure was ~80%.

The remaining lines were analyzed by recovering a smaller amount of DNA using inverse PCR according to the method of J. Rehm (http://www.fruitfly.org/methods/). This method was successfully adapted to a 96-well format where the success rate in obtaining 25 bp or more of flanking sequences has been >85%.

Association with ESTs: BDGP is generating a collection of 80,000 Drosophila EST sequences with support from Howard Hughes Medical Institute (accessible at http://www.fruitfly.org/EST/). During the preparation of this article, ~48,000 ESTs were available for comparison. Each flanking sequence was searched against this EST database, matches validated by inspection, and the position of the *P* insertion relative to the EST-homologous portion of the flanking sequence determined. The names of ESTs with strong matches are given in Tables 4 and 5. Only ESTs that were located within ~100 bp of the *P* element are reported; more distant sequence matches might represent adjacent transcripts and were not included in the tables.

Stock distribution: To hasten the availability of the gene disruptions, verified lines from the primary collection were sent to the Bloomington Stock Center in several batches beginning in 1993; the number of strains reached 700 by late in 1994. All 1052 primary collection strains have been available from the Bloomington Stock Center since October 1997. Reserve alleles are maintained at the Carnegie Institution (chromosome *3*) or at Berkeley (chromosome *2*), and have also been available on request since 1993. Information about stocks is updated periodically on the BDGP website and strains found to be inappropriate are removed from the Bloomington Stock Center. Information derived from further study of any of the BDGP stocks is welcome and should be forwarded to the corresponding author's e-mail address.

Statistical analysis of saturation: Previous attempts to esti-

TABLE 4

Chromosome 2 stocks

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence Verified?		Reserve	EST	Reference
l(2)k01206	21A1-4				1					
l(2)04207	21B2-3			1			Ver			
l(2)03350	21B4-6	4		1		G00736	Ver	l(2)k13601		
l(2)07812	21B4-6	27	*kis*	1	2	AQ073293	Ver	l(2)k13416		Verheyen *et al.* (1996)
l(2)k01901	21B4-6		*Rpp30*	1		AQ034135	Ver		HL08073	This study
l(2)k07612	21B4-6	4		1			Ver	l(2)k07721		
l(2)k06805	21B4-6				1	AQ025806				
l(2)k08102	21B4-6				1	AQ025855				
n(2)k10237	21B4-6	2			1	AQ025931	Ver	l(2)k16510		
l(2)k14504	21B7-8	2	*U2af38*		1,2	AQ026079	Ver	l(2)06751	LD07472	Rudner *et al.* (1996)
l(2)k11324	21C1-2				2	AQ025961				
l(2)06694	21C1-2		*α-Adaptin*	2	1,3	G00611	Ver		LD01019	Gonzalez-Gaitan and Jackle (1997)
l(2)k16513	21C1-2		*RpI135*	2		AQ034040	Ver			This study
l(2)k16213	21C2-3		*Tb11*	2		AQ026103	Ver		LD32876	This study
l(2)01270	21C4-5	5	*ex*	2	3	AQ073263	Ver	l(2)k12913		This study
l(2)05142	21C5-6			2		G00608	Ver			
l(2)k06506	21C4-5				2	AQ025798				
l(2)k07005	21C4-5	2			2	AQ034156	Ver	l(2)k08218		
l(2)k08915	21C5-6				2	AQ025875				
l(2)02858	21C6-7				1,2,3	AQ025601				
l(2)05486	21C6-7	2	*lwr*		2,3	G00739	Ver	l(2)01519	GM08125	FBrf0101086
l(2)k13714	21C7-8				2,3	AQ026066				
l(2)05341	21C7-D1		*Gsc*		2,3	AQ025636	Ver			Hahn and Jackle (1996)
l(2)01855	21D1-2			3		AQ073266	Ver			
ms(2)06619	21D1-4		*hsp60B*	3		AQ026417	Ver		GH05807	This study
l(2)04723	21D3-4	3	*dock*	3		G01444	Ver	l(2)k13421		Garrity *et al.* (1996)
l(2)k05428	21D4-E1	3	*dbe*	3		AQ025761	Ver	l(2)k00108	LD22189	This study
l(2)07056	21D2,3	10	*S*	3		G00465	Ver	l(2)k09530		Kania *et al.* (1995)
l(2)06955	21F1-2	2		3		G00453	Ver	l(2)k01217		
l(2)10685	21F1-2	3		3	4	G00626	Ver	l(2)k00420	GM06352	
l(2)k00619	21F1-2		*Dcap*	3		AQ034171	Ver		LD18894	This study
l(2)k11704	22A3-4		*RFeSP*	3,4		AQ034029	Ver		HL02717	This study
l(2)04111	22A5-6	4		3,4		G00530	Ver	l(2)k13009		
n(2)k09624	22B1-2				3,4	AQ025901				
n(2)k07918	22B6-7		*GlyP*		3,4	AQ025849			GM02594	This study
n(2)k09932	22C1-2				3,4	AQ025915			LD15963	
l(2)s5379	22D3-4	2			4,5	AQ026151	Ver	l(2)k08027	LD23816	
v(2)03953	22D1-2		*aop*	5	3,4	AQ025619	Ver			This study
l(2)k08232	22E1-2		*Rab5*	5	3	AQ025862	Ver		LD03788	This study
l(2)00231	22E2-3			5	4	AQ073257	Ver			
l(2)10638	22F1-4		*dpp*	5		G00760	Ver	l(2)k17036		Twombly *et al.* (1996)
l(2)k05909	23B1-2			5,6		AQ025774	Ver			
l(2)k16525	23B1-2			6			Ver			
l(2)03575	23B5-6	2	*oho23B*	6	5,7	AQ025612	Ver	l(2)k16814	GM13392	Török *et al.* (1993)
l(2)00632	23C1-2	2		6,7		AQ025583	Ver	l(2)k05431		
l(2)01361	23D1-2	4	*toc*	7,8	6	G01437	Ver	l(2)k08224	LD27161	This study
l(2)k00237	23D3-4	3	*Mad*	7,8	6	AQ034169	Ver	l(2)k05807	LD03112	This study
l(2)k10101	23F3-4		*Pdsw*		7,8	AQ025920			GM03559	This study
l(2)k07736	23F5-6		*Phas1*		8	AQ025845			HL08053	This study

(*continued*)

TABLE 4
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence Verified?		Reserve	EST	Reference
l(2)01863	24A1-2	2	*odd*		8,10		Ver	l(2)rF111	GH01449	Rauskolb *et al.* (1995)
l(2)06860	24A1-2	5	*for*	10		AQ073291	Ver	l(2)k04703	LD08322	This study
v(2)k08012	24A1-2		*Dot*			AQ025853	Ver			Rodriguez *et al.* (1996)
l(2)k08617	24C1-2		*bowl*		10	AQ025865				This study
l(2)k16918	24C8-9				10	AQ026119			GH05923	
l(2)05965	24C8-D1		*slp1*	9,10		AQ025644	Ver			Park *et al.* (1996)
l(2)k01102	24D3-4			10		AQ034173	Ver		GM09285	
l(2)01085	24E1-2	2		9	10	G00578	Ver	l(2)k14703	HL01565	
l(2)k08903	24F1-2			9		AQ034018	Ver			
l(2)k10004	25B1-2					AQ025918				
l(2)k10217	25B1-2					AQ025930				
l(2)05714	25B4-6				9	AQ025642			LD03394	
l(2)01209	25C1-2	13	*vkg*		11,12	AQ073262	Ver	l(2)k00236		Yasothornsrikul *et al.* (1997)
l(2)k00405	25C1-2	6	*Cg25C*		11,12	AQ025687	Ver	l(2)k03009	GM04010	This study
l(2)k09003	25C1-2	3	*eIF-3*		11,12	AQ025877	Ver	l(2)k16615	LD05962	This study
l(2)k10127	25C1-2					AQ025926				
l(2)k11206	25C5-6				11,12	AQ025954				
l(2)k01302	25D1-2			11	12	AQ025703	Ver		LD23535	
l(2)04415	25D1-2		*tkv*	11,12		AQ073280	Ver			George and Terracol (1997)
l(2)k05901	25D4-5	4	*vri*	12		AQ034151	Ver	l(2)k09602		George and Terracol (1997)
l(2)03771	25D4-6			11	13,14	AQ025616	Ver			
l(2)k11511	25E5-6	3	*Lam*	13	11,14	AQ034028	Ver	l(2)04643	LD10531	Petersen *et al.* (1997)
l(2)k06502	25F3-4	4		12,14		AQ025796	Ver	l(2)02839	LD16669	
ms(2)04875	26A1-9		*ifc*	13			Ver			Endo *et al.* (1996)
l(2)k13321	26A5-6	13	*chi*	13		AQ026056	Ver	fs(2)01320	LD08034	Cooley *et al.* (1992)
l(2)10424	26A8-9	2		13		G01406	Ver	l(2)k06801		
l(2)02439	26B1-2	11	*eIF-4a*	12,13,14		G01428	Ver	l(2)k01501		Dorn *et al.* (1993)
l(2)10642	26B8-9	5	*Kr-h*		12,13,14	G00625	Ver	l(2)k04411		Roch *et al.* (1998), this study
l(2)k13720	26C2-3				14	AQ026067				
l(2)k07502b	26D1-2				14	AQ025833			HL02956	
l(2)k09923	26D1-2	3			14	AQ034023	Ver	l(2)k09847		
l(2)k04917	26D6-8	5				AQ025751	Ver	l(2)k05435		
l(2)k14206	26F3-5				15,17	AQ026076				
fs(2)01355	27A	3	*cup*				Ver	fs(2)06890		Keyes and Spradling (1997)
l(2)k00605	27A1-2				15	AQ025688			GM14348	
l(2)k13315	27B1-2			15		AQ026055	Ver			
l(2)k04223	27B1-2			15			Ver			
l(2)03300	27B4-C1		*Rca1*	15	17,18	AQ025609	Ver		LD13031	Dong *et al.* (1997)
l(2)k09022	27C1-2			15	16,17,18	AQ025880	Ver			
l(2)k00230	27C4-5			17	16,18	AQ025682	Ver			
l(2)02647	27C4-5	6	*Hrb27C*	15,16,17	18	AQ073272	Ver	l(2)k02814		Campbell *et al.* (1995)
l(2)k10617	27C6-8	2	*Coprox*	16	17,18	AQ025939	Ver	l(2)k11018	LD07292	This study

(*continued*)

TABLE 4
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence Verified?		Reserve	EST	Reference
l(2)04493	27C7-8	8		15	17	AQ025627	Ver	l(2)k09603		
l(2)02107	27D1-2		*snRNP27D*	16	18	AQ025594	Ver		GM08995	Mancebo *et al.* (1990)
l(2)k04704	27D5-6	2		15,16,18		AQ025746	Ver	l(2)k06704		
l(2)02657	27F1-2	6	*wg*	19	15	AQ073273	Ver	l(2)04894		Mathies *et al.* (1994)
l(2)00434	27F1-2			16,18,19			Ver			
l(2)k00202	27F4-5	2		18,19			Ver	l(2)k04301a		
l(2)10607	27F4-6	5		16,18	15	AQ073299	Ver	ms(2)05158		
l(2)k10113	27F4-6			16,18	19	AQ025925	Ver			
l(2)k09238	28A1-2			18	19		Ver			
l(2)k10609	28B1-2				21	AQ025938				
l(2)k05404	28C7-9			21	16,18,19		Ver			
l(2)02496	28D1-2	4	*mts*	21	18,19	G01458	Ver	l(2)s5286	LD12341	Wassarman *et al.* (1996)
l(2)k10210	28D7-9					AQ025928			LD30420	
l(2)05836	28E1-2	3			21,23	AQ025643	Ver	l(2)s1883		
ms(2)01659	28E1-9		*poe*			AQ026403				Castrillon *et al.* (1993)
l(2)06243	28E3-4	8	*gel*	23	21	AQ025647	Ver	l(2)k13638	LD01834	This study
l(2)k14308	28F1-2			23		AQ026077	Ver			
l(2)k00206	29A1-2	2	*Btk29A*	23		AQ034168	Ver	l(2)k05610		Roulier *et al.* (1998)
l(2)k09614	29B1-2	2	*RpS13*		23,24,25	AQ025899	Ver	l(2)k15708		This study
l(2)k14902	29B1-2				23	AQ026085				
l(2)k01105	29C1-2				24,25	AQ025697				
l(2)k12914	29C1-2	2			24,25	AQ026046	Ver	l(2)k14509		
l(2)k07118	29C1-3				23,24	AQ025816				
l(2)k16715	29C3-4	2		24	23,25	AQ026113	Ver	l(2)k06303		
l(2)01482	29C3-5				23,25					
l(2)k07704	29D1-2		*Acer*		25	AQ025842			LD17687	This study
l(2)03424	29D4-5		*Orct*	24,25		AQ025610	Ver		LD20119	Taylor *et al.* (1997)
l(2)k05125a	29D1-2			25			Ver			
l(2)k13702	29E1-2	2	*sema-I*	25		AQ026063	Ver	l(2)k03509	HL03652	Yu (1998)
l(2)rH280	29E1-2			25		AQ026132	Ver			
l(2)k04003	29E3-4			25		AQ025734	Ver			
l(2)06825	29F1-2	6		25	26	AQ025659	Ver	l(2)k01021		
ms(2)07717	29F1-8		*ms(2)29F*			AQ026418				Castrillon *et al.* (1993)
l(2)s2978	29F8-A1			25	26	AQ026145	Ver			
l(2)01351	30A3-5	6		25,26		G00546	Ver	l(2)k15101		
l(2)k05809	30A3-6			26		AQ025769	Ver			
ms(2)05289	30B1-12		*scat*	26		AQ026412	Ver		LD22446	Castrillon *et al.* (1993)
l(2)03235	30B5-6	3	*numb*	26		G00599	Ver	l(2)s2201		Uemura *et al.* (1989)
l(2)k05113	30B5-6			26			Ver			
l(2)01272	30C1-2	13	*Pka-C1*	26		G01420	Ver	l(2)k00804		Lepage *et al.* (1995)
l(2)k07104	30C1-2	4	*hoip*	25,26		AQ034157	Ver	l(2)k07607	GH03082	Kania *et al.* (1995)
ms(2)01559	30C1-9		*pelo*	26		AQ026402	Ver			Eberhart and Wasserman (1995)
ms(2)07822	30C1-9		*ms(2)30C*		26					Castrillon *et al.* (1993)
l(2)k02506	30C7-8		*zf30C*	25	26	AQ025709	Ver		LD19288	This study
l(2)k14204	30D1-2				25	AQ026075				
l(2)06320	30D3-4				26					
l(2)k09010	30D3-4	2			25	AQ025879	Ver	l(2)k08408		
l(2)02695	30E1-2					AQ025597			LD07208	

(*continued*)

TABLE 4
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence Verified?	Reserve	EST	Reference
l(2)08014	30E1-2					AQ025666			
l(2)k01215	30E1-2	3	*sop*			AQ025701 Ver	l(2)k06507	LD24077	This study
l(2)k13305	30F5-6					AQ026052			
l(2)k14401	31A1-2		*Pen*			AQ026078 Ver			Kussel and Frasch (1995)
l(2)k06607	31B1-2		*me31B*		27	AQ025801		GM01268	This study
l(2)k09310	31C1-2			27		Ver			
l(2)08774	31D1-2	3	*Msp1*	27		AQ025668 Ver	l(2)k10909	LD20618	This study
l(2)k06709	31D8-9	2	*RnrL*	27		AQ034154 Ver	l(2)k13717	LD06546	This study
l(2)k00311	31E1-2		*KdelR*	27		AQ025685 Ver		LD06574	This study
l(2)k16801	31E1-2			27		Ver			
l(2)04820	31E3-5		*RpS27A*	27		AQ034164 Ver			This study
l(2)k10307	31F4-5	3	*Fatp*	27		AQ034026 Ver	l(2)k10801	GH04319	This study
ms(2)04818	32A1-5		*dbf*	27		AQ026409 Ver			FBrf0064394
l(2)k02605	31F3-4				27	AQ025713			
l(2)k09116	31F4-5				27	AQ025886			
l(2)k13206	32A4-5		*UbcD2*			AQ026047		LD17255	This study
l(2)k05123	32B1-2	2	*porin*		27,28	AQ034146 Ver	l(2)k08405	GH11331	This study
fs(2)06843	32BC	2	*piwi*			Ver	fs(2)P-1		Lin and Spradling (1997)
l(2)03788	32C1-2				27				
l(2)k03107	32C1-2	2				AQ025718 Ver	l(2)k13811		
l(2)06225	32C4-5					AQ025646		GH12084	
l(2)k05812	32C4-5					AQ025770			
l(2)k15817	32D1-2				28	AQ026093			
l(2)01501	32D1-2	4	*Nup32D*	28		G00759 Ver	l(2)k07717	LD21772	Gigliotti *et al.* (1998)
n(2)k13807	32D4-5		*Rfc38*		28	AQ026070		LD13549	This study
l(2)04008	32E1-2			28		G01158 Ver			
l(2)04431	32E1-2			28		AQ073281 Ver			
l(2)k07716	32E1-2			28		AQ025844 Ver		LD03334	
l(2)03602	32F1-2		*salm*	28		G00737 Ver			Treisman and Rubin (1996)
l(2)04418	33A1-2	9	*crol*	28		G00541 Ver	l(2)k05205		D'Avino and Thummel (1998)
l(2)08307	33A3-7			28		Ver			
l(2)01810	33B8-12			28		AQ025589 Ver			
ms(2)01284	33C1-D5		*aret*	28		AQ026400 Ver			Castrillon *et al.* (1993)
l(2)k04203	33C4-5	2	*Aats-thr*	28,29		AQ034143 Ver	l(2)k04910	GM10740	This study
l(2)08323	33D1-2	2	*Rab6*	28,29		G00615 Ver	l(2)k13606	GM04112	This study
l(2)04518b	33D1-2			28		Ver			
l(2)k06909	33E5-7		*Elf*	28		AQ034155 Ver		GM07765	This study
l(2)00255	33E7-8	20	*bun*	28		G00410 Ver	l(2)k06713		Treisman *et al.* (1995b)
l(2)01433	34A1-2			28,29		G00461 Ver			
n(2)k07332	34A1-2					AQ025829 Ver		GM02380	
l(2)rK639	34A1-2			28,29		AQ026136 Ver			
l(2)01510	34A1-2	5	*Vha68-2*	29	28	AQ073265 Ver	l(2)s4214		Sozen *et al.* (1997)
l(2)k17004	34A3-4				28	AQ026123			
l(2)05327	34A5-6				28				
l(2)k10105	34A5-6				29	AQ025922		LD18129	
l(2)k11328	34A5-6				29	AQ025962			
l(2)k05911	34B1-2					AQ025775			
l(2)k07826	34B6-7	2	*Nnp-1*			AQ034006 Ver	l(2)k08704	LD07345	This study
l(2)k07914	34B6-7					AQ025848			
l(2)k00302	34B8-9					AQ025684			

(*Continued*)

TABLE 4
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence Verified?		Reserve	EST	Reference
n(2)05337	34C1-2		*B4*			AQ025634			LD07101	Sotillos *et al.* (1997)
l(2)k01403	34C4-5	7	*kuz*			AQ073300	Ver	l(2)k09934		Rooke *et al.* (1996)
n(2)k07245	34D1-2				30	AQ025825				Ashburner *et al.* (1999)
v(2)k05524	34D4-6	3	*Sos*	30		AQ025755	Ver	l(2)k05705	GH01796	Ashburner *et al.* (1999)
l(2)k05605	34D6-7		*RpII33*	30		AQ025764	Ver		LD05121	Ashburner *et al.* (1999)
n(2)06646	34E1-2					AQ025655				Ashburner *et al.* (1999)
n(2)k09909	34F1-2				30	AQ025909				Ashburner *et al.* (1999)
l(2)k00811	35A1-2		*l(2)34Fa*	30		AQ025693	Ver			Ashburner *et al.* (1999)
v(2)k11509	34F3-4		*smi35A*			AQ025968	Ver			Ashburner *et al.* (1999)
l(2)09437	35A1-2	7	*wb*			G00417	Ver	l(2)k13507		Ashburner *et al.* (1999)
n(2)k08712	35A3-4		*Rab14*			AQ025867			LD41067	Ashburner *et al.* (1999)
n(2)k07706	35B1-2		*elB*			AQ025843				Ashburner *et al.* (1999)
v(2)rJ571	35B1-4		*osp*	30	32	AQ026134	Ver		LD15891	Treisman and Rubin (1996)
l(2)k11524	35B6-7		*l(2)35Bb*			AQ026036	Ver		LD16050	Ashburner *et al.* (1999)
l(2)k08808	35B6-10		*l(2)35Bc*			AQ025869	Ver			Ashburner *et al.* (1999)
l(2)10408	35B8-9		*l(2)35Bd*	30,32		AQ073297	Ver		LD09819	Ashburner *et al.* (1999)
l(2)k10011	35B8-10		*l(2)35Bg*	30,31,32		AQ025919	Ver		GM10279	Ashburner *et al.* (1999)
l(2)k07904	35B8-9		*Su(H)*	30,31,32		AQ025847	Ver		LD24729	This study
l(2)07130	35C1-2		*ck*	31,32		AQ025664	Ver		LD10736	This study
fs(2)00424	35B		*vas*	31			Ver			This study
l(2)05441	35C1-2	2	*stc*	32		G00414	Ver	l(2)k11112		Stroumbakis *et al.* (1996)
l(2)06430	35D1-4		*gft*	31,32		G00415	Ver		LD10516	Ashburner *et al.* (1999)
ms(2)02316	35D1-2		*ms(2)35Ci*			AQ026404	Ver			Ashburner *et al.* (1999)
l(2)07082	35D1-2	14	*esg*	32		G00416	Ver	l(2)k00606		Samakovlis *et al.* (1996)
l(2)k05305	35D3-4	3	*lace*			AQ034147	Ver	l(2)k02303	LD17449	Ashburner *et al.* (1999)
l(2)05206	35D3-4	6	*CycE*	33	31	G00412	Ver	l(2)k05007		Knoblich *et al.* (1994)
l(2)k14423	35D5-7		*l(2)35Df*			G01316	Ver		LD02559	Ashburner *et al.* (1999)
n(2)k09033	35D6-7		*Gli*			AQ925882				Ashburner *et al.* (1999)
n(2)05271	35D1-2		*l(2)35Ea*							Ashburner *et al.* (1999)
l(2)k09834	35F1-5		*PRL-1*			AQ025903	Ver			Ashburner *et al.* (1999)

(*continued*)

TABLE 4
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence Verified?		Reserve	EST	Reference
ms(2) k08310	35F1-2		*twe*				Ver			Ashburner *et al.* (1999)
l(2)00232	35F1-2	11	*crp*	33	34	G00409	Ver	l(2)k00809		Ashburner *et al.* (1999)
l(2)k11403	35F6-7	3	*heix*	33		AQ025963	Ver	l(2)k12401	LD08373	Ashburner *et al.* (1999)
n(2)k17003	35F10-11		*cactus*			AQ026122				Ashburner *et al.* (1999)
l(2)k14608	35F11-12		*l(2)35Fe*			G01317	Ver			Ashburner *et al.* (1999)
n(2)k04216	35F11-12		*chif*			AQ025739				Ashburner *et al.* (1999)
l(2)rK364	36A1-2		*dac*	34		AQ026135	Ver			Mardon *et al.* (1994)
ms(2)04445	36A1-B6		*bln*		33	AQ026408				Castrillon *et al.* (1993)
fs(2)06034	36A1-10	7	*grp*	35			Ver	fs(2)02257		Sullivan *et al.* (1993)
l(2)k13905	36A10-11			34,35	33	AQ026073	Ver			
l(2)k15102	36A10-11			34,35	33	AQ026086	Ver			
l(2)k08819	36A12-14		*glu*	34,35	33	AQ025873	Ver	l(2)k06821	LD20207	Kania *et al.* (1995)
l(2)k03902	36B1-2		*Cas*	34,35		AQ025732	Ver		LD14270	This study
l(2)k10423	36B1-2		*Mhc*	34,35		AQ025936	Ver			This study
fs(2)01313	36C	2	*dl*				Ver	fs(2)k10816		This study
l(2)k06710	36C8-11	2	*Aac11*	35	36	AQ025805	Ver	l(2)k07112	LD09852	This study
l(2)04553	36E1-4		*RpS26*	36	35	AQ025629	Ver			This study
fs(2)neo2	36EF		*kel*				Ver			Xue and Cooley (1993)
l(2)k13805	37A1-2			36,37			Ver			
l(2)k05402	37B7-10		*Catsup*	37			Ver			P. F. Lasko, personal communication
l(2)k05424	37B8-12			37		AQ025760	Ver		LD23513	
l(2)01265	37B8-9	7		37		G00579	Ver	l(2)k00308		
l(2)02660	37B8-9			37			Ver			
l(2)k16106	37B8-9		*l(2)37Db*	37		AQ026099	Ver		LD20470	P. F. Lasko, personal communication
l(2)k02104	37C1-2		*Ddc*	36,37		AQ025704	Ver			This study
l(2)k06028	37C6-7	2	*brat*	37		AQ025779	Ver	l(2)k1l538	LD16270	This study
l(2)01068	37F1-2	6	*spi*	37	39,40	G00577	Ver	l(2)s3547		Perrimon *et al.* (1996)
l(2)k08115	38A5-6	2	*fs(2)ltoPP43*	37		AQ025856	Ver	l(2)k15716	LD13084	P. F. Lasko, personal communication
l(2)k10239	38A7-8			38	37	AQ034025	Ver			
l(2)k14014	38B1-2	7	*barr*	38,39		AQ034033	Ver	l(2)k08103		Bhat *et al.* (1996)
l(2)03552	38B4-6	5	*neb*	38,40		G00600	Ver	l(2)k05702		Ruden *et al.* (1997)
l(2)02306	38E1-2		*Hr38*	40		G00581	Ver			Kozlova *et al.* (1998)
l(2)k07135	38E5-6	2	*dia*			AQ025819	Ver	ms(2)04138	LD16963	Castrillon *et al.* (1993)
l(2)01528	38F3-4	4		38,40	39		Ver	l(2)k08613		
l(2)04530	38F5-6			40	37	G00584	Ver			
l(2)05287	39A1-2	3		38,40		G00609	Ver	l(2)k16804B	LD28636	
l(2)07054	39B1-2	9	*AconM*	38	39	AQ025661	Ver	l(2)k02301	LD24561	This study
l(2)10523	39B1-2	4	*bur*	38,40		G00613	Ver	l(2)k07130	LD10169	Neufeld *et al.* (1998)
l(2)k07215	39B1-2	2		38,39		AQ034159	Ver	l(2)k16009	LD13720	
l(2)k09410	39B1-2	2	*snRNA:U4:39B*	38		AQ034021	Ver	l(2)k06410		This study
l(2)k05106	39C1-2	2		39		AQ034145	Ver	l(2)k08036		

(*continued*)

TABLE 4
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence Verified?	Reserve	EST	Reference
l(2)k11226	39C1-2			38,40		AQ025956 Ver			
l(2)k14505	39C1-2			38,40		AQ034036 Ver		GM08665	
l(2)05095	39E1-2				38	AQ025632			
l(2)k05815	39E3-4	2	*anon1A4*		38,40	AQ034150 Ver	l(2)k05501	LD15206	This study
l(2)k06113	39E5-6	2			38,40	AQ025782 Ver	l(2)k08034		
l(2)02074	39F1-2	2		38,40		G00597 Ver	l(2)05230		
l(2)03832	39F1-2	2		38,40		G01434 Ver	l(2)k05429		
l(2)k08110	39F1-3	2		40		AQ034012 Ver	l(2)k11104		
l(2)k16406	40A1-2			38		AQ034039 Ver			
l(2)04319	40A1-4		*tsh*	38		G00738 Ver			FBrf0091269
l(2)rA135	40A1-4			38		AQ026130 Ver			
l(2)k03002	41C	2			37,41	Ver	l(2)k04601		
l(2)02047	41C1-6			41		G00522 Ver	l(2)k13422a		
l(2)07022	41F8-9	2		41–43	44	G01159 Ver	l(2)k01109	LD32728	
ms(2)06410	42A1-19		*ms(2)42A*		44	AQ026415		LD17664	Castrillon *et al.* (1993)
l(2)09851	42A1-2			42,43	41	AQ025672 Ver		GM02153	
l(2)k06109	42A1-2	2	*Bub1*	42,43	41	AQ025781 Ver	l(2)k03113	LD24007	This study
l(2)k10108	42A1-2		*Src42A*	42,43	41	AQ025923 Ver		LD15045	This study
l(2)k09848	42A8-12			42,43		AQ025905 Ver			
l(2)k06210	42A10-12	3	*EcR*	42,43		AQ025786 Ver	l(2)k04504		FBrf0086762
l(2)k14710	42A15-19	2		42,43		Ver	l(2)k15603	LD26521	
l(2)k14805	42B1-2			43	42,44	AQ026082 Ver			
l(2)01094	42B1-3	9		43	42,46	G00529 Ver	l(2)k02002		
l(2)k03204	42B1-3			43	44,46	AQ025721 Ver			
l(2)04535	42C1-2	7		43,44	47	Ver	l(2)k02710		
l(2)k14019	42C1-2	2		43,44,45	46	AQ034034 Ver	l(2)k03202		
l(2)01349	42C1-4	2	*Adf-1*	43–46	47	AQ025586 Ver	l(2)k09919	LD09689	This study
v(2)k09107	42C1-2				43,44	AQ025885		LD07974	
l(2)04065	42C1-2			44	43,46,47	AQ025622 Ver	l(2)k00620		
l(2)03055	42C6-9			46		Ver			
l(2)01289	42C8-9	2		43–46	48	G00460 Ver	l(2)07769		
l(2)k03203	42D1-3			43–46	47	AQ034141 Ver			
n(2)09967	42D1-6				43,46,47	AQ026420		LD11166	
l(2)k08011	42D4-5			43,45		AQ025852 Ver			
l(2)k16722	42E3-4	2	*vimar*	43,45	44,48	AQ034041 Ver	l(2)k10203	LD07418	Lo and Frasch (1998)
l(2)04524	42D1-2		*Eb1*	43,45,50	47,48,49	AQ024628 Ver		LD08743	This study
l(2)04614	43B1-2		*l(2)43Bb*	47,50		AQ073285 Ver			FBrf0086245
l(2)05518	43B1-2		*l(2)43Bc*	47,48,50	49	AQ034165 Ver			FBrf0086245
l(2)k16101	43B1-2		*cos*	45,48,50		AQ034038 Ver		LD15871	Sisson *et al.* (1997)
l(2)03610	43D1-2	2	*Dhap*	45,47,50	48,49	G00524 Ver	l(2)05467	LD02207	This study
l(2)k08815	43D1-4	2	*dpld*	45,50	48	AQ025871 Ver	l(2)k14202	LD19006	Rodriguez (1996)
l(2)03427	43E1-5	2	*Ani*	47,50		G00563 Ver	l(2)k09008	LD23741	This study
l(2)01857	43E4-5			47		AQ025591 Ver			
l(2)05643	43E4-6	3		45,47,49	51	AQ025641 Ver	l(2)k11110		
l(2)08492	43E9-13	2		45,47,49,50	51	G01429 Ver	l(2)k10320	GM08726	
l(2)k07619	43E15-16			45	47,50	AQ025839 Ver		LD05439	
l(2)k01207	43F1-2		*lin19*	45,51	50,52	AQ025699 Ver		LD20253	This study
l(2)k07624	43F1-2	2	*rnh1*	45,51	50,52	AQ025841 Ver	l(2)k07409	LD20030	This study
l(2)k08018	44A1-2		*Cops4*	45,51	52,53	AQ034010 Ver		LD11968	This study
l(2)k08504	44A1-2			45	51,52	AQ034015 Ver			
l(2)k16503	44A4-5				45,51,52	AQ026107		GH04443	
l(2)s9998	44A4-5				45,52				
l(2)k16120	44B5-6		*Adk2*		45,51–53	AQ026101			
l(2)02045	44B5-9				45,51–53	AQ025593			
l(2)k07521	44C1-2	2			52,53	Ver	l(2)k15216		

(*continued*)

TABLE 4
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence	Verified?	Reserve	EST	Reference
l(2)k09514	44C1-2		*Drs1*		45,52,53	AQ025895			GM05306	This study
l(2)02502	44C1-2	3	*pnut*	53	45,49,51	AQ073270	Ver	l(2)rN498	LD33747	Neufeld and Rubin (1994)
l(2)k03110	44C1-2	3		53	45,52,54	AQ034140	Ver	l(2)k04002		
fs(2)02465	44D		*ptc*				Ver			Forbes *et al.* (1996)
l(2)k08904	44D4-5			52,53,54		AQ034019	Ver			
l(2)k02507	44D5-6	3	*rubr*	52,53,54		AQ025710	Ver	l(2)rH075		Kania *et al.* (1995)
l(2)s1878	44D5-6	2		53	52,54		Ver	l(2)rN173		
l(2)05847	44E1-2	5		53,54	52	G00494	Ver	l(2)k10313		
l(2)k02107	44E1-2			54			Ver			
l(2)k05304	44F1-2			53,54			Ver			
l(2)k04913	44F1-2			54	52,53,55–57	AQ025750	Ver			
l(2)03996	44F3-4					AQ025620				
l(2)k08017	44F3-4		*Ggamma1*	54	53,55,56	AQ034009	Ver		LD03453	This study
l(2)k16109	44F3-4	2	*Dmn*	54	55,56,57	AQ026100	Ver	l(2)k16218	LD07994	This study
l(2)k16912	44F11-12	2	*babo*	54,55,56	53,57	AQ034042	Ver	l(2)k07737		Brummel (1999)
l(2)03697	45A4-8			54,56,57	58	AQ025614	Ver		LD13319	
l(2)k00116	45A4-8	2		56	58	AQ025679	Ver	l(2)k00413		
l(2)k11201	45B1-2			54,58		AQ025953	Ver			
n(2)k04512	45B1-2				56	AQ025744			HL05962	
1(2)k13412	45B1-2			54,57,58	56	AQ026058	Ver			
l(2)06736	45C1-2			54,57,58	56	AQ025657	Ver			
l(2)k05611b	45C1-2			54,57	56		Ver			
v(2)rG232	45C3-4	2		55,57			Ver	v(2)k11209		
l(2)03659	45D1-2		*hspr*	57,58	54,55,56	AQ025613	Ver		LD06376	This study
l(2)k16806	45D4-5			57	54–56,58	AQ026117	Ver			Zhang *et al.* (1997)
l(2)k09507	45D4-5	3	*wun*	57	55,56	AQ025894	Ver	l(2)k10201	GH02203	Zhang *et al.* (1997)
l(2)03497	45D4-5				56,57,58	AQ025611				
l(2)k12402	45D4-5				57	AQ026044				
l(2)06424	45D5-6				56,57,58					
l(2)k01301	45F1-2					AQ025702				
l(2)k08914	45F1-2				57					
l(2)k10213	45F1-2					AQ025929				
l(2)k17035	45F1-2									
l(2)k09501	45F4-5					AQ025893				
l(2)02353	46A1-2	8			56,57	G00598	Ver	l(2)k00604		
l(2)03405	46A1-2	9	*Uba1*		60	G00562	Ver	l(2)s3484	LD20374	This study
l(2)04454	46B1-2	3	*dap*		55-60	AQ073282	Ver	1(2)k07309	LD11071	de Nooij *et al.* (1996)
l(2)k03111	46B1-2	2			59,60	AQ025719	Ver	l(2)k02003		
l(2)k09221	46B1-2					AQ025891				
l(2)k05420	46B4-5					AQ025758				Kania *et al.* (1995)
ms(2)05704	46C1-12		*Ft1*			AQ026413			HL08032	This study
l(2)k07237	46C1-2				59,60	AQ025824				
l(2)k08816	46C1-2				59,60	AQ025872			GH07336	
v(2)k06408	46C7-8			59		AQ025793	Ver			
l(2)k08601	46C6-8			59	60	AQ034016	Ver			
l(2)03775	46D1-2	2	*Vcp*	59,60		AQ025617	Ver	l(2)k15502	LD15631	This study
l(2)k13906	46D1-2			59,60		AQ034032	Ver		GH12681	
l(2)k07103	46E4-8		*14-3-3zeta*	59	60,61	AQ025662	Ver			Kockel *et al.* (1997)
l(2)06339	46F1-2		*Pfk*	59	61	AQ025651	Ver		HL03554	This study
l(2)k03610	46F1-2	2		59	62	AQ025727	Ver	1(2)k03703		
l(2)k04308	46F1-2			59	61,62	AQ025742	Ver			
l(2)k16104	46F1-2			59	62		Ver			
l(2)k10308	46F5-6	4	*Hr46*	59,61		AQ025932	Ver	l(2)k01017		Rottgen *et al.* (1998)
l(2)k07703	46F9-10		*Syb*	59,61,62		Ver				This study
l(2)k05201	47A3-5			61,62	59	AQ025753	Ver		LP04652	
l(2)00642	47A11-12	24	*lola*	62,63		AQ073258	Ver	l(2)k09901		FBrf0086256

(*continued*)

TABLE 4
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence	Verified?	Reserve	EST	Reference
fs(2)j13b6	47B	4	*psq*				Ver	fs(2)02403		Siegel *et al.* (1993), Horowitz and Berg (1996)
l(2)10565	47B7-8			62			Ver			
l(2)10425	47C3-4			62,63	61	AQ025674	Ver			
l(2)04738	47E1-2	7	*shn*	61	62	G00605	Ver	l(2)k00401		Arora *et al.* (1995)
l(2)k06103	47F1-2		*Fpps*	63	62,64	AQ034152	Ver		LD24632	This study
v(2)k03514	47F1-2				63	AQ025725				
l(2)k17005	47F4-9		*Tap&dgr;*		63,64	AQ026124			GM08470	This study
l(2)k15001	48A		*rack*							Kania *et al.* (1995)
l(2)k14708	48A3-5			64	63,65	AQ034037	Ver		LP02372	
l(2)k14602	48B1-2			64			Ver			
l(2)k06524	48B1-2		*oho48A*		64	AQ025799	Ver			Török *et al.* (1993)
l(2)02516	48C1-2	2	*Etf*		64	AQ025595	Ver	l(2)k14026	LD07532	This study
l(2)01275	48C5-6	8	*Efl&agr;48D*			G00459	Ver	l(2)k06102	GM03706	This study
l(2)k03905	48C5-6					AQ025733				
l(2)k06612	48D1-2				66	AQ025802				
l(2)k13312	48D5-6				66					
l(2)k05644	48E1-2	2			66	AQ034149	Ver	l(2)k13306		
l(2)02833	48E4-5	5	*guf*			AQ025599	Ver	l(2)k11803	GM14566	Salzberg *et al.* (1996)
l(2)06444	48E10-11	3	*Cct5*	65		AQ025653	Ver	l(2)k06005	GM12270	This study
l(2)03909	48F1-2	6	*Cam*		66	AQ073279	Ver	l(2)k04213		Harvie *et al.* (1998)
l(2)k11404	48F10-11				65,66	AQ025964				
l(2)k01103	48F3-6					AQ025696				
l(2)k17040	49A7-11				66	AQ026129				
l(2)k03003	49B1-2			66		AQ025717	Ver			
l(2)08269	49B3-6	7	*Sin3A*	66	67	G00542	Ver	l(2)k02703		Neufeld *et al.* (1998)
fs(2)k09833	49B7-8						Ver			
l(2)k05316	49B7-11			66	68	AQ034148	Ver			
l(2)k15501	49BC		*unch*	66,68		AQ026088	Ver		LD13852	Kania *et al.* (1995)
ms(2)00815	49C1-4		*ox*	66		AQ026399	Ver		GH03312	Castrillon *et al.* (1993)
l(2)04329	49D1-3			66,68	67	G00564	Ver			
l(2)k09328	49D1-3			67	66,68	AQ034020	Ver			
l(2)k10712	49D5-6		*Btf*	67,68		AQ034027	Ver		GM05329	This study
l(2)k05722	49E1-2			66,67			Ver			
l(2)k04508	49E1-2	3		66,67,68			Ver	l(2)k05408		
l(2)01424	49E1-2			66,68	67	G00733	Ver			
l(2)k06344	49E6-7		*su(z)2*	67,68	66	AQ034153	Ver			This study
l(2)k07834	49E6-7		*Psc*	67,68	66,72	AQ034007	Ver			C. Russell, personal communication
l(2)03531	49F7-8			67	66	G00583	Ver		LD07176	
l(2)k14804	49F7-8	4	*Aats-val*	67		AQ026081	Ver		GM09906	This study
l(2)10626	50A12-14	6	*drk*	68	67	G00788	Ver	l(2)k02401	GH14963	Simon *et al.* (1993)
l(2)05488	50C1-2		*fas*	68		AQ025638	Ver		HL03670	Lekven *et al.* (1998)
l(2)k10814	50C3-4			68		AQ025941	Ver			
v(2)k16105	50C4-5			68		AQ026098	Ver		LD33796	
l(2)k08708	50C6-7			68		AQ025866	Ver			
l(2)k03010	50C9-10	4		68		AQ034139	Ver	l(2)k03405		
l(2)k04204	50C11-12	2		68		AQ025736	Ver	l(2)k05803		
l(2)04845	50C12-21			68	69,70	AQ025630	Ver			
l(2)k08121	50C14-15	2		68		AQ025857	Ver	l(2)k16601	LD09501	
v(2)k15606	50C17-19	2	*Cp1*	69,70	68	AQ026089	Ver	v(2)k15819		This study
l(2)k00208	50C17-19			69,70	68	AQ025680	Ver			
l(2)03505	50D1-2	11	*mam*	68,69,70			Ver	l(2)k02214		Treisman and Rubin (1996)
l(2)k04301b	50D5-6			70			Ver			

(*continued*)

TABLE 4
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence Verified?	Reserve	EST	Reference
l(2)k05416	50E1-2								
l(2)s3475	50E4-5				71	AQ026146		GM07864	
l(2)k04907	50E4-7	2	*Hsc70-5*			AQ034144 Ver	l(2)k006618	GM05829	This study
l(2)k12303	50E4-7	3	*SelD*		68	AQ073303 Ver	l(2)k11320		Alsina *et al.* (1998)
l(2)06949	51A1-2	3			71,72	AQ025660 Ver	l(2)k06203		
l(2)02637	51A4-5				71,72	AQ025596		GM01679	
l(2)03563	51A1-2			71		AQ073278 Ver			
l(2)k16805	51A4-5			71,72		AQ026116 Ver			
l(2)00681	51B1-5	10		71,72		AQ026040[a] Ver	l(2)k11904		
l(2)07214	51B7-10	3		72	71	G00586 Ver	l(2)k03307		
l(2)k00103	51C1-2	2	*Pros44.5*		72	AQ034167 Ver	l(2)k14707	HL05493	This study
l(2)k09935	51C1-2				72,73				
l(2)k08015	51D3-5				73	AQ025854			
ms(2)02530	51D1-12		*boc*	73	71	AQ026405 Ver			Castrillon *et al.* (1993)
l(2)k06403	51D7-8			73		AQ025792 Ver			
l(2)02064	51E1-2	2	*chn*	73		G00580 Ver	l(2)k04218		Kania *et al.* (1995)
l(2)01288	51E10-11	2	*scb*	73		G00732 Ver	l(2)k17029		Stark *et al.* (1997)
l(2)k03308	51F11-12			73	74	AQ034142 Ver		LD03765	
l(2)k07207	52A9-11		*VhaD*	73,74		AQ034158 Ver		GM07722	This study
l(2)k05713	52C8-9				73,74,75	AQ025767		LP05195	
l(2)k08407	52D1-2				74,75	AQ025864			
l(2)05248	52D1-2			75	73,74	AQ025633 Ver			
l(2)k02205	52D1-2			74,75	73	AQ034136 Ver			
l(2)k05419	52D9-10			73,74,75		Ver			
l(2)01466	52D9-10			73,74		AQ034162 Ver			
n(2)k09217	52D11-12		*ATPCL*		73,74,75	AQ025890		LD21334	This study
l(2)05070	52E1-2			73,74,76		G00607 Ver			
l(2)15617	52E1-2			76	73,74	AQ026090 Ver		LD33152	
l(2)10403	52E5-6	2		74	73	G01436 Ver	l(2)08639	LD05605	
n(2)k07236	52E5-6		*Rho1*		73	Ver			Strutt *et al.* (1997)
l(2)k09905	52E5-8	4		74	73,76	AQ34022 Ver	l(2)k14705		
l(2)k13209	52F5-7		*Paf-AH&bgr;*	74,75		AQ034030 Ver	n(2)k11702	GM10718	This study
l(2)13314	53A3-5		*Khc*			AQ026054		LD26478	This study
l(2)02836	53B1-2	3		75	74	AQ073274 Ver	l(2)k05440		
l(2)06214	53C1-2		*Hmgs*			AQ025645		LP04424	This study
l(2)k07824	53C1-4	2				AQ034005 Ver	l(2)k16901	LD06385	
l(2)03021	53C6-10	3	*sema-II*			AQ073276 Ver	l(2)k11240	LD09306	Kolodkin *et al.* (1993)
l(2)05428	53C9-10					AQ025637			
l(2)s4639	53C9-10								
l(2)k06503	53C14-15					AQ025797			
l(2)k12701	53D11-13								
l(2)k11009	53D13-14								
l(2)04154	53D13-15					AQ025623			
l(2)k09318	53D13-15								
n(2)k04810	53E1-2		*Ef1β*			AQ025749	n(2)k10209	GM03568	This study
l(2)k07202	53E1-2	2	*veg*			AQ025821 Ver	l(2)k17010		Kania *et al.* (1995)
l(2)K07408	53E4-5					AQ025831			
l(2)k13704	53E4-5					AQ026064			
l(2)k09837	53E6-8					AQ025904		LD18969	
l(2)04291	53F1-2		*RhoGEF2*			AQ025625			Barrett *et al.* (1997)
l(2)k11502	53F1-2					AQ025966			
n(2)06253	53F1-5	6	*Gst2*			AQ025648	n(2)k09303	LD14356	This study
l(2)k15914	53F1-5					AQ026096		LD34058	

(*continued*)

TABLE 4
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence	Verified?	Reserve	EST	Reference
l(2)05091	53F4-5					AQ025631			GH13488	
l(2)k10815	53F6-9					AQ025942				
l(2)06655	54A1-2					AQ034108				
l(2)k09202	54A1-2					AQ025888				
l(2)01038	54B1-2					AQ025584				
l(2)k01212	54B1-2	2				AQ034133	Ver	l(2)k05507		
l(2)k04222b	54B1-2					AQ025740				
n(2)k07110	54B1-2		*mm*			AQ025815				Kania *et al.* (1995)
l(2)k10914	54B1-2									
l(2)06373	54B4-5		*Sip1*			AQ025652			LD04116	This study
l(2)10491	54B4-5					AQ025676				
l(2)k14517	54B4-5									
l(2)k16314	54B10-13		*cnk*			AQ026105	Ver			Therrien *et al.* (1998)
l(2)k08901	54B15-16	3				AQ034114	Ver	l(2)k14501		
l(2)k11012	54B15-16					AQ025945				
l(2)k06904	54B15-18					AQ025809				
fs(2)02086	54C1-12		*rhi*				Ver			FBti0009086
l(2)k07433	54C1-2					AQ025832				
l(2)k11303	54C1-2									
l(2)k10408	54C1-4					AQ025935				
l(2)k07406	54C7-8					AQ025830				
l(2)k15815	54C7-8	4			79	AQ034122	Ver	l(2)k15512	GM09451	
l(2)10505	54C11-12					AQ025677				
l(2)k03303	54E1-2									
l(2)k07509	54E1-2				79	AQ025834				
l(2)k09924	54E1-2				79	AQ025913				
l(2)k11311	54E1-2				79	AQ025959				
l(2)k11505	54F1-2				78,79,80	AQ025967			LD03241	
l(2)06850	54F1-2	3		78,79		G00456	Ver	l(2)s2140		
l(2)k07805b	54F4-5		*thr*	78			Ver			Kania *et al.* (1995)
l(2)k02512	55B5-10	3	*lolal*	78		AQ034138	Ver	l(2)k07907	LD06695	This study
l(2)03091	55B5-6		*Hsf*	78,80		G00582	Ver		LD18486	This study
l(2)k10109	55B5-6		*pAbp*	78,79	81	AQ034024	Ver		GM09987	This study
l(2)s1859	55B5-6	2	*Pcl*	78		AQ026142	Ver	l(2)k08920	GH02674	This study
l(2)04440	55C1-2	10		79,80,81	78	AQ025626	Ver	l(2)k00702		
l(2)07838	55D1-2		*Prp19*	79,80,81	82	G00454	Ver		LD08810	This study
l(2)08770	55D1-2	4		81	79,82	G00698	Ver	l(2)k04808	GM02307	
l(2)k03007	55E1-2	2	*oho55DE*				Ver	l(2)k13104		Török *et al.* 1993
l(2)k06602	55E6-9	2		79,81	82	AQ025800	Ver	l(2)k12907		
1(2)08717	55F2-3			81,82	79,80	AQ025667	Ver			
1(2)03709	55F5-6			81,82	79,80	AQ025615	Ver			
l(2)k08810	56A1-2		*prod*	82	81	AQ025870	Ver		LD09957	Török *et al.* 1997
l(2)02029	56C1-2		*enb*	82	81	AQ025592	Ver		LD19771	This study
l(2)k08713	56C1-2		*cora*	82		AQ034017	Ver		HL02495	This study
l(2)k16207	56D1-2									
sl(2)01103	56D5-6	6	*hts*			G01435	Ver	fs(2)10089	LD10717	Yue and Spradling (1992)
l(2)k16914	56D5-6					AQ026118				
l(2)k00705	56D7-9	11			82	AQ034172	Ver	l(2)k06027		
l(2)k06323	56D8-11					AQ025787			GH14769	
l(2)05338	56E1-2		*sm*			AQ025635	Ver			zur Lage *et al.* (1997)
ms(2)06268	56E1-6		*emm*			AQ026414				Castrillon *et al.* (1993)
l(2)03068	56F1-2		*5SRNA*		83	AQ025605				This study
l(2)k16210	56F1-2									

(*continued*)

TABLE 4
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence	Verified?	Reserve	EST	Reference
l(2)k08002	56F6-9				83	AQ025850				
l(2)k02701	56F6-9	2	*18w*	83		AQ025714	Ver	l(2)00053		Eldon *et al.* (1994)
l(2)k09918	56F8-15		*RpL30*		83	AQ025911			GM06054	This study
l(2)02448	56F10-11	9	*mus209*	83	84	AQ073269	Ver	l(2)k00704		This study
l(2)k10809	56F10-12			83	84	AQ025940	Ver		GM09993	
l(2)s1866	56F11-12			83	84	AQ026143	Ver		GH12502	
l(2)s4831	57A3-4				84	AQ026148				
l(2)05510	57A3-6				83,84	AQ025639			GM10152	
l(2)k09920	57A3-6				83,84	AQ025912				
l(2)k16204	57A5-6				84	AQ026102				
l(2)k07713	57A5-6			84			Ver			
l(2)05056	57A5-6	3		84			Ver	l(2)k07917	LD24469	
l(2)k02206	57A8-9			84	85	AQ025706	Ver		GM10963	
l(2)k06409	57B1-3				84	AQ025794				
l(2)05475	57B2-3	8	*sktl*	85		G00585	Ver	l(2)k02517	LP03320	Hassan *et al.* (1998)
l(2)07806	57B4-6			85		AQ073292	Ver			
l(2)03050	57B13-14			85	84,86	G00547	Ver			
l(2)07206	57B13-14			85			Ver			
l(2)10469	57B13-14	3	*shg*	85	86	G00624	Ver	l(2)k03401		Uemura (1996), Tepass *et al.* (1996)
l(2)k13803	57C1-2		*Xbp1*	85,86		AQ034031	Ver		GM04013	This study
l(2)k08108	57D11-12	3	*domino*	86		AQ034011	Ver	l(2)k00904	GM10082	Braun *et al.* (1997)
l(2)10608	57E1-2	2		85,86		AQ034166	Ver	l(2)k10215		
l(2)00734	57E3-4		*Fkbp13*	85,86		AQ073260	Ver		GM07659	This study
l(2)01467	57E6-7	9		85			Ver	l(2)k00119		Roch *et al.* (1998)
l(2)k10317	57E6-7			85		AQ025934	Ver		GH14568	
l(2)03033	57F1-2	4	*Egfr*	85,86		AQ025603	Ver	l(2)k05115		This study
l(2)k05614	57F5-6			85			Ver			
l(2)k07204	57F5-6			85,86			Ver			
l(2)03605	57F8-10			85		G00574	Ver			
l(2)07837	58A3-4			85,86	87	AQ073294	Ver			
l(2)07768	58D1-2		*ari2*	87		G00612	Ver		LD02916	This study
l(2)01738	58D1-2	7		87	85,86	AQ034163	Ver	l(2)k06515		
l(2)k09801	58D4-5	2		87		AQ025902	Ver	l(2)k15821	LD14720	
l(2)k13211	58D6-7				87	AQ026048				
l(2)k05603	56D8-10		*mei-W68*			AQ025763			LD20471	McKim and Hayashi-Hagihara (1998)
l(2)k08316	58E1-2	2			87	AQ034014	Ver	l(2)k08134		
l(2)k08320	58E3-5									
l(2)k00611	58F4-5				88	AQ025690				
l(2)k17002	58F4-5	2				AQ026121	Ver	l(2)k11534		
1(2)k10825	58F10-12									
l(2)01410	59B1-2	9	*ATPsyn-&agr*	88		AQ073264	Ver	l(2)k00212	GM10728	This study
l(2)10444	59B6-7				88	AQ025675			LD18550	
l(2)01034	59C1-2				88					
l(2)06496	59C1-4	2			88,89,91	AQ025654	Ver	l(2)k14618		
l(2)k09913	59C3-5				88,89	AQ025910			LD07353	
l(2)k07136	59D3-4				88,89	AQ025820				
l(2)k10411	59D3-4				89					
l(2)s4830	59E1-2				88,89,91	AQ026147				
l(2)k06908	59E1-2		*chrw*	89	91	AQ025810	Ver			Kania *et al.* (1995)
l(2)02132	59E3-4	6	*Dcp-1*	89,91		G00735	Ver	l(2)k05606	LD13945	Song *et al.* (1997)
l(2)06369	59F1-2			91	89	AQ073289	Ver			
l(2)03041	59F3-4	3	*apt*	89,91		AQ025604	Ver	l(2)k15608		Eulenberg and Schuh (1997), Gellon *et al.* (1997)
l(2)02535	59F3-4	2		89,91		AQ073271	Ver	l(2)05096		
l(2)k13214	60A3-4			89,91		AQ026049	Ver		GH08760	

(*continued*)

TABLE 4
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence	Verified?	Reserve	EST	Reference
l(2)02970	60A5-9	9	*ken*	89,90,91		AQ073275	Ver	l(2)k11035	GM12839	Castrillon *et al.* (1993)
l(2)k09025	60A8-11			89,90,91			Ver			
l(2)05006	60B1-2			90	89,91,92		Ver			
l(2)k04201	60B1-2		*Adk2*	90	89	AQ025735	Ver		GM04566	This study
l(2)k07623	60B1-2		*Phm*	89,90	91	AQ025840	Ver		GH12243	Kolhekar *et al.* (1997)
l(2)k05633	60B4-5	9	*tsr*	90	89,91,93	AQ025765	Ver	l(2)k09317	GM13866	This study
l(2)01296	60B4-5	2			89-93	G00596	Ver	l(2)10530		
l(2)rQ313	60B8-11				89,90,92	AQ026141				
l(2)09373	60B10-11		*MdaPk*		89,92	AQ034109			LD08416	This study
l(2)k13705	60B11-13		*Zfrp8*	92	89,90,93	AQ026065	Ver		GM13546	This study
l(2)k12101	60B12-13		*spag*	92	89,90,93		Ver			Iliopoulos *et al.* (1997)
fs(2)01310	60C		*slbo*				Ver			Montell *et al.* (1992)
l(2)k05318	60C1-2	2	*Nop60B*	92	89,93	AQ025756	Ver	l(2)k06308	LD17317	This study
l(2)03267	60C6-8	2	*bs*	89,92,93		G00561	Ver	l(2)k07909		Guillemin *et al.* (1996)
l(2)k08003	60D1-2	2	*Mov34*	89,92		AQ025851	Ver	l(2)k16724	HL02986	This study
l(2)04012	60D1-2				89,92,93	AQ025621				
l(2)k10502	60D1-2				89,92	AQ025937				
l(2)k16102	60D5-6		*Nurf-38*	93	89,92	AQ026097	Ver			This study
l(2)k00808	60D6-8				89,93	AQ025692				
l(2)k04809	60D13-14				89,93	AQ025748				
l(2)01092	60E1-2		*Dll*		89,92-95	AQ025585				Goto and Hayashi (1997)
l(2)k03205	60E5-6				89,94,95	AQ025722				
l(2)k03704	60E8-9				89,94,95	AQ025729			LD24657	
l(2)k11118	60E10-12			94	95	AQ025951	Ver			
l(2)02957	60F1-3	2	*zip*		94,95	AQ025602	Ver	l(2)k15609		Perrimon *et al.* (1996)
l(2)10481	60E11-12	2		95	89,94	AQ073298	Ver	l(2)03263		
l(2)01155	60F1-3		*gsb-d*	95	89,94		Ver			Duman-Scheel *et al.* (1997)
l(2)01848	60F1-3			95		G00734	Ver			
l(2)00895	60F1-5	2	*Kr*	95		G00455	Ver	l(2)k05826		Hoshizaki (1994)

Second chromosome stocks from the primary collection are listed in the estimated physical order of their *P*-element insertions along the chromosome. For each strain the table gives the following information: name (Strain), cytogenetic location of *P* insertion (Site), number of alleles characterized among starting strains (Alleles), the disrupted gene (Gene) if known, deficiency chromosomes from Table 2 that fail to complement the strain (Noncomp), deficiency chromosomes from Table 2 that complement the strain (Comp), the accession number of the genomic sequence flanking the insertion (Sequence), whether the *P* insertion has been Verified genetically to cause the associated mutant phenotype—Ver = Yes (Verified?)—the name of an allelic strain that was saved (Reserve), the BDGP identifier of an EST associated with the insertion (EST), a reference in which the strain was used to characterize the disrupted locus (Reference).

[a] Accession numbers that derive from a nonprimary allele.

TABLE 5

Chromosome 3 stocks

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence	Verified?	Reserve	EST	Reference
l(3)L5150	61A1-3				2					
l(3)06240	61B1-2	2			2		Ver	l(3)rH321		
l(3)10512	61C1-2		*trh*		1,2	AQ034076	Ver			Issac and Andrew (1996), Wilk *et al.* (1996)
l(3)neo1	61C1-9									
l(3)05967	61C7-8	3		1,2		AQ034068	Ver	l(3)06318		
l(3)L1170	61C7-8			2		AQ026239	Ver			Demakov *et al.* (1993)
l(3)L3130	61C7-8			2		AQ026252	Ver			
l(3)00835	61D1-2			1,2		AQ034043	Ver		LD03158	
l(3)04322	61D1-2	5	*emc*	1,2		G01462	Ver	l(3)j4E11		Rottgen *et al.* (1998)
l(3)07012	61F3-4	3	*Klp61F*	2,3		AQ026225	Ver	l(3)06345	LD15641	Heck *et al.* (1993)
l(3)L2049	61F3-4	2		2		AQ026247	Ver	l(3)L3879		
l(3)02640	61F6-7	5		2,3		G00701	Ver	l(3)j4E3		
l(3)j4A6	61F6-7		*ND-AcC*	2,3		AQ026317	Ver			This study
l(3)neo5	62A1-12			2		AQ026347	Ver			
ms(3)02509	62A3-4		*cue*			AQ026213			LD11871	Castrillon *et al.* (1993)
l(3)02104	62A4-5				4,5,6					
l(3)neo7	62B1-12		*l(3)62Be*	4,5	6		Ver			Sliter *et al.* (1989)
l(3)04276	62B4-5			4,5	2	AQ026196	Ver			
l(3)j1D7	62C1-2	2		4,5		AQ026295	Ver	l(3)04860	LD07388	
l(3)rL182	62C3-4			4,6	7	AQ026367	Ver			
l(3)L1910	62C1-3			6	4	AQ02646	Ver			
l(3)j1E2	62E5-8		*Nik*	5		AQ026296	Ver		LD34191	This study
l(3)06946	62E6-7	8	*msn*	5	6	G00763	Ver	l(3)06286		Treisman *et al.* (1997a)
l(3)neo8	62F1-6				5,6,7	AQ026353				
ms(3)08445	63A1-B									Castrillon *et al.* (1993)
l(3)06803	63A3-4				7				HL01251	
l(3)j5C2	63B7-8		*Hsp83*	7	20	AQ026323	Ver		GH03850	van der Straten *et al.* (1997)
l(3)L7160	63C1-2			7			Ver			
l(3)01029	63B10-11			7			Ver			
l(3)L3659	63D1-2			7	8	AQ026254	Ver			
l(3)05634	63F5-6		*Ubi-p63E*	8	9	G01167	Ver		GH05622	This study
l(3)L1459	64A4-5				9	AQ026243			LD03011	
l(3)rG166	64A4-5			9		AQ026355	Ver			
l(3)09291	64B5-6			9	10		Ver			
l(3)04556	64C1-2		*Rpd3*	9	10	G00703	Ver		LD06915	This study
l(3)01418	64C9-10		*Srp54*	9,10		AQ034047	Ver		GM09489	This study
l(3)06524	64D3-4			10	9	AQ034069	Ver		CK02636	
fs(3)07084	64E8-12		*p70s6K*	10		AQ025577	Ver		GH02870	U66562
l(3)02331	64E8-12			10		AQ034051	Ver			
l(3)01640	64E8-9			10		AQ034048	Ver		LD20747	
l(3)10567	64F1-3	2	*vn*	10		AQ034078	Ver	l(3)rF264		Yarnitzky *et al.* (1997)
l(3)04026	65A5-6			10		AQ034060	Ver		GM12884	
l(3)02094	65A7-9	2		10		G01461	Ver	1(3)03042	LD03346	
l(3)06811	65A10-11			10		AQ073334	Ver			
l(3)L3999	65A10-11			10		AQ026259	Ver			
l(3)rP047	65A10-11			10		AQ073359	Ver			
l(3)03844	65C1-2		*mdm*	*10*		*AQ034059*				Hong and Hashimoto (1995)
1(3)L4060	65D3-4					AQ026261				

(*continued*)

TABLE 5
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence	Verified?	Reserve	EST	Reference
l(3)08310	65D4-5	4	*sgl*		10	G00549	Ver	l(3)j3E11	LD07210	Hacker *et al.* (1997)
ms(3)04202	65E1-12				11	AQ026430				
l(3)L2249	65E10-11		*Neos*			AQ026249			HL05936	This study
l(3)L7123	65F1-2		*Cdc27*			AQ026278				This study
l(3)j1D5	65F5-8			11		AQ026294	Ver		LD04205	
l(3)07217	66A1-2	3		11			Ver	l(3)02067		
l(3)j6B8	66A1-2	4	*smid*	11		AQ034093	Ver	l(3)s2898	LD17070	This study
l(3)09645	66A17-18			11		AQ073344	Ver			
l(3)j1C7	66B10-11		*Nmt1*	12	11	AQ026293	Ver		GM13220	This study
l(3)L0139	66C1-2				12	AQ026234				
l(3)L4111	66C1-4				12	AQ026263				
l(3)L3852	66C8-9					AQ026256			GH01524	
l(3)03928	66D1-2	5			13	G01407	Ver	l(3)j5A3		
l(3)L4222	66D3-4				13					
l(3)j8E8	66D5-6			13		AQ026334	Ver			
l(3)8247	66D10-12	9	*h*	13		G00473	Ver	l(3)00257		Forjanic *et al.* (1997)
l(3)rK561	66D10-13	2	*SrpR&bgr;*	13		AQ026363	Ver	l(3)rL537	LD12339	This study
l(3)j5B6	66D13-14			13		AQ034090	Ver		LD32443	
l(3)10631	66D14-15		*Prm*	13		AQ034079	Ver		GH14085	This study
l(3)01629	66E1-2			13		G00741	Ver			
l(3)06464	66E1-2		*dally*	13			Ver			Nakato *et al.* (1995)
l(3)04264	66E1-2				13	AQ026215				
l(3)10534	66E2-3	3	*Mrp17*		13	AQ034077	Ver	l(3)j4B2	GM03767	This study
ms(3)07272	66F1-6		*bol*		14	AQ026435				Eberhart *et al.* (1996)
l(3)07238	67B1-2		*eIF-4E*	14		AQ034073	Ver		LD13949	This study
l(3)j2B9	67B4-5			14		AQ026304	Ver			
l(3)j2B8	67B10-11		*Uch-L3*	14		AQ026303	Ver		LD02862	This study
l(3)02240	67C4-5			14		G00700	Ver	l(3)rK145		
l(3)01859	67C5-8			14		G00627	Ver			
l(3)L0539	67E1-2				15	AQ026236				
l(3)rL562	67E1-4		*Dhh1*		14,15	AQ026370			LD05563	This study
l(3)01814	67E5-7	4		15	14,16	AQ034049	Ver	l(3)j4A5		
l(3)rI075	67F1-2	2		16		AQ073356	Ver	l(3)L6731		
l(3)09036	68A1-2	3	*klu*	16	18	AQ034074	Ver	l(3)10052		Yang *et al.* (1997)
l(3)01239	68A4-5			16		AQ026167	Ver	l(3)j9B4	LD06973	
fs(3)neo1	68C1-8		*rt*	16		AQ025579	Ver			Martin-Blanco and Garcia-Bellido (1996)
l(3)05408	68C12-13	2		16, 17	18	AQ073329	Ver	l(3)j1B3		
l(3)03946	68E1-2	9	*CycA*	17, 18		G00572	Ver	l(3)j3C8		Lehner and O'Farrell (1989)
l(3)08232	68F1-3	2		18		G01171	Ver	l(3)j2A6		
l(3)neo18	68F1-8			18		AQ073352	Ver		GH10129	
l(3)j2D3	68F2-3			18		AQ034085	Ver		LD20590	Baumgartner *et al.* (1996)
l(3)05088	69A1-3			18	17	G00525	Ver			
l(3)06924	69E3-4					AQ026223				
l(3)00305	69E3-5	2				G00463	Ver	l(3)01413	LD18893	
l(3)s2783	69F5-6		*RpS12*			AQ026389			LD23808	This study
l(3)01470	70A1-2				19	AQ026170			LD27581	
l(3)04220	70A1-2	2			19	AQ034063	Ver	l(3)j11E7		
l(3)j10B6	70A1-2				19	AQ026283				
l(3)02937	70A1-5	3		19		AQ026183	Ver	l(3)05121	LD07388	
l(3)neo19	70A1-8				19	AQ026338			LD06122	

(*continued*)

TABLE 5
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence	Verified?	Reserve	EST	Reference
l(3)00543	70B1-3				19,20,21	AQ026158				
l(3)L5212	70B7-C1				20	AQ026268			GM14474	
l(3)05871	70C1-2				20,21,22	AQ026210				
l(3)s4868	70C1-2				20	AQ026396				
l(3)s3635	70C4-6					AQ026395				
l(3)00082	70C12-13	2		20,21	22	G00462	Ver	l(3)06704		
l(3)07621	70D1-2			21	22,23	AQ026227	Ver			
l(3)00208	70D1-2	3	*dev*	21,22		G00614	Ver	l(3)j2e11		FBrf0075380
l(3)02402	70D1-2		*l(3)70Da*	21,22		AQ034052	Ver	l(3)s4868		FBti0004865
fs(3)02024	70D4-6	7	*stwl*	21,22,23		AQ025573	Ver	fs(3)03217	LD09806	Clark and McKearin (1996)
l(3)L0499	70D1-2			22	21	AQ026235	Ver			
l(3)00564	70E1-2	2	*PpsM*	22,23	21	AQ026159	Ver	l(3)04680	GM03113	This study
l(3)s2325	70F1-4		*Trl*	23		AQ026386	Ver			This study
l(3)j2A2	71B4-5			23		AQ026300	Ver			
l(3)s2172	71B4-5			23		AQ034107	Ver			
l(3)03576	71D1-2		*CrebA*	23,24		G00569	Ver			Andrew *et al.* (1997)
l(3)s1754	71D1-2			24	23	AQ026378	Ver		GM05443	
l(3)j6B9	71E1-2			24		AQ026329	Ver		LD04071	
l(3)03802	72D1-2	3	*l(3)72Dd*	24	25	AQ034058	Ver	l(3)j3B4		FBti0005488
l(3)j5C8	72D1-2		*th*	24	25	AQ034091	Ver		LP01716	Hay *et al.* (1995)
l(3)s1939	72D8-9		*SsR&bgr;*	24	25	AQ026382	Ver		LD10457	This study
l(3)j5A4	72D10-11		*l(3)72Dn*	24		AQ034087	Ver			FBrf0082216
l(3)s3123	72E1-2			24,25		AQ026393	Ver			
ms(3)03957	72D1-12			25	24	AQ026428	Ver			Castrillon *et al.* (1993)
l(3)05845	73A1-2	6	*gil*	25, 26		AQ073330	Ver	l(3)j5E11	GH06179	Okano *et al.* (1992), Freeman *et al.* (1992)
l(3)j10E8	73A1-2			25,26		AQ034081	Ver			
l(3)02540	73A9-10		*l(3)73Ah*	25, 26		AQ034054	Ver			This study
l(3)neo20	73B1-7			26	27	AQ073353	Ver			Nozaki *et al.* (1996)
l(3)00274	73B1-2		*Mo25*	26,27		AQ026154	Ver		LD09950	Nozaki *et al.* (1996)
l(3)04674	73B1-2		*Abl*	27		AQ026201	Ver		GH09917	This study
l(3)02281	73B5-6	8		27		AQ026179	Ver	l(3)01895	LD17363	
l(3)10547	73D1-2	2	*Int6*	27	26	AQ073349	Ver	l(3)j9E8	LD13907	This study
l(3)04069	73D3-4			26, 27		AQ073320	Ver			
l(3)s1629	73D3-6			27		AQ026377	Ver		LD04071	
l(3)01658	74B1-2	2		27	26	G00742	Ver	l(3)j2B12		
l(3)00073	74C1-2	2	*ttv*	27		AQ026152	Ver	l(3)02619	CK00510	Bellaiche *et al.* (1998)
l(3)02619	74C1-2	2		27		AQ026182	Ver	l(3)j2E3		
l(3)j11B2	74D3-5	2		27		AQ034082	Ver	l(3)j3E5	LD13506	
l(3)neo24	74F1-A1		*Eip74EF*	27		AQ034100	Ver		LP01488	Fletcher and Thummel (1995)
l(3)07041	75B1-2	22	*Eip75B*	28	27, 29	AQ073339	Ver	l(3) 03247		This study
l(3)rL061	75B1-2			28	29	AQ026365	Ver			
l(3)05014	75C1-2		*W (hid)*	28,29,30	29	AQ026203	Ver			Grether *et al.* (1995)
l(3)02069	75C3-4	4		30		G00744	Ver	l(3)j9A6	LD08565	
l(3)neo26	75C1-7		*Cat*		28,30	AQ026340			GM06015	This study
l(3)03649	75D4-5		*ftz-f1*	29	30		Ver			Yu *et al.* (1997)
l(3)00864	75E1-2	2		29		AQ026162	Ver	l(3)rG084		
l(3)j4E6	75E3-5				29	AQ026321				
l(3)neo27	76A1-10					AQ026341			LD10629	
l(3)06945	76A3-4					AQ026224			LD20394	

(*continued*)

TABLE 5
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence	Verified?	Reserve	EST	Reference
l(3)L3809	76B9-10			31		AQ026255	Ver		LD04962	
l(3)j3D4	76D3-4	2	*Pha*	31		AQ026314	Ver	l(3)01058	GM02843	FBrf0082216
l(3)L1243	76D3-4			31		AQ026242	Ver		LD03114	
l(3)01673	77B1-2	2	*polo*	32			Ver	l(3)06089		T. Xie (personal communication)
ms(3)00414	77B1-9	2	*fbl*	32		AQ026422	Ver	ms(3)00713		Castrillon *et al.* (1993)
l(3)00103	77B4-7	4		32	33	G00512	Ver	l(3)05637	LD12337	
l(3)j10B2	77B6-7		*DNAprim*	32		AQ026282	Ver		LD02632	This study
l(3)04521	77D4-E2			32,33		AQ073321	Ver		LD08057	
l(3)rG554	78A1-2				33,34	AQ026356				
l(3)rK760	78A1-2				33,34	AQ026364				
l(3)L7062	78A2-3					AQ026277				
l(3)L5541	78A5-6					AQ026270				
l(3)07615	78C1-2				34	AQ026226				
l(3)00217	78C3-4	2	*l(3)78Cb*	34		G00740	Ver	l(3)06913	HL01257	Russell *et al.* (1996)
l(3)04063	78D1-2			34		AQ073319	Ver			
ms(3)04066	78D1-8		*bet*		34					
l(3)00836	79B1-2				34	AQ026161				Castrillon *et al.* (1993)
l(3)04093	79B1-2	3	*mub*		34	AQ034062	Ver	l(3)j8B3	LD20678	This study
l(3)01544	79D1-2		*RpP0*		34	AQ026171			GM02387	This study
l(3)neo30	79D1-4				34	AQ026342				
l(3)00827	79E1-2		*Aats-ile*		34	AQ026160			GM04407	This study
l(3)03335	79E1-2		*Hem*		34	AQ026185				Baumgartner *et al.* (1995)
l(3)03988	79E1-2		*Csp*		34	AQ026193			GM01170	This study
l(3)05309	79E1-2	9	*Ten-m*		34	AQ073328	Ver	l(3)02017	LD10511	Levine *et al.* (1994) Baumgartner *et al.* (1994)
l(3)09070	79E1-2					AQ026229				
l(3)04053	79E6-7	2				AQ034061	Ver	l(3)j8B2		
l(3)00506	79F1-2					AQ026157				Levine *et al.* (1994),
l(3)04281	79F1-2					AQ026197				
l(3)L7251	79F1-2					AQ026279				
ms(3)03817	80A1-F9				35	AQ026427				
l(3)06713	81F1-6				35	AQ026220				
l(3)00620	82A1-2	3	*abs*	35		G00716	Ver	l(3)06862		Schmucker *et al.* (1997)
l(3)j1E6	82A3-5			35		AQ026298	Ver		LD08389	
l(3)L1233	82B1-2			35		AQ026241	Ver			
l(3)L0021	82C1-2			35		AQ034080	Ver		CK01877	
ms(3)07735	82C1-5		*shank (shk)*		35					Castrillon *et al.* (1993)
l(3)10112	82D1-2			35		AQ073347	Ver			
l(3)j3A4	82D1-2	2	*Kary&bgr;3*	35		AQ026311	Ver	l(3)j7E8	LD12881	This study
l(3)j4D1	82D1-2	2		35		AQ026319	Ver	l(3)j5D1		
ms(3)06208	82D1-8		*bob*	35		AQ026432	Ver			Castrillon *et al.* (1993)
l(3)01456	82D4-5	2		35		AQ073310	Ver	l(3)02466	LD12651	
l(3)02733	82D4-7	2		35		G00694	Ver	l(3)10619		
l(3)01010	82E1-2					AQ026163				
l(3)rK509	82E4-5			35		AQ026362	Ver		LD20584	
l(3)j4B9	82E4-5				35					
l(3)07128	82E5-7	4			35	AQ073341	Ver	l(3)05259	LD16501	
l(3)neo32	82F1-8	2	*corto*			AQ034101	Ver	l(3)neo31		This study
l(3)L3051	82F4-5				35	AQ026250				
l(3)02255	82F4-7				35	AQ026177				
l(3)09904	82F8-9				35					

(*continued*)

TABLE 5
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence	Verified?	Reserve	EST	Reference
l(3)01319	83A5-6		*Snr1*		35	AQ026168	Ver		LD18257	Dingwall *et al.* (1995)
l(3)03644	83A5-6				35	AQ026187				
l(3)05616	83A5-6	3	*Itp-r83A*		35	AQ034066	Ver	l(3)j5B4		Venkatesh and Hasan (1997)
l(3)j5E2	83A5-6		*KSR*			AQ026325	Ver			Therrien *et al.* (1995)
l(3)03022	83B1-2					AQ026184			LD10048	
l(3)j3E7	83B1-2		*noi*			AQ026316			LD16525	This study
l(3)j5E7	83B1-2					AQ026326			LD04714	
l(3)03834	83B4-5		*Rga*			AQ026190			LD19865	This study
l(3)j9B6	83B4-5					AQ026337			LD14028	
l(3)s1938	83B4-7		*Atu*			AQ026381			LD02958	This study
l(3)02248	83B6-7		*Xe7*			AQ026176			GM07812	This study
l(3)j13C8	83B6-7					AQ026287			LD05302	
l(3)04696	83C1-2	2				G00704	Ver	l(3)04712	LD14028	
l(3)j1C2	83C1-2	2	*cas*			AQ026292	Ver	l(3)neo33		Cui and Doe (1992), Mellerick *et al.* (1992)
l(3)01086	83D4-5	9	*Lip*			G01404	Ver	l(3)j3D2	LD07117	This study
l(3)03342	83F1-2	3		36	37	G01163	Ver	l(3)sA3484		
l(3)01241	84A1-2		*lab*	36,37	38	G01161	Ver			This study
l(3)04498	84A4-5		*pb*	37	38	G01164	Ver			This study
l(3)05614	84A4-5		*twr*	37,38	38	G01166	Ver		LD15759	FBrf0082216
l(3)j7A6	84B1-2			38		AQ034097	Ver			
l(3)L2100	84B2-3			38		AQ026248	Ver		LD18350	
l(3)02267	84C1-2		*Aly*	38	66	AQ026178	Ver		LD10551	This study
l(3)j8C8	84C1-2	3		38		AQ034099	Ver	l(3)13A6		
l(3)s2214	84C4-6			38		AQ026383	Ver			
l(3)neo34	84D1-14		*LAP*	38		AQ034102	Ver		GH06360	Zhang *et al.* (1998)
l(3)02732	84D11-12			38	39		Ver			
l(3)neo35	84D1-13				38,39	AQ026343	Ver			
ms(3)02245	84F1-16	2	*sdl*		39		Ver	ms(3)05090	LP04056	Castrillon *et al.* (1993)
l(3)03692	84E10-11		*tom34*	39	38	AQ073315	Ver		HL06590	This study
l(3)j4E1	84E10-11			39		AQ026320	Ver			
l(3)05930	84F1-2		*Gata-c*	39		AQ034067	Ver			Lin *et al.* (1995)
l(3)rL074	84F6-7		*Mcm2*	39		AQ026366	Ver		LD05520	Treisman *et al.* (1995a)
ms(3)00940	85A4-5		*cap*	39		AQ026423	Ver			Castrillon *et al.* (1993)
l(3)L4740	85A5-6			39		AQ026266	Ver			
l(3)s3512	85A5-7		*dhod*	39		AQ026394	Ver		LD09208	This study
l(3)j1B9	85A9-10		*tRNA:Y1:85Aa*		39	AQ026291			GH13240	This study
l(3)00281	85B8-9				39	AQ026155				
l(3)neo36	85C1-13		*Rel*		39	AQ026344			GH1881	This study
l(3)L6332	85C1-2									
l(3)05652	85C3-5				40	AQ026209				
l(3)j6B12	85C9-10	6	*neur*		40	AQ034092	Ver	l(3)j9B8		Boulianne *et al.* (1991)
ms(3)03565	85D1-27				40	AQ026426				
l(3)01688	85D2-3	10	*pum*		40,41	AQ025572	Ver	l(3)01688		Lin and Spradling (1997)
l(3)04837	85D5-6				40					
l(3)01728	85D7-8				40,41					
l(3)03559	85D7-8	2	*Aats-trp*	40	41	G00589	Ver	l(3)04410	LD24552	FBti0005478
l(3)L4092	85D11-12	2		40		AQ026262	Ver	l(3)L4091	LD14230	
l(3)10615	85D15-17					AQ026233				
l(3)06677	85D18-20	3	*Ras85D*		40,41	AQ034071	Ver	l(3)s1747	CK01231	Schnorr and Berg (1996)

(*continued*)

TABLE 5
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence	Verified?	Reserve	EST	Reference
n(3)s2681	85F7-8				40,41	AQ026388				
l(3)02414	85F12-13	6	*tws*	41	40	G01405	Ver	l(3)j11C8	LD14078	Uemura *et al.* (1993)
l(3)j9A5	85F15-16				40,41	AQ026335				
l(3)06142	86B1-2				40,41	AQ026212				
l(3)06439	86B1-2	5			40,41	AQ073331	Ver	l(3)03676		
l(3)j8B6	86B1-2				42	AQ026332				
l(3)05745	86C1-2	2	*hth*		40,42	G00762	Ver	l(3)06762		Rieckhof *et al.* (1997)
l(3)j3C1	86C3-4		*TfIIFbeta*		42	AQ026313			LD10269	This study
l(3)09656	86E1-2			42	40	G01172	Ver			
l(3)10419	86E1-2	19	*pros*	42		AQ073348	Ver	l(3)j6E2		Kauffmann *et al.* (1996)
ms(3)04112	86E1-20		*tho*	42	40	AQ073360	Ver			Castrillon *et al.* (1993)
l(3)04629	86E16-19	2		42		AQ026200	Ver	l(3)05275		
l(3)j1D8	86E16-19			42		AQ034084	Ver			
l(3)neo38	86F1-11			42		AQ034103	Ver		GH13205	
l(3)10621	86F6-7			42		AQ073350	Ver			
l(3)07842	87B4-6	6	*svp*	42, 43	44	G00472	Ver	l(3)j2E2		Begemann *et al.* (1995)
l(3)j6E7	87B10-13		*Pp1-87B*	43	44	AQ034095	Ver		LD10068	This study
l(3)j2E9	87C2-3		*Vha55*	43		AQ026310	Ver		GM02970	Davies *et al.* (1996)
l(3)05043	87C6-8			43,44	42	AQ073323	Ver			Davies *et al.* (1996)
l(3)03463	87D7-9		*CtBP*	44	43	AQ073314	Ver			Nibu *et al.* (1998)
l(3)neo39	87E1-12			44	43,45		Ver			
l(3)j5A1	87E5-6					AQ026322				
l(3)s2149	87E10-11	2	*l(3)87Eg*		44	AQ034106	Ver	l(3)s3582		FBal0028397
l(3)05137	87E7-8				44	AQ026205				
l(3)j6E3	87F2-3	2	*sqd*			AQ026331	Ver	fs(3)00048	LD16014	Kelley (1993)
l(3)j4B4	87F3-4	2			45	AQ026318	Ver	l(3)j8E7		
l(3)L4179	87F7-8				45	AQ026264			LD13350	
l(3)s2249	87F7-8		*B52*	45		AQ026385	Ver		LD03622	This study
n(3)01949	87F10-11				44,45	AQ026174				
l(3)03477	88A4-5	7		45		AQ034055	Ver	l(3)j1D1		
l(3)j1E7	88A4-5			45		AQ026299	Ver			
l(3)L5340	88B1-2			45		AQ026269	Ver			
l(3)00347	88B1-3	12	*trx*	45		G00467	Ver	l(3)j14A6	GH06495	This study
l(3)neo41	88C1-10			45	46	AQ026345	Ver		LD12042	
l(3)06951	88C1-4			45,46		AQ073336	Ver			
l(3)10460	88C9-10	2	*put*		45,46	AQ034075	Ver	l(3)j5A5	LD09722	Ruberte *et al.* (1995), Letsou *et al.* (1995)
l(3)L1231	88C9-11				46	AQ026240				
l(3)02404	88D1-2	4			46	AQ034053	Ver	l(3)j5C3		
l(3)03719	88D1-2	2			46	AQ073317	Ver	l(3)j3B3	GH14307	
l(3)01462	88D5-6	15	*eff*		46	G00767	Ver	l(3)j2C8		Berg and Spradling (1991)
l(3)10418	88D5-6					AQ026231				
l(3)04713	88E1-2									
l(3)03550	88E8-9	5	*Hsc70-4*		47	AQ034056	Ver	l(3)j7A4	GM02246	This study
l(3)neo43	88E9-10					AQ026346			LD18041	
l(3)j6A3	88E11-12					AQ026327			LD22492	
l(3)02299	88F1-2	9	*Tm1*		47	Dm2383	Ver	l(3)s2958	LD05487	Tetzlaff *et al.* (1996)
l(3)j6A6	88F7-8				47	AQ026328				
l(3)04210	89A1-2		*Pros26S*		47	AQ026194			GM06024	This study
l(3)05057	89A1-2				47	AQ026204				

(*continued*)

TABLE 5
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence Verified?	Reserve	EST	Reference
l(3)rN346	89A1-2				47	AQ026372			
l(3)06490	88F7-8			47		AQ026218 Ver		GH13857	
l(3)06536	89A4-5	2		47		AQ034070 Ver	l(3)j2E5	LD18382	
l(3)01618	89A8-9			47	48	AQ026173 Ver			Perrimon *et al.* (1996)
fs(3)03987	89B1-2		*spn-E*			AQ025574 Ver			Gillespie and Berg (1995)
l(3)01549	89B1-3	2	*srp*	47	48	AQ026172 Ver	l(3)neo45	GH11649	Rehorn *et al.* (1996)
ms(3)08724	89B1-3			47	48	AQ026437 Ver			
l(3)04226	89B6-7		*Akt1*	47	48	AQ026195 Ver		GM04486	This study
ms(3)04895	89B1-22		*Hrr25*	48		AQ073361 Ver		LD08007	This study
l(3)L1820	89B11-13			48		AQ026245 Ver			
l(3)05203	89B12-13	6		47,48		AQ026207 Ver	l(3)07636		
l(3)03881	89B12-15	6		47,48		AQ073318 Ver	l(3)j1C5		
l(3)L4032	89D1-2		*JadBp*	48		AQ026260 Ver		LD14392	This study
fs(3)05649	89E1-8		*AbdB*			AQ025575 Ver			Lin and Spradling (1993)
l(3)L4569	89E7-9		*EfTuM*		49	AQ026265		GM04894	This study
l(3)j1E4	89E10-11		*Dad*		49	AQ026297			Tsuneizumi *et al.* (1997)
l(3)06442	89E10-11			49		AQ073332 Ver			
l(3)07882	90B3-4					AQ026228		LD15947	
l(3)00090	90C5-8	7	*eld*		50	G00550 Ver	l(3)j9C3		Treisman *et al.* (1997b)
l(3)01432	90D1-2	37	*cpo*	50		AQ073308 Ver	l(3)j4A1		This study
l(3)00643	90E1-2			50		G00468 Ver			
l(3)03999	90E1-2	3	*sr*	50		G00573 Ver			Frommer *et al.* (1996)
l(3)05697	90E1-2			50		G01168 Ver			
l(3)neo48	90E1-7			50		AQ034104 Ver		LD02313	
l(3)03702	90F1-2	3	*repo*	50		AQ073316 Ver	l(3)00692	GH11090	Xiong *et al.* (1994), Campbell *et al.* (1994)
l(3)s2956	90F1-2			50		AQ026391 Ver		CK00124	
l(3)j2B10	90F6-7		*14-3-3epsilon*	50		AQ026302 Ver			Chang and Rubin (1997)
l(3)05284	91A4-6	3	*sprd*	50,51		AQ073327 Ver	l(3)j3C3		This study
ms(3)06411	91B1-8	2	*fru*			AQ026434 Ver	ms(3)08366		Ito *et al.* (1996)
l(3)07551	91B5-6	2	*fray*	51		G00764 Ver	l(3)s2427	GH10417	Russell *et al.* (1998)
l(3)08126	91B5-6	2		51		G01170 Ver	l(3)j2B3		
l(3)02515	91D3-5				51	AQ026180			
l(3)07013	91F1-5			52	51	G00601 Ver		GM10731	
l(3)07117	91F4-5	5	*nos*	52	51	G00548 Ver	l(3)j3B6		This study
l(3)03346	91F6-9			52		AQ026186 Ver		LD13685	
l(3)03675	91F7-9	2		52		G00570 Ver	l(3)03750		
l(3)02102	91F10-11			52	51	AQ034050 Ver			
l(3)j5A6	91F10-11	4		52		AQ034088 Ver	l(3)j7C8		
l(3)05151	92A1-2	6	*Dl*	52		AQ073325 Ver	l(3)j8C3		Rottgen *et al.* (1998)
l(3)05820	92A1-2			52		G00746 Ver			
l(3)05113	92A13-14		*VhaG*	52		Dm0379 Ver		GH07606	This study
l(3)06916	92B2-3	3		52	53	G00595 Ver	I(3)00857	GH08887	
l(3)01344	92B3-4				52	AQ026169			
l(3)10585	92B3-5			53	52	AQ026232 Ver			
fs(3)08482	92D1-2		*bwk*			AQ025578 Ver			Rittenhouse and Berg (1995)
l(3)06346	92E2-4	2	*Stat92E*	53		AQ026216 Ver	l(3)j6C8		Hou *et al.* (1996)

(*continued*)

TABLE 5
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence Verified?	Reserve	EST	Reference
l(3)03806	92F1-2			53		AQ026189 Ver		LD18575	
l(3)j5C7	93A4-5				57	AQ026324			
l(3)neo54	93B1-13			54		AQ026349 Ver			
l(3)01453	93B1-2	20	*Atpalpha*	54	55	G01160* Ver	l(3)j6A4		Feng *et al.* (1997)
l(3)07086	93B8-11			54	55	AQ073340 Ver		LD12473	
l(3)00295	93B10-11	3	*slmb*	54		Ver	l(3)04295		Jiang and Struhl (1998)
l(3)02231	93B10-11	2		54		G00745 Ver	l(3)j6B6		
l(3)03773	93C1-2	2		54,55		G00717 Ver	l(3)02312		
l(3)j2D1	93C1-2		?	54		AQ026306 Ver			
n(3)05241	93D4-7	2		54	56	G00592 Ver			
l(3)03852	93E1-2	6	*mod(mdg4)*	55,56	54	G00571 Ver	l(3)j2B7	LD06927	Dorn *et al.* (1993)
l(3)05545	93E4-5		*InR*	55,56		G01165 Ver			Fernandez *et al.* (1995)
l(3)07172	93E8-9	4	*E2f*	55,56		Ver	l(3)j3B1	GM02934	Duronio *et al.* (1995)
fs(3)00617	93F6-14		*tsl*			AQ025571 Ver			Savant-Bhonsale and Montell (1993)
l(3)j5B5	94A1-2	2	*how*	55	57	AQ034089 Ver	l(3)j5D5	LD20232	This study
l(3)05712	94A3-4	2		55	56,57	G01169 Ver	l(3)07484		
l(3)L3560	94A5-7				57	AQ026253			
l(3)03685	94A8-12				55,56,57	AQ026188		LD20323	
l(3)rQ178	94A9-10				57	AQ026375			
l(3)L4910	94B4-5		*Dph5*		57	AQ026267		LD12153	This study
l(3)L0580	94C1-2				57	AQ026237		LD11053	
l(3)rN712	94D1-4		*klg*			AQ026373		LD10776	This study
fs(3)00107	94E1-10		*orb*		57	Ver			Lantz *et al.* (1992)
l(3)rJ413	94E1-4	2	*hh*	57	55	AQ026358 Ver	l(3)neo56		Lee *et al.* (1992), Heberlein *et al.* (1993)
l(3)03921	94E3-7	3	*cnc*		57	G00591 Ver	I(3)05134	LD04714	Mohler *et al.* (1995)
l(3)01031	94F1-2		*sec13*			AQ026165		LD22416	This study
l(3)07825	94F1-2	12	*pnt*		57	G00471 Ver	l(3)j1B7		O'Neill *et al.* (1994)
l(3)rF149	94F1-2					AQ026354			
l(3)06906	95A7-8				57	AQ026222			
l(3)04684	95B1-3					AQ026202		LD06189	
l(3)01152	95B5-6		*HmG-CoAR*			AQ026166		LD15354	This study
l(3)rQ303	95D1-2								
l(3)00096	95D10-11				58	AQ026153		LD25146	
l(3)06737	95E1-2	5	*syx1A*	58		G01403 Ver	l(3)j4D9	HL03543	Schulze *et al.* (1995)
ms(3)06746	95F1-15	2	*jar*	58,59		Ver	ms(3)06840		Castrillon *et al.* (1993)
l(3)05842	95F11-12	3		58,59		G00594 Ver	l(3)s5349		
l(3)07207	95F11-12	2		58,59	58	AQ073342 Ver	l(3)07883		
l(3)j1B5	95F11-12			59		AQ026290 Ver			
l(3)L6710	95F14-A1				59	AQ026273		LD02665	
l(3)rI809	96A7-8				59,60				
l(3)L6540	96B3-5				60	AQ026272			
l(3)01207	96B10-11			60	59	AQ034045 Ver		LD17262	
l(3)j2D9	96B19-20		*OstStt3*		60	AQ026308		GM01838	This study
l(3)rL205	96C1-2		*Fur1*		60	AQ026369		LD05213	This study
fs(3)neo61	96C1-9		*bam*		60	AQ025580 Ver			McKearin and Spradling (1990)
l(3)01969	96C7-8				60	AQ026175		LP02719	
l(3)05461	96C8-9	2	*Aats-gln*		60	AQ034065 Ver	l(3)1926	GM13383	This study

(*continued*)

TABLE 5
(Continued)

Strain	Site	Alleles	Gene	Non-comp	Comp	Sequence	Verified?	Reserve	EST	Reference
l(3)j12B4	96C8-9					AQ026285				
l(3)rJ880	96D1-2				60	AQ026359				
l(3)rQ197	96F1-2			61	60	AQ026376	Ver			
l(3)j7B3	97B8-9			61,62		AQ034098	Ver			
ms(3)03445	97D1-15		*Rb97D*	62		AQ026425	Ver			Karsch-Mizrachi and Haynes (1993)
l(3)rK344	97D1-2					AQ026361				
l(3)05146	97D3-6	2	*H2AvD*	62		AQ026206	Ver	l(3)L1602	LD15832	This study
n(3)03884	97D6-9				62	AQ026192				
l(3)rL203	98B1-2					AQ026368			LD13603	Murphy *et al.* (1995)
ms(3)06302	98B1-8					AQ026433				
l(3)06487	98C3-4					AQ026217				
l(3)s2976	98E1-2				63	AQ026392				
l(3)s2784	98F1-2			63		AQ026390	Ver			
l(3)01705	98F4-5	2	*Doa*	63			Ver	l(3)04743		FBti0005439
l(3)04708	99A1-2			63	64	AQ034064	Ver			
l(3)06743	99A4-5			64	64	G00628	Ver			
l(3)01235	99A5-6	10	*stg*	64	63	G00587	Ver	l(3)j1D3		Edgar and O'Farrell (1989)
l(3)L6241	99A5-6				64	AQ026271	Ver			
l(3)05218	99B7-10				64					
l(3)06734	99B8-10				64					
l(3)05884	99C1-2		*ncd*			AQ026211			LD12267	This study
l(3)s2222	99D1-2					AQ026384				
l(3)j11B7	99E1-2					AQ026284				
l(3)00035	99F1-2	14	*Fer2LCH*			AQ073305	Ver	l(3)j2A3	LD10239	B. Dunkov (personal communication)
l(3)00451	99F1-2	4	*Fer1HCH*			AQ026156	Ver	l(3)j10B4	LD16801	This study
l(3)j2D5	99F1-2					AQ026307				
l(3)j8B9	99F8-9					AQ026333				
l(3)00848	99F10-11	3	*spdo*			AQ034044	Ver	l(3)02288	LD07593	This study
l(3)00865	100A1-2		*zfh1*		65	AQ073307			LD15891	Justice *et al.* (1995)
l(3)s2500	100A5-6				65	AQ026387				
l(3)07028	100B1-2				65	AQ073337				
l(3)03670	100B1-2			65		G00590	Ver		LD05921	
l(3)j3B9	100B2-4	2	*dbt*	65		AQ026312	Ver	l(3)rK215	LD04938	Kloss *et al.* (1998)
l(3)00720	100B5-7			65		AQ073306	Ver			
fs(3)06936	100C		*Gprk2*			AQ025576	Ver		LD33670	Schneider and Spradling (1997)
l(3)rM731	100C1-2			65	66	AQ073358	Ver			
l(3)02667	100D1-2	11	*ttk*	66,67	65	AQ073312	Ver	l(3)j2A1		Xiont and Montell (1993)
l(3)j2A4	100E1-2	2	*awd*	67		AQ026301	Ver		GM07644	This study
l(3)s1921	100E1-2			67		AQ026380	Ver		LD09536	
ms(3)10515	100E1-3	2	*heph*	67		AQ026438	Ver	ms(3)07446		Castrillon *et al.* (1993)
ms(3)07570	100E1-F5		*mod*	67		AQ026436	Ver		GM04021	This study
l(3)03429	100F1-5	3		67		G00761	Ver	l(3)j11B9		
l(3)L1022	100F1-2				67	AQ026238			LD11808	
l(3)L7321	100F1-2				67	AQ026280			HL05832	
v(3)03847	100F4-5				67	AQ026191				

Third chromosome stocks from the primary collection are listed in the estimated physical order of their *P*-element insertions along the chromosome. For each strain the table gives the following information: name (Strain), cytogenetic location of *P* insertion (Site), number of alleles characterized among starting strains (Alleles), the disrupted gene (Gene), if known, deficiency chromosomes from Table 2 that fail to complement the strain (Noncomp), deficiency chromosomes from Table 2 that complement the strain (Comp), the accession number of the genomic sequence flanking the insertion (Sequence), whether the *P* insertion has been Verified genetically to cause the associated mutant phenotype—Ver = yes (Verified?)—the name of an allelic strain that was saved (Reserve), the BDGP identifier of an EST assocaited with the insertion (EST), and a reference in which the strain was used to characterize the disrupted locus (Reference). Accession numbers marked with an asterisk derive from a nonprimary allele.

mate the saturation behavior of *P* elements have utilized inadequately characterized data sets. We focused on the 737 independent lines from chromosome *2* and 535 independent lines from chromosome *3* that contain a single verified *P* insertion lying clearly within the validated deficiencies used in the verification analysis. Within this group, for a known number of total lines (transposition events), the number of genes mutated and how many times each was hit should have been determined with complete accuracy. Because the deficiencies included a majority of chromosome *2* and *3* genes (60.3 and 62.0%, respectively), and should be distributed effectively at random, this sample should accurately represent all insertions that cause a phenotype.

Focusing on the less-mutagenized chromosome first, we determined that 154 of the 535 third chromosome genes had been hit once, 43 twice, 16 three times, 6 four times, and 5 five times and that 18 were previously discussed hotspot loci hit six or more times (Table 1). Despite the small number of hotspot loci, they accounted for 204 of the 535 insertions (38%). First, we attempted to fit the data to a Poisson distribution, ignoring for the moment the obvious presence of hotspot genes. The best distribution ($\lambda = 0.558$; Table 1) fit the data poorly because it predicted only 8.0 genes (instead of 16) hit three times, only 1.1 (instead of 6) hit four times, and 0.1 instead of 5 hit five times ($\chi^2 = 270$, $P \ll 0.001$). To determine if the observed "excess" of genes hit three to five times was caused by the statistical tail from the hotspot loci, we used a binomial distribution to model their contribution (Table 1). The distribution used maximizes the contribution of hotspot genes to the classes of genes hit 3–5 times, while yielding the observed number of genes hit 6–12 times. Despite this, the results reveal that there are too few hotspot genes to account for the excess of genes hit 3–5 times (Table 1; $\chi^2 = 20$, $P \ll 0.001$).

Consequently, a class of genes of intermediate mutability must exist (warmspot genes). To estimate the size of this class, we fit the data for genes hit one to five times on the assumption of two mutability classes, warmspot and coldspot genes. Postulating 115 "warmspot" genes ($\lambda = 1.51$) and 613 "coldspot" genes ($\lambda = 0.241$) produced a good fit to the data (Table 1; $\chi^2 = 0.81$, $P \gg 0.05$). Extrapolating the warmspot and coldspot data to the entire chromosome and adding the whole chromosome hotspot data, the following was predicted for the third chromosome: 27 hotspot loci + 115/0.603 = 191 warmspot loci + 613/0.603 = 1017 coldspot loci.

We next considered the second chromosome and found that 737 independent verified lines defined 190 genes that had been hit once, 57 twice, 19 three times, 17 four times, 5 five times, and 32 that were hotspot loci hit six or more times (Table 1). Hotspot insertions accounted for 288 of these lines (39%). Again, at least two general classes of mutability were required to fit the data from non-hotspot lines, even after correcting for the contribution of hotspot genes (Table 1; $\chi^2 = 36$, $P \ll 0.001$). Because we expected genes on the second and third chromosomes to have the same average mutabilities, we reasoned that the Poisson parameters for chromosome *2* warmspot and coldspot loci should correspond to parameters of the corresponding chromosome *3* genes corrected for the more extensive mutagenesis that was carried out on chromosome *2*. The relative fraction of independent single-insert lines analyzed on chromosome *2* compared to chromosome *3* was 737/535 = 1.38. Multiplying the warmspot and coldspot class Poisson parameters determined for chromosome *3* by 1.38 gave the expected values on chromosome *2* ($\lambda = 2.08$ and 0.331). Values of 110 warmspot and 680 coldspot loci on chromosome *2* were then determined to fit the distribution (Table 1; $\chi^2 = 3.8$, $P \gg 0.05$). Thus, chromosome *2* is predicted to house 47 hotspot genes, 110/0.62 = 177 warmspot genes and 680/0.62 = 1097 coldspot genes.

RESULTS

Rationale: The gene disruption library was assembled from ~3900 starting lines that had been produced in seven separate single *P*-element mutagenesis screens (Table 2). Each starting line contained one (or a few) *P*-element insertion on an autosome bearing a newly induced scorable recessive phenotype. The process of going from this amalgamated raw collection to the finished library involved (1) localizing the insertions by *in situ* hybridization to polytene chromosomes at high resolution; (2) identifying strains with allelic insertions by *inter se* complementation crosses; (3) verifying that insertions were responsible for the mutant phenotype by crossing to chromosomes bearing deficiencies; and (4) sequencing DNA flanking the insertions and comparing it to EST and genomic sequence databases (Spradling *et al.* 1995). Single-insert-bearing strains that appear to disrupt distinct genes based on all these criteria constitute the final library, or "primary collection." Because of the requirement for complementation testing, the project was designed initially to focus on genes that mutate to a recognizable lethal, sterile, or visible phenotype.

Identifying single-insert lines: The *P* insertion(s) in each line was cytogenetically localized by *in situ* hybridization as described previously (Spradling *et al.* 1995; see also materials and methods). The number of lines localized from each screen is given in Table 2. Highly consistent and accurate localizations were required for the success of the complementation analysis. Images of each localization were digitized and stored (http://www.fruitfly.org/p_disrupt/). Based on these results, a significant number of strains were immediately eliminated from consideration for the primary collection. Two hundred seven lines were discarded because they did not represent independent insertions (see materials and methods) and >900 lines were eliminated because they contained two or more insertions or a rearrangement on the mutation-bearing chromosome. A total of 2695 independently derived strains bearing single *P*-element insertions (1643 on II, 1052 on III) on intact chromosomes were retained. The *P*-element insertion site within each primary collection strain is listed in Tables 4 and 5.

Identifying allelic mutations: Complementation crosses were carried out *inter se* between lines whose insertions were located near each other (see materials and methods). The maximum cytogenetic distance between the reported positions of insertions that is required to ensure they are not allelic depends critically

on the accuracy of the *in situ* localization. We complementation tested lines when the distance between their elements was six to eight bands or less. This should have been sufficient to eliminate errors even in cytogenetically difficult regions, because the divergence in the reported positions of allelic insertions averaged less than one band (Spradling *et al.* 1995). The molecular analyses reported below provide further independent verification that allelic lines were not missed due to localization errors.

The complementation analysis provided considerable insight into the frequency with which individual loci are mutated by *P* elements (see Tables 4 and 5). In particular, 74 complementation groups on the autosomes were identified that are hotspots for *P*-element insertion with between 6 and 37 alleles each. Because of the size of our data set, these loci likely comprisevirtually all the *P*-element insertion hotspots on the autosomes. Following completion of the complementation analysis, one allele of each complementation group was retained for the primary collection (see Table 2).

Verifying that the insertions cause mutations: Three criteria were used to determine if the *P*-element insertion in a given single insert line was likely to be responsible for the observed mutant phenotype. First, if the mutation failed to complement one or more independently derived strains whose insertions had been localized nearby, then it was considered verified along with all the other insertions in the complementation group so defined. (The chance that such lines actually contained identical secondary or background mutations was negligible as indicated by test crosses with lines whose insertions were at different sites.) To apply the second verification test, the strain in question was crossed to deficiency chromosomes whose cytogenetically determined breakpoints (Table 3) indicated that they might lack the disrupted gene. Crosses were scored based on the presumed phenotype of the insertion (Tables 4 and 5, "Df comp" and "Df noncomp"). Lines that failed to complement were considered to be verified, because the chance that a background mutation was closely linked to the *P* element was acceptably small. If complementation was observed, the line was discarded if its insertion clearly fell within the deficiency boundaries; otherwise it was retained but remained unverified. These two tests, combined with further verification based on the analysis of flanking DNA sequences as described below, allowed the total number of lines in the primary collection to be reduced to 1045, of which 725 (69%) are verified (see Table 2). Of these lines, 93% disrupt vital genes, while most of the remainder cause male or female sterility. The phenotype and verification status of each line are shown in Tables 4 and 5.

We can estimate the approximate number of bogus lines that remain in the library. First, the overall fraction of verified lines arising from each screen is calculated by restricting our analysis to those lines whose insertions clearly fall within the boundaries of valid deficiencies and hence can be reliably tested (Table 2, "in Df"). This subgroup represents >60% of all the lines and should be representative of each screen as a whole. The proportion of lines that were verified ranges from 48–88% among the seven screens. Assuming that insertions falling outside the deficiencies are as likely to be valid as those inside allowed us to estimate the number of unverified primary collection lines from each starting screen that are likely to be valid. After making this final correction, the final number of different genes disrupted by *P* insertions in the collection is estimated to be 953 (Table 2). Using this information we also determined an overall efficiency for each of the seven screens, defined as the percentage of raw lines that contain a single insertion causing its associated phenotype (Table 2, "screen efficiency").

Deficiency chromosomes with accurate breakpoints are a valuable genetic resource. Knowledge of the true extent of material deleted in deficiency stocks was improved as a by-product of verifying the *P* insertions. The location of each verified insertion predicted the expected complementation behavior with relevant deficiencies. In cases where contradictions were observed, the breakpoints of the deficiency could sometimes be refined on the basis of the cytogenetic localizations of the terminal *P* elements (see Tables 4 and 5). A number of such corrections have been incorporated into FlyBase (see flybase.bio.indiana.edu:80/.bin/fbabsq.html). Table 2 shows current estimates of the deficiency breakpoints used in these studies.

Characterizing insertions using flanking DNA sequence: The genomic DNA sequence flanking the insertion sites in the primary collection lines was needed to complete the verification process and to begin associating lines with specific genes. Physically associating as many insertions as possible with specific sites in the genome would also enhance the usefulness of the primary collection for gene mapping and for directed mutational screening using accurately positioned starting strains (Spradling *et al.* 1995). Consequently, we attempted to recover genomic DNA adjacent to the 5′, 3′, or both sides of the *P* element from every remaining candidate primary collection line following completion of the genetic verification tests. Both plasmid rescue and inverse PCR were used. A single sequencing run was carried out beginning at the insertion site of all recovered flanks (see materials and methods). Despite a shorter average amount of sequence recovered, inverse PCR was successful at a slightly higher average frequency (85% *vs.* 80%) and could be carried out in a 96-well format that allowed lines to be analyzed more

rapidly. If both a 5′ and 3′ sequence was obtained, the two runs were merged in a single contig.

The sequences flanking the insertions were initially compared among themselves as an additional verification test. We wished to eliminate lines whose insertions were very close together but that behaved genetically like separate genes. Such lines are likely to be produced when chromosomes bearing nonallelic background mutations acquire insertions within the same nonvital gene. The genetic behavior of the resulting strains will cause them to survive into the primary collection if their insertions lie outside existing deficiencies. On the other hand, we did not want to eliminate valid insertions in adjacent genes. Consequently, in the absence of additional information, nonallelic insertions separated by 100 bp or more were assumed to represent distinct genes. When the separation was <100 bp, usually only one (if verified) or neither line was retained in the primary collection. Rarely, this might have led to the loss of valid lines, for example, in cases of overlapping genes or intragenic complementation, but it allowed us to discard nearly 100 questionable strains for the collection.

After completing these tests 1045 lines remained in the primary collection. Flanking sequence information has been obtained from 921 of the lines in this final group (88%). Accession numbers for each strain are listed in Tables 4 and 5 ("Sequence"). These sequences, including the position of the insertion, are listed on the BDGP website (http://www.fruitfly.org/p_disrupt/).

Associating primary collection lines with genes: The primary collection provides an opportunity to link ~1000 Drosophila genes with a genetic phenotype. Because these strains and genetic data have been publicly available from the inception of the project, the Drosophila research community has extensively utilized many lines from the primary collection (and the precursor raw collections). Publications describing at least 250 different Drosophila genes have employed strains from the collection (see Tables 4 and 5, "References"). In many cases, the *P*-element disruption strain played a major role in the initial characterization of the gene in question.

To identify as many additional genes as possible the *P*-element flanks were searched against all Drosophila sequences in GenBank and ~26 Mb of genomic sequence (most searches are current as of December 1998). To test the accuracy of flanking sequence recovery, the polytene location of the *P* element in each of the 286 lines whose flanking sequences matched genomic sequence determined by BDGP was compared to the independently mapped polytene location of the corresponding P1 clones. Only a few discrepancies resulted, presumably due to the rare recovery of sequence from a cryptic *P* element, and in these cases a correct flanking sequence was sought. These searches provided a wide variety of valuable information. They confirmed most of the 250 published gene assignments, identified many additional characterized Drosophila genes disrupted by strains in the collection, and molecularly positioned the insertion sites within all these loci. Of the additional Drosophila genes, 55 had previously been characterized only at the molecular level (Table 6).

Further links to well-characterized genes were discovered by associating the insertions with Drosophila transcripts defined by EST sequencing. About 48,000 Drosophila EST sequences were available for these comparisons. A total of 376 insertions were located close to or within an EST sequence, usually near the 5′ end (see Tables 4 and 5). Mutation-causing *P* elements are known to preferentially cluster in the 5′ region of the affected genes (see Spradling *et al.* 1995), a tendency that probably increases the chance of recovering overlaps between the short flanking sequences and 5′ ESTs. For each line with a matching EST, the relevant "clot" (consensus sequence of overlapping ESTs) sequence was conceptually translated and used to search protein databases. These comparisons associated 76 more primary collection lines with previously undescribed Drosophila genes encoding proteins related to characterized genes from other species (Table 7). These new Drosophila genes have been named on the basis of the name of their ortholog. Genes are listed in Table 7 only if there is a strong match within the region of overlap and if a study of the ortholog's properties has been published. BDGP has determined complete complementary DNA (cDNA) sequences for some of these genes (Table 7). The approaches described so far linked 450 primary collection lines with known Drosophila genes or with orthologs of characterized genes in other organisms (Tables 4 and 5).

Although the insertions in the remaining lines were not associated with a well-characterized gene or ortholog, it was still possible to link many of them with predicted transcripts and ORFs. The sequence comparisons associated the insertions in 135 additional lines with ESTs whose clots either predicted novel proteins or matched proteins conceptually encoded by ESTs or ORFs from other organisms. BLAST reports of these searches, including periodic updates, are available by searching the BDGP website using the appropriate EST (Tables 4 and 5). Finally, the insertions within 138 of the remaining lines not associated with genes or ESTs were localized within sequenced portions of the Drosophila genome. Bioinformatic analyses of the sequences flanking these insertions reveal candidate ORFs, although such studies have not yet been carried out systematically. In sum, therefore, 706 of the 1045 primary collection strains (67%) already link known or candidate genes with mutant phenotypes. It should be

TABLE 6

New gene-mutant associations

Gene	Strain	Site	Accession
5SRNA (5S RNA gene repeat)	l(2)03068	56F1-2	X06938
Acer (angiotensin-converting enzyme-like)	l(2)k07704	29D1-2	X96913
Adf-1 (Adh distal factor)	l(2)01349	42C1-4	M37787
Ani (anilin, actin-binding protein)	l(2)03427	43E1-5	X89858
anon1A4 (fast evolving gene)	l(2)k05815	39E3-4	AF005846
ari2	l(2)07768	58D1-2	AJ010169
Atu (putative transcriptional regulator)	l(3)s1938	83B4-7	U75467
brat (brain tumor)	l(2)k06028	37C6-7	
Cdc27	l(3)L7123	65F1-2	U18298
dbe (dribble)	l(2)k05428	21D4-E1	Z96931
Drs1 (putative ribosome biogenesis regulator)	l(2)k09514	44C1-2	1360171
Ef1&agr;48D (trans. elongation factor EF-1alpha)	l(2)01275	48C5-6	X06869
eIF-4E (translation initiation factor 4E, cap-binding)	l(3)07238	67B1-2	U63033
Elf (Ef-1 alpha-like peptide chain release factor)	l(2)k06909	33E5-7	U88868
Fer1HCH (25kD and 26kD ferritin subunits)	l(3)00451	99F1-2	U91524
Fur1 (furin-like protein)	l(3)rL205	96C1-2	L12368
Ggamma1 (G protein gamma)	l(2)k08017	44F3-4	L28751
Est2	n(2)06253	53F1-5	M95198
HmG-CoAR (HMG Coenzyme A reductase)	l(3)01152	95B5-6	M21329
how (KH RNA binding protein)	l(3)j5B5	94A1-2	U72331
Hsc70-4 (hsp70 cognate 4)	l(3)03550	88E8-9	M36113
Hsc70-5 (hsp70 cognate 5)	l(2)k04907	50E4-7	L01502
hspr (hsp related)	l(2)03659	45D1-2	V00219
lin19 (lin-19/cul-1/cdc53 related)	l(2)k01207	43F1-2	L41642
Lip (Rm62 helicase)	l(3)01086	83D4-5	X52846
me31B (DHH1|Putative DEAD/DEAH RNA helicase)	l(2)k06607	31B1-2	M59926
Mov34 (proteasome subunit p40)	l(2)k08003	60D1-2	M64641
mub (mushroom bodies)	l(3)04093	79B1-2	X99340
ND-AcC (NADH-ubiq. oxidoreductase acyl-carrier)	l(3)j4A6	61F6-7	AJ000879
Nop60B (nucleolar protein)	l(2)k05381	60C1-2	AF017230
Nurf-38 (inorganic pyrophosphatase)	l(2)k16102	60D5-6	AF085600
Nup32D (nucleoporin-like mammalian Nup155p)	l(2)01501	32D1-2	X94613
pAbp (poly(A)-binding protein)	l(2)k10109	55B5-6	L05109
Pha (PHD-containing ATPase)	l(3)j3D4	76D3-4	
porin (mitochondrial porin)	l(2)k05123	32B1-2	AJ000880
Rab6	l(2)08323	33D1-2	D84314
Rel (Rel/NF-kappa B family member)	l(3)neo36	85C1-13	U62005
Rga	l(3)03834	83B4-5	U75467
rnh1 (ribonuclease H1 (rnh1) gene)	l(2)k07624	43F1-2	AF032921
RnrL	l(2)k06709	31D8-9	U09369
RpI135	l(2)k16513	21C1-2	X17298
RpL30	l(2)k09918	56F8-15	
RpP0 (ribosomal protein P0)	l(3)01544	79D1-2	M25772
RpS13	l(2)k09614	29B1-2	X91853
RpS26	l(2)04553	36E1-4	X14247
S6k (RPS6-p70-protein kinase)	fs(3)07084	64E8-12	
smid (smallminded: AAA ATPase family)	l(3)j6B8	66A1-2	X99207
snRNA:U4:39B	l(2)k09410	39B1-2	D00043
Src42A	l(2)k10108	42A1-2	D42125
spdo	l(3)00848	99F10-11	U92490
Syb (synaptobrevin)	l(2)k07703	46F9-10	L14270
tRNA:Y1:85Ab	l(3)j1B9	85A9-10	M21611
TfIIFbeta (transcription initiation factor TfIIFβ)	l(3)j3C1	86C3-4	U25188
Uba1 (ubiquitin activating enzyme 1)	l(2)03405	46A1-2	Y15895
UbcD2 (ubiquitin conjugating enzyme 2)	l(2)k13206	32A4-5	X92663
Ubi-p63E	l(3)05634	63F5-6	M22428
zf30C (zinc finger protein 30C)	l(2)k02506	30C7-8	AF035275

The table lists new gene-mutant associations defined by the primary collection. Most are Drosophila genes (Gene) known previously from molecular data only (see sequence Accession). In each case, the name of the primary collection strain that disrupts the gene (Strain) and the cytogenetic location of its *P*-element insertion (Site) are given. Where no accession is listed, either the gene was previously known by mutation only, and the primary collection strain (Strain) indicated the nature of its encoded product, or no accession was available. New associations in the Adh region are omitted; see Ashburner *et al.* (1999).

TABLE 7

New genes

Gene	Strain	Site	Accession	cDNA
Aac11 (putative apoptosis inhibitor)	l(2)k06710	36C8-11	U83857 (H)	
Aats-ile (isoleucyl-tRNA synthetase)	l(3)00827	79E1-2	I59314 (H)	
Aats-thr (threonyl-tRNA synthetase)	l(2)k04203	33C4-5	M63180 (H)	
Aats-gln (glutaminyl-tRNA synthetase)	l(3)04561	96C8-9	X76013 (H)	
Aats-val (glutaminyl-tRNA synthetase)	l(2)k14804	49F7-8	P07806 (S)	
AconM (aconitase, mitochondrial)	l(2)07054	39B1-2	J05224 (O)	
Adk2 (adenylate kinase 2)	l(2)k16120	44B5-6	D13061 (R)	
Aly (transcriptional coactivator)	l(3)02267	84C1-2	U89876 (M)	
ATPCL (ATP:citrate lyase)	l(2)k09217	52D11-12	U18197 (H)	
Btf (transcription factor)	l(2)k10712	49D5-6	X74070 (H)	
Bub1 (spindle-assembly checkpoint kinase)	l(2)k06109	42A1-2	AF002823 (M)	AF132565
Cas (CAS/CSE1|segregation protein)	l(2)k03902	36B1-2	U33286 (H)	AF132562
Cct5 (T-complex Chaperonin)	l(2)06444	48E10-11	X75777 (A)	
Coprox (coproporphyrinogen oxidase)	l(2)k10617	27C6-8	Z28409 (H)	
Cops4 (COP9 complex subunit 4)	l(2)k08018	44A1-2	AF071314 (M)	
Dcap (adenylyl cyclase-associated protein)	l(2)k00619	21F1-2	M98474 (H)	AF132562
Dhap (aldehyde dehydrogenase type II)	l(2)03610	43D1-2	P30838 (H)	
Dhh1 (DEAD/DEAH RNA helicase)	l(3)rL562	67E1-4	1431254 (H)	
DNAprim (DNA primase)	l(3)j10B2	77B6-7	D13545 (M)	
Dph5 (Diphthamide methyltransferase)	l(3)L4910	94B4-5	M83375 (S)	
Dmn (dynamitin)	l(2)k16109	44F3-4	U50733 (H)	
Eb1 (APC-binding protein)	l(2)04524	42E1-3	U24166 (H)	AF132560
*Ef1*β (elongation factor 1 beta)	n(2)k04810	53E1-2	D13339 (I)	
EfTuM (elongation factor Tu mitochondrial)	l(3)L4569	89E7-9	X84694 (H)	
elF-3 (initiation factor el3 p40 subunit)	l(2)k09003	25C1-2	U54559 (H)	
Etf (electron transfer flavoprotein)	l(2)02516	48C1-2	J04058 (H)	
Fatp (long chain fatty acid transport protein)	l(2)k10307	31F4-5	P97849 (R)	
Fkbp13 (rapamycin-binding protein)	l(2)00734	57E3-4	M77831 (M)	AF132555
Fpps (farnesyl pyrophosphate synthetase)	l(2)k06103	47F1-2	X76026 (L)	AF132554
Ft1 (related to fused toes 1 protein)	ms(2)05704	46C1-12	S33513 (M)	
GlyP (glycogen phosphorylase, brain)	l(2)k07918	22C1-2	U47025 (H)	
Hmgs (HMG CoA synthase)	l(2)06214	53C1-2	X96617 (S)	
Hrr25 (casein kinase I related)	ms(3)04895	89B1-22	1370424 (S)	
hsp60B (heat shock protein 60 related)	ms(2)06619	21D1-4	X99341 (E)	
Int6 (Int6 homologue)	l(3)10547	73D1-2	L35556 (M)	AF132551
JadBp (Jun activation domain binding protein)	l(3)L4032	89D1-2	U65928 (H)	AF132563
Kary&bgr;3 (karyopherin beta 3)	l(3)j3A4	82D1-2	U72761 (H)	AF132561
KdelR (KDEL receptor)	l(2)k00311	31E1-2	X63745 (H)	AF132559
lolal (resembles Lola-like D. Hydei protein)	l(2)k02512	55B5-10	Y14994 (Y)	
MdaPk (myotonic dystrophy associated kinase)	l(2)09373	60B10-11	L19268 (H)	
MI02 (inhibits mitosis when overexpressed)	v(3)03847	100F4-5	Q09329 (P)	
Mrp17 (mitochondrial ribosomal protein L5/L17)	l(3)10534	66E2-3	X79865 (H)	
Msp1 (mitochondrial sorting protein)	l(2)08774	31D1-2	1323004 (H)	
Neos (neosin)	l(3)L2249	65E10-11	X94344 (M)	
Nik (Ste20-related kinase)	l(3)j1E2	62E5-8	U88984 (M)	
Nmt1 (N-myristoyltransferase)	l(3)j1C7	66B10-11	U14913 (S)	AF132556
Nnp-1 (resembles mouse NNP-1)	l(2)k07826	34B6-7	U79773 (M)	
Paf-AH&bgr; (platelet-activating factor fl/LIS1)	l(2)k11702	52F3-5	AF016049 (R)	
Pdsw (ubiquinone oxidoreductase cx)	l(2)k10101	23F3-4	X63224 (B)	
Phas1 (insulin-stimulated elF-4E binding protein)	l(2)k07736	23F5-6	U75530 (M)	AF132557
PpsM (mitochondrial proton/phosphate symporter)	l(3)00564	70E1-2	X60036 (H)	
Pros44.5 (p44.5 proteasome subunit)	l(2)k00103	51C1-2	AB003102 (H)	
Pros26S (26S proteasome subunit)	l(3)04210	89A1-2	Q63569 (R)	
Prp19 (Pre-mRNA splicing factor)	l(2)07838	55D1-2	H35544 (R)	
Rab5 (Rab5 homologue)	l(2)k08232	22E1-2	Z27110 (D)	
Rfc38 (DNA replication factor, 38k subunit)	n(2)k13807	32D4-5	L07541 (H)	
RFeSP (Rieske iron-sulfur protein)	l(2)k11704	22A3-4	M34336 (B)	

(continued)

TABLE 7

(Continued)

Gene	Strain	Site	Accession	cDNA
Rpp30 (RNaseP protein p30)	l(2)k01901	21B4-6	U77665 (H)	
RRM domain protein	l(3)02094	65A7-9	X06347 (H)	
sec13 (secretory pathway gene)	l(3)01031	94F1-2	L09260 (H)	
Sip1 (SRY interacting protein)	l(2)06373	54B4-5	U82108 (H)	
SrpR&bgr; (signal recognition particle receptor fl)	l(3)rK561	66D10-13	U17343 (M)	
Srp54 (signal recognition particle 54k)	l(3)01418	64C9-10	X86373 (H)	
SsR&bgr; (signal sequence receptor fl)	l(3)s1939	72D8-9	X53529 (D)	
OstStt3 (Oligosaccharyl transferase)	l(3)j2D9	96B9-10	P46975 (C)	AF132552
Tap&dgr; (translocon-associated protein)	l(2)k17005	47F4-9	X90582 (M)	
Tbl1 (beta transducin-like 1 protein)	l(2)k16213	21C2-3	Y12781 (H)	
Tom34 (outer mitochondrial 34 kD translocase)	l(3)03692	84E10-11	U58970 (H)	
Uch-L3 (ubiquitin C-terminal hydrolase related)	l(3)j2B8	67B10-11	M30496 (H)	AF132567
VhaD (vacuolar H-ATPase subunit D)	l(2)k07207	52A9-11	U11927 (B)	
VhaG (vacuolar H-ATPase subunit G)	l(3)05113	92A13-14	Q25532 (N)	
Vcp (valosin-containing protein)	l(2)03775	46D1-2	M30143 (O)	AF132553
Xe7 (activated lymphocyte surface protein)	l(3)02248	83B6-7	Q02040 (H)	
Xbp1 (X box binding protein-1)	l(2)k13803	57C1-2	M31627 (H)	
Zfrp8 (zinc finger protein RP-8)	l(2)k13705	60B11-13	U1090 (M)	

New Drosophila genes (Gene) orthologous to previously studied genes (Accession) from the indicated species (A, *A. sativa*; B, *B. tarus*; C, *C. elegans*; D, *C. lupus*; E, *D. melanogaster*; H, *H. sapiens*; I, *B. mori*; M, *M. musculans*; L, *K. lactis*; N, *Manduca*; O, *S. scrofa*; R, *R. norveigicus*; S, *Saccharomyces cerevisiae*; P, *S. pombe*; Y, *D. hydei*). In each case, the name of the primary collection strain that disrupts the gene (Strain) and the cytogenetic location of its *P*-element insertion (Site) is given. Accession numbers for genes whose cDNAs have been completely sequenced by BDGP and deposited in GenBank are also listed (cDNA).

possible to make most of the remaining gene-mutant associations by the time genome and EST sequencing nears completion.

***P*-element selectivity:** This study reveals the identity of most genes that are hotspots for *P*-element insertion on the autosomes (Tables 4 and 5, "Alleles"). We searched for common properties that might explain their efficiency as *P*-element insertional targets. Hotspot genes are not associated with generally high transcription levels, because only 30% of the genes in the primary collection with more than five alleles have an associated EST sequence, compared to 36% for the collection as a whole. Hotspot genes might be those actively transcribed in premeiotic germline cells, where *P* elements usually transpose; however, the few genes in the collection whose transcripts are abundant in early germ cells, including *vasa*, *bam*, and *hsp83*, were each hit only once. Indeed, our comparisons uncovered no common biological features such as size, location, or regulation that might explain why hotspot genes are highly susceptible to *P*-element insertion.

We also considered whether strong preferences exist for insertion within certain classes of genes among all those disrupted in the collection. The primary collection includes an estimated 30% of readily mutable autosomal genes. Genes involved in signal transduction were usually well represented, because the collection mutates ~50% of all autosomal genes known to be involved in the EGFR, *dpp*, *ras*, *wg*, *hh*, or *N* signaling pathways. In addition, disruptions were obtained in 46% of autosomal posterior group genes, 31% of trithorax and Polycomb group genes, but only 14% of ribosomal protein genes. It remains unclear if these differences reflect more than the research priorities of the Drosophila research community.

Not all insertion sites were associated within protein-coding genes. One *P* element was located within a 5S rDNA repeat and four interrupted tRNA clusters. Nine lines, two of which disrupt the genes *Distal-less* and *fruitless*, were found by sequence analysis to contain insertions within the LTR sequence of a Drosophila retrotransposon related to the yoyo element of the Mediterranean fruit fly *Ceratitis capitata* (Zhou and Haymer 1997; see also FBgn0021759). The abundance of this element was low overall and all the insertions clustered in a small part of the LTR, a likely hotspot. Two other multicopy target sites were the telomere associated sequence (TAS) element, with six insertions, and the hoppel element, with one insertion. Both elements have been shown previously to be frequent targets of *P*-element insertion (Karpen and Spradling 1992; Zhang and Spradling 1995; D. Stewart and A. Spradling, personal communication). Because most insertions within repetitive sequences would not be expected to disrupt vital functions, these observations probably reflect which repetitive target sequences are frequently located within the introns or immediate flanks of vital genes in the strains used.

Modeling mutational saturation: The gene disruption project provides a much larger and better-characterized data set than has been previously available for analyzing the site specificity of *P*-element transposition. This is an important question for determining the appropriate strategy to expand the collection. The insertional specificity of *P* elements must be extremely broad to achieve complete or nearly complete coverage of all Drosophila genes. In contrast, previous studies inferred that a significant percentage of Drosophila genes, perhaps as great as 50%, are refractory to mutation using *P* elements (see Kidwell 1986; Török *et al.* 1993). If true, this would imply that a different method of mutagenesis is needed to complete the gene disruption project (Spradling *et al.* 1995). However, these conclusions remain highly uncertain, because previous studies of saturation behavior utilized raw collections of unverified lines that differ in *P*-element content and did not correct for locus-specific differences in mutagenesis rates. The total number of different genes mutagenized clearly rises more slowly than expected by assuming that nearly all genes are equally susceptible to *P*-element insertion. However, this observation alone cannot distinguish between the presence of genes refractory to *P*-element insertion and the presence of gene classes that differ significantly in *P*-element mutability. Fortunately, the very information gathered to build the primary collection also allows one to more accurately deduce the saturation behavior of *P* elements.

We focused on the large subset of the *P*-element lines from the collections whose insertions lie within the boundaries of validated deficiencies. Within this group, for a known number of total lines (transposition events), the number of genes mutated and how many times each was hit has been determined with complete accuracy. Because the deficiencies included a majority of chromosome *2* and *3* genes (60.3 and 62.0%, respectively), and should be distributed effectively at random, this sample should accurately represent all insertions that cause a phenotype. When we analyzed the distribution of insertional mutations among this set of genes, it was clear that the data did not fit a simple Poisson distribution (see materials and methods; Table 1). The most obvious problem was the hotspot loci. On chromosomes 2 and 3, just 18 or 32 loci account for 38 or 39% of all insertions, respectively. However, even after subtracting the contribution of these hotspot loci, the distribution of gene mutabilities remained skewed (see materials and methods; Table 1). Consequently, a class of warmspot genes was inferred whose mutability is intermediate between the hotspot loci and the large group of low mutability coldspot genes. Assuming the existence of three major mutability classes allowed a good fit to the data.

This model provides several useful insights into *P*-element behavior. The third chromosome is predicted to contain 27 hotspot loci + 191 warmspot loci + 1017 coldspot loci, while the second chromosome should house 47 hotspot genes + 177 warmspot genes + 1097 coldspot genes. Despite accounting for only 17% of all genes, the 368 warmspot and 74 hotspot genes account for ~70% of all transposition events. As a result, virtually all the hotspot loci and 80–90% of the warmspot loci have already been defined by strains in the primary collection. On the other hand, only 22–28% of the coldspot loci have so far been disrupted. However, assuming that there are 1400 vital loci per major autosome (Miklos and Rubin 1996), and considering that 93% of the disruptions in our collection are of vital genes, then the model predicts that at least $2556 \times 0.93/2800 = 85\%$ of vital genes can eventually be mutated using *P* elements. Thus, the existence of the hotspot and warmspot genes is the reason that mutational saturation proceeds more slowly than expected on the basis of a single class Poisson analysis, but the final level of saturation is higher than previously appreciated. Indeed, if gene mutabilities actually vary more broadly than three discrete classes, as seems likely, the true level of saturation will exceed 85%. There is no reason to suspect that *P*-element insertional preferences differ between vital and nonvital genes, so the conclusions drawn here should apply to Drosophila genes generally. These results suggest that a much larger fraction of Drosophila genes than previously supposed, at least 85% and possibly 100%, are susceptible to inactivation by *P*-element insertion.

DISCUSSION

Collections of gene disruptions as tools for functional genomics: It is now possible in theory to mutate virtually any gene that has been molecularly identified in the major multicellular model organisms and to isolate the mutant allele on a standard genetic background free of secondary lesions. In practice, obtaining mutants remains a time-consuming task that constitutes the largest current impediment to progress in understanding gene function *in vivo*. While it has become widely accepted that gene sequence and structure can be more efficiently analyzed on a genome-wide scale, a similar consensus on the value of whole genome gene disruption has been slow to develop. As a result, linking genes with mutations remains a cottage industry pursued by individual laboratories. The work reported here has been motivated by the belief that complete gene mutation libraries are feasible and have the potential to greatly accelerate the rate at which gene function can be analyzed. We feel that whole genome mutant collections belong together with complete genome and cDNA sequences as essential tools for future biological research.

The BDGP gene disruption library represents a significant step toward the ultimate goal of stockpiling an identified mutation in every Drosophila transcription unit. The current collection of single *P*-element insertions provides a particularly useful type of link between the genetic and molecular properties of ~1000 different autosomal genes that can mutate to a readily recognizable phenotype. This is more than the number of genes that have been characterized at both the genetic and molecular levels in any of the other widely used model multicellular eukaryotes, including Arabadopsis, *C. elegans*, zebrafish, or mice, and exceeds the number of gene-mutation links known in humans. As a reflection of its utility, lines from the BDGP collection have been utilized in publications characterizing more than 250 different genes since 1988 (Tables 4 and 5).

Expanding the collection: Because the Drosophila genome is believed to house ~12,000 genes (Miklos and Rubin 1996), the current primary collection is still far from complete. Two basic approaches can be considered for expanding its coverage. A targeted strategy would avoid reisolating new mutations in genes that have already been disrupted in the existing collection or by individual Drosophila researchers. A general strategy for identifying mutations in any gene encoding a protein that can be detected with a specific antiserum has been developed (Dolph *et al.* 1992). However, a substantial number of genes that express proteins only at low levels may be refractory to disruption by this approach. Consequently, continuing the insertional mutagenesis strategy used previously in some form appears to be the most promising approach to completing the collection.

Significant improvements are possible in the short term by incorporating several new collections of insertions that have already been constructed since the project was initiated (Erdelyi *et al.* 1995; Deak *et al.* 1997; Rørth *et al.* 1998). The third chromosome collection described by Deak *et al.* (1997) is similar in size to the collection of Török *et al.* (1993) on chromosome 2, but preliminary estimates by the authors indicate a higher screen efficiency. Incorporating these lines into the existing collection should increase the number of third chromosome lines to >600 and equalize the saturation levels of the two major autosomes.

It will also be of value to carry out new mutagenesis screens. A major variable in the generation of single *P*-element-induced mutations is the wide variation in screen efficiency that is documented here (Table 2). One factor that can affect screen efficiency is the overall rate of *P* transposition. High transposition rates like those in the screen of Török *et al.* (1993) produce an excess of lines with more than one *P*-element insertion (>23% in this case). High transposition rates probably also cause secondary mutations as elements transpose and excise at multiple sites over several germ cell division cycles. However, our results imply that the rate of transposition and amount of secondary damage are not always correlated and are not simply a function of the *P* elements used (Table 2). Both Bier *et al.* (1989) and Török *et al.* (1993) employed the PlacW and Δ2-3 *P* elements but obtained very different frequencies of multiple insert lines, rates of background mutation, and overall screen efficiencies. In contrast, the screen of Cooley *et al.* (1988) using PUChsneo and a weak mobilizing *P* element exhibited a low transposition rate but still gave an efficiency of only ~50%. Consequently, our results suggest that currently unidentified factors in the genetic backgrounds used for *P*-element mutagenesis affect the prevalence of damage at chromosomal sites that do not retain *P*-element sequences. Unfortunately, the nature of these factors remains poorly understood.

The number of new lines that needs to be characterized to substantially complete the gene disruption project can be estimated from our analysis of saturation. The genome contains ~3600 vital genes, at least 3100 of which fall into the coldspot class. Statistically, twice this number of insertions, 6200, must be recovered in this class of genes to achieve 87% saturation. Because only 30% of raw insertions target the coldspot class, and because the best screens produce only 85% verified single insert lines, achieving 87% saturation would require the isolation and analysis of $6200/(0.3 \times 0.85) = 24{,}300$ autosomal insertions associated with phenotypes. This represents about six times as many lines as were analyzed in the current project.

A molecular strategy for finishing the mutation library: Even a project of this size is feasible, although a very large effort would be required. However, a continuation of the current approach would not address the estimated two-thirds of all genes that do not mutate to a readily detectable phenotype in genetic screens. To obtain *P*-element insertions that disrupt such genes, it will be necessary to look directly for changes in their structure. With large amounts of genomic and EST sequences becoming available and a strong commitment to completing the Drosophila genome sequence within 1–3 years (Collins *et al.* 1998; Venter *et al.* 1998), a strategy based entirely on molecular mapping is becoming feasible. A new generation of *P*-element misexpression vectors (Rørth 1996) are attractive candidates for use with this approach. These insertions not only can disrupt genes but also are frequently able to program the controlled misexpression of the affected protein. This option should accelerate the collection of functional information, especially on the many genes whose loss does not produce an immediately recognizable phenotype.

We propose to inaugurate a phase two gene disruption project whose goal would be to disrupt all Drosoph-

ila genes, regardless of phenotype. Flanking DNA will be recovered from a large number of raw insertion lines and sequenced, much as was done with the primary collection lines in the current collection. The short sequences obtained will allow most new insertions to be precisely positioned on the genomic sequence. Consulting EST and cDNA sequences, gene predictions, ORF homologies, and other relevant data in the vicinity of the insertion sites will allow rapid predictions as to whether each new insertion is likely to disrupt or misexpress an ORF not currently represented in the collection. Lines that do not appear to do so would be quickly discarded. Recently, this strategy has received a valuable test within the fully sequenced 2.9-Mb Adh region (Ashburner *et al.* 1999). By mapping all available *P* elements onto the genomic DNA sequence, not just those causing phenotypes as described here, the number of gene-mutation links was increased substantially (see Table 4).

The phase two strategy has several distinct advantages. First, it broadens the project to include all Drosophila genes. In addition, it greatly simplifies the work required to characterize new candidate lines, compensating in part for the much larger number of lines that will need to be analyzed. Polytene localizations are unnecessary, because multiinsert lines can be detected through their production of more than one distinct *P*-element flanking sequence. Balancing most of the newly mutagenized chromosomes is not required. Genetic complementation is not necessary, because redundant lines can increasingly be identified on the basis of their location. However, there are several requirements for success. First, the Drosophila genome sequence must be completed in a timely manner. Second, semiautomated methods for recovering and sequencing flanking DNA segments must be further improved. Finally, bioinformatic tools to assist decision making about line retention must be developed.

We can calculate the approximate number of lines that will need to be analyzed during the phase two project. About 11,000 of the estimated 12,000 Drosophila genes are predicted to fall into the coldspot class, assuming that the *P*-element mutability of all genes is similar to that of vital genes. Therefore, if 30% of new insertions fall in the coldspot class as in the case with lethal insertions, and 95% of raw lines contain only one insertion, then $2 \times 11{,}000/(0.3 \times 0.95) = 77{,}000$ lines would be required for 87% saturation. However, two observations suggest that some unselected insertions will fail to disrupt any gene, increasing the total number of lines that will need to be analyzed. First, *P* elements are attracted to at least some repetitive sequences such as yoyo, TAS, and hoppel, which are often located at nonmutagenic sites within the genome. The fraction of insertions that land in such sites might be significant. Second, *P* insertions that cause phenotypes cluster around the 5′ region of genes (Spradling *et al.* 1995; data not shown). Previously, insertions located too far upstream from transcription start sites, or at nonmutagenic sites within large introns, have been edited out by the requirement for a phenotypic effect. In phase two, they would be recovered and analyzed, lowering efficiency.

The relative fraction of unselected insertions that disrupt genes can be estimated, however. If all insertions mutated genes, then 33% of new transpositions should cause a recognizable phenotype, because about one-third of genes are thought to mutate in this manner. Instead, only ~15% of raw insertions recovered on clean chromosomes cause a recognizable phenotype (see citations in Table 2). Consequently, as many as $77{,}000/0.5 = 154{,}000$ insertions might need to be screened to obtain 87% saturation across all Drosophila genes. However, in practice, this may be an overestimate. *P* elements can be excised imprecisely to generate deletions adjacent to the insertion site. Because of the large number of mapped insertions that will be available by the time phase two is only partially complete, a strategy in which some genes are disrupted by excising nearby nonmutagenic insertions might substantially reduce the final number of strains that need to be generated and analyzed.

A gene disruption library represents a fundamental and indispensable resource for analyzing gene function on a genome-wide scale. The BDGP gene disruption project has already accelerated studies of Drosophila gene function and is likely to be even more valuable as coverage increases. A pilot screen for phase two has already been completed in collaboration with several laboratories (Rørth *et al.* 1998). A total of 2400 lines from this project have been mapped and initially analyzed (BDGP, unpublished results; see http://www.fruitfly.org/bfd/). We believe that researchers using Drosophila (and other model multicellular organisms) are rapidly approaching an era where obtaining mutations, the basic tools for understanding gene function *in vivo*, will no longer limit the progress of research.

BDGP acknowledges all those researchers who participated in constructing the strains that were used in this project. These include L. Ackerman, M. Alvarado, S. Barbel, C. Berg, E. Bier, S. Bockheim, M. Boedingheimer, R. Carretto, Z. Chang, L. Cooley, M. Fuller, U. Gaul, R. Glaser, E. Grell, B. Harkins, M. Heck, L. Higgins, L. Jan, Y.-N. Jan, G. Karpen, R. Kelley, I. Kiss, A. Laughon, K. Lee, L. Lee, G. Mardon, K. McCall, D. McKearin, C. Montell, D. Montell, T. Overbode, B. Price, J. Riesgo, M. Scott, S. Shepherd, R. Smith, D. Thompson, T. Tick, T. Török, J. Tower, T. Uemura, H. Vassin, E. Verheyen, S. Wasserman, and L. Yue. We are also grateful to many workers who in the course of this study communicated complementation results and other information on specific *P*-element strains. In particular, John Roote and Paul Lasko shared complementation data for *2L* divisions 24-36 and 37-38. H. Bellen (various), Erica Roulier (29A), Ken Howard (45), Jordan Raff (46A), Elliott Goldstein (46), Robert Burgess (47EF), Claire Russell (49EF), Paul Wes (52E), and Boris Dunkov (99F) contributed and confirmed results in the cytogenetic

regions indicated. We thank A. deGrey for assistance in analyzing chromosome *2* data. This work was supported by a genome center grant (P50NIHHG750) from the National Institutes of Health. A.C.S. and G.M.R. are Howard Hughes Medical Institute Investigators.

LITERATURE CITED

Allende, M. L., A. Amsterdam, T. Becker, K. Kawakami, H. Gaianoand *et al.*, 1996 Insertional mutagenesis in zebrafish identifies two novel genes, *pescadillo* and *dead eye*, essential for embryonic development. Genes Dev. **10:** 3141–3155.

Alsina, B., F. Serras, J. Baguna and M. Corominas, 1998 *patufet*, the gene encoding the *Drosophila melanogaster* homologue of selenophosphate synthetase, is involved in imaginal disc morphogenesis. Mol. Gen. Genet. **257:** 113–123.

Andrew, D. J., A. Baig, P. Bhanot, S. M. Smolik and K. D. Henderson, 1997 The Drosophila *dCREB-A* gene is required for dorsal/ventral patterning of the larval cuticle. Development **124:** 181–193.

Arora, K., H. Dai, S. G. Kazuko, J. Jamal, M. B. O'Connor *et al.*, 1995 The Drosophila *schnurri* gene acts in the Dpp/TGF-beta signaling pathway and encodes a transcription factor homologous to the human MBP family. Cell **81:** 781–790.

Ashburner, M., 1990 *Drosophila: A Laboratory Manual*. Cold Spring Harbor Laboraory Press, Cold Spring Harbor, NY.

Ashburner, M., S. Misra, J. Roote, S. Lewis, R. Blazej *et al.*, 1999 An exploration of the sequence of a 2.9-Mb region of the genome of *Drosophila melanogaster*: the *Adh* region. Genetics **153:** 179–219.

Barrett, K., M. Leptin and J. Settleman, 1997 The Rho GTPase and a putative RhoGEF mediate a signaling pathway for the cell shape changes in Drosophila gastrulation. Cell **91:** 905–915.

Baumgartner, S., D. Martin, C. Hagios and R. Chiquet-Ehrismann, 1994 *Ten-m*, a Drosophila gene related to tenascin, is a new pair-rule gene. EMBO J. **13:** 3728–3740.

Baumgartner, S., D. Martin, R. Chiquet-Ehrismann, J. Sutton, A. Desai *et al.*, 1995 The HEM proteins: a novel family of tissue-specific transmembrane proteins expressed from invertebrates through mammals with an essential function in oogenesis. J. Mol. Biol. **251:** 41–49.

Baumgartner, S., J. T. Littleton, K. Broadie, M. A. Bhat, R. Harbecke *et al.*, 1996 A Drosophila neurexin is required for septate junction and blood-nerve barrier formation and function. Cell **87:** 1059–1068.

Begemann, G., A. M. Michon, L. van der Voorn, R. Wepf and M. Mlodzik, 1995 The Drosophila orphan nuclear receptor Seven-up requires the Ras pathway for its function in photoreceptor determination. Development **121:** 225–235.

Bellaiche, Y., I. The and N. Perrimon, 1998 Tout-velu is a Drosophila homologue of the putative tumour suppressor EXT-1 and is needed for Hh diffusion. Nature **394:** 85–88.

Bellen, H. J., C. J. O'Kane, C. Wilson, U. Grossniklaus, R. K. Pearson *et al.*, 1989 P-element-mediated enhancer detection: a versatile method to study development in Drosophila. Genes Dev. **3:** 1288–1300.

Berg, C., and A. Spradling, 1991 Studies on the rate and site-specificity of *P* element transposition. Genetics **127:** 515–524.

Bhat, M. A., A. V. Philp, D. M. Glover and J. Bellen, 1996 Chromatid segregation at anaphase requires the *barren* product, a novel chromosome-associated protein that interacts with Topoisomerase II. Cell **87:** 1103–1114.

Bhatt, A. M., T. Page, E. J. Lawson, C. Lister and C. Dean, 1996 Use of Ac as an insertional mutagen in Arabidopsis. Plant J. **9:** 935–945.

Bier, E., H. Vassin, S. Shepherd, K. Lee, K. McCall *et al.*, 1989 Searching for pattern and mutation in the Drosophila genome with a P-lacZ vector. Genes Dev. **3:** 1273–1287.

Boulianne, G. L., A. de la Concha, J. A. Campos-Ortega, L. Y. Jan and Y. N. Jan, 1991 The Drosophila neurogenic gene *neuralized* encodes a novel protein and is expressed in precursors of larval and adult neurons. EMBO J. **10:** 2975–2983.

Braun, A., J. A. Hoffmann and M. Meister, 1997 Drosophila immunity: analysis of larval hemocytes by P-element-mediated enhancer trap. Genetics **147:** 623–634.

Brummel, T., S. Abdollah, T. E. Haerry, M. J. Shimell, J. Merriam, L. Raftery, J. L. Wrana and M. B. O'Connor, 1999 The Drosophila activin receptor baboon signals through dSmad2 and controls cell proliferation but not patterning during larval development. Genes Dev. **13:** 98–111.

Burns, N., B. Grimwade, P. B. Ross-Macdonald, E. Y. Choi, K. Finberg *et al.*, 1994 Large-scale analysis of gene expression, protein localization, and gene disruption in *Saccharomyces cerevisiae*. Genes Dev. **8:** 1087–1105.

Campbell, G., H. Goring, T. Lin, E. P. Spana, S. Anderson *et al.*, 1994 RK2, a glial-specific homeodomain protein required for embryonic nerve cord condensation and viability in Drosophila. Development **120:** 2957–2966.

Campbell, S. D., F. Sprenger, B. A. Edgar and P. H. O'Farrell, 1995 Drosophila Wee1 kinase rescues fission yeast from mitotic catastrophe and phosphorylates Drosophila Cdc2 in vitro. Mol. Biol. Cell **6:** 1333–1347.

Castrillon, D. H., P. Gonczy, R. Rawson, C. G. Eberhart, S. Viswanathan *et al.*, 1993 Toward a molecular genetic analysis of spermatogenesis in *Drosophila melanogaster*: characterization of male-sterile mutants generated by single P element mutagenesis. Genetics **135:** 489–505.

Chang, H. C., and G. M. Rubin, 1997 *14-3-3 epsilon* positively regulates Ras-mediated signaling in Drosophila. Genes Dev. **11:** 1132–1139.

Chang, Z., B. D. Price, S. Bockheim, M. J. Boedigheimer, R. Smith *et al.*, 1993 Molecular and genetic characterization of the Drosophila *tartan* gene. Dev. Biol. **160:** 315–332.

Clark, K. A., and D. M. McKearin, 1996 The Drosophila *stonewall* gene encodes a putative transcription factor essential for germ cell development. Development **122:** 937–950.

Collins, F. S., A. Patrinos, E. Jordan, A. Chakravarti, R. Gesteland *et al.*, 1998 New goals for the U.S. human genome project: 1998–2003. Science **282:** 682–689.

Cooley, L., R. Kelley and A. C. Spradling, 1988 Insertional mutagenesis of the Drosophila genome with single *P* elements. Science **239:** 1121–1128.

Cooley, L., E. Verheyen and K. Ayers, 1992 *chickadee* encodes a profilin required for intercellular cytoplasm transport during Drosophila oogenesis. Cell **69:** 173–184.

Cui, X., and C. Q. Doe, 1992 *ming* is expressed in neuroblast sublineages and regulates gene expression in the Drosophila central nervous system. Development **116:** 943–952.

Davies, S. A., S. F. Goodwin, D. C. Kelly, Z. Wang, M. A. Sozen *et al.*, 1996 Analysis and inactivation of *vha55*, the gene encoding the vacuolar ATPase B-subunit in *Drosophila melanogaster* reveals a larval lethal phenotype. J. Biol. Chem. **271:** 30677–30684.

D'Avino, P. P., and C. S. Thummel, 1998 *crooked legs* encodes a family of zinc finger proteins required for leg morphogenesis and ecdysone-regulated gene expression during Drosophila metamorphosis. Development **125:** 1733–1745.

Deak, P., M. M. Omar, R. D. Saunders, M. Pal, O. Komonyi *et al.*, 1997 P-element insertion alleles of essential genes on the third chromosome of *Drosophila melanogaster*: correlation of physical and cytogenetic maps in chromosomal region 86E-87F. Genetics **147:** 1697–1722.

Demakov, S. A., V. F. Semeshin and I. F. Zhimulev, 1993 Cloning and molecular genetic analysis of *Drosophila melanogaster* interband DNA. Mol. Gen. Genet. **238:** 437–443.

de Nooij, J. C., M. A. Letendre and I. K. Hariharan, 1996 A cyclin-dependent kinase inhibitor, dacapo, is necessary for timely exit from the cell cycle during Drosophila embryogenesis. Cell **87:** 1237–1247.

Dingwall, A. K., S. J. Beek, C. M. McCallum, J. W. Tamkun, G. V. Kalpana *et al.*, 1995 The Drosophila snr1 and brm proteins are related to yeast SWI/SNF proteins and are components of a large protein complex. Mol. Biol. Cell **6:** 777–791.

Dolph, P. J., R. Ranganathan, N. J. Colley, R. W. Hardy, M. Socolich *et al.*, 1992 Arrestin function in inactivation of G protein-coupled receptor rhodopsin in vivo. Science **260:** 1910–1916.

Dong, X., K. H. Zavitz, B. J. Thomas, M. Lin, S. Campbell *et al.*, 1997 Control of G1 in the developing Drosophila eye: *rca1* regulates *Cyclin A*. Genes Dev. **11:** 94–105.

Dorn, R., J. Szidonya, G. Korge, M. Sehnert, H. Taubert *et al.*,

1993 Transposon-induced dominant enhancer mutations of position-effect variegation in *Drosophila melanogaster*. Genetics **133:** 279–290.

Duman-Scheel, M., X. Li, I. Orlov, M. Noll and N. H. Patel, 1997 Genetic separation of the neural and cuticular patterning functions of *gooseberry*. Development **124:** 2855–2865.

Duronio, R. J., P. H. O'Farrell, J. E. Xie, A. Brook and N. Dyson, 1995 The transcription factor E2F is required for S phase during Drosophila embryogenesis. Genes Dev. **9:** 1445–1455.

Eberhart, C. G., and S. A. Wasserman, 1995 The *pelota* locus encodes a protein required for meiotic cell division: an analysis of G2/M arrest in Drosophila spermatogenesis. Development **121:** 3477–3486.

Eberhart, C. G., J. Z. Maines and S. A. Wasserman, 1996 Meiotic cell cycle requirement for a fly homologue of human Deleted in Azoospermia. Nature **381:** 783–785.

Edgar, B. A., and P. H. O'Farrell, 1989 Genetic control of cell divisions patterns in the Drosophila embryo. Cell **57:** 177–187.

Eldon, E., S. Kooyer, D. D'Evelyn, M. Dunman, P. Lawinger *et al.*, 1994 The Drosophila *18-wheeler* is required for morphogenesis and has striking similarities to *Toll*. Development **120:** 885–899.

Endo, K., T. Akiyama, S. Kobayashi and M. Okada, 1996 *Degenerative spermatocyte*, a novel gene encoding a transmembrane protein required for the initiation of meiosis in Drosophila spermatogenesis. Mol. Gen. Genet. **253:** 157–165.

Erdelyi, M., A. M. Michon, A. Guichet, J. B. Glotzer and A. Ephrussi, 1995 Requirement for Drosophila cytoplasmic tropomyosin in *oskar* mRNA localization. Nature **377:** 524–527.

Eulenberg, K. G., and R. Schuh, 1997 The *tracheae* defective gene encodes a bZIP protein that controls tracheal cell movement during Drosophila embryogenesis. EMBO J. **16:** 7156–7165.

Feng, Y., L. Huynh, K. Takeyasu and D. M. Fambrough, 1997 The Drosophila Na,K-ATPase alpha-subunit gene: gene structure, promoter function and analysis of a cold-sensitive recessive-lethal mutation. Genes Funct. **1:** 99–117.

Fernandez, R., D. Tabarini, N. Azpiazu and M. Frasch, 1995 The Drosophila insulin receptor homolog: a gene essential for embryonic development encodes two receptor isoforms with different signaling potential. EMBO J. **14:** 3373–3384.

Fletcher, J. C., and C. S. Thummel, 1995 The Drosophila *E74* gene is required for the proper stage- and tissue-specific transcription of ecdysone-regulated genes at the onset of metamorphosis. Development **121:** 1411–1421.

Forbes, A. J., A. C. Spradling, P. W. Ingham and H. Lin, 1996 The role of segment polarity genes during early oogenesis in Drosophila. Development **122:** 3283–3294.

Forjanic, J. P., C. K. Chen, H. Jackle and M. Gonzalez-Gaitan, 1997 Genetic analysis of stomatogastric nervous system development in Drosophila using enhancer trap lines. Dev. Biol. **186:** 139–154.

Freeman, M., C. Klambt, C. S. Goodman and G. M. Rubin, 1992 The *argos* gene encodes a diffusible factor that regulates cell fate decisions in the Drosophila eye. Cell **69:** 963–975.

Frommer, G., G. Vorbruggen, G. Pasca, H. Jaeckle and T. Volk, 1996 Epidermal egr-like zinc finger protein of Drosophila participates in myotube guidance. EMBO J. **15:** 1642–1649.

Gaiano, N., A. Amsterdam, K. Kawakami, M. Allende, T. Becker *et al.*, 1996 Insertional mutagenesis and rapid cloning of essential genes in zebrafish. Nature **383:** 829–832.

Garraway, L. A., L. R. Tosi, Y. Wang, J. B. Moore, D. E. Dobson *et al.*, 1997 Insertional mutagenesis by a modified in vitro Ty1 transposition system. Gene **198:** 27–35.

Garrity, P. A., Y. Rao, I. Salecker, J. McGlade, T. Pawson *et al.*, 1996 Drosophila photoreceptor axon guidance and targeting requires the dreadlocks SH2/SH3 adapter protein. Cell **85:** 639–650.

Gaul, U., G. Mardon and G. M. Rubin, 1992 A putative Ras GTPase activating protein acts as a negative regulator of signaling by the Sevenless receptor tyrosine kinase. Cell **68:** 1007–1019.

Gellon, G., K. W. Harding, N. McGinnis, M. M. Martin and W. McGinnis, 1997 A genetic screen for modifiers of *Deformed* homeotic function identifies novel genes required for head development. Development **124:** 3321–3331.

George, H., and R. Terracol, 1997 The vrille gene of Drosophila is a maternal enhancer of *decapentaplegic* and encodes a new member of the bZIP family of transcription factors. Genetics **146:** 1345–1363.

Gigliotti, S., G. Callaini, S. Andone, M. G. Riparbelli, R. Perrnas-Alonso *et al.*, 1998 *Nup154*, a new Drosophila gene essential for male and female gametogenesis is related to the *nup155* vertebrate nucleoporin gene. J. Cell Biol. **142:** 1195–1207.

Gillespie, D. E., and C. A. Berg, 1995 *Homeless* is required for RNA localization in Drosophila oogenesis and encodes a new member of the DE-H family of RNA-dependent ATPases. Genes Dev. **9:** 2495–2508.

Gonzalez-Gaitan, M., and H. Jackle, 1997 Role of Drosophila alpha-adaptin in presynaptic vesicle recycling. Cell **88:** 767–776.

Gossler, A., A. L. Joyner, J. Rossant and W. C. Skarnes, 1989 Mouse embryonic stem cells and reported constructs to detect developmentally regulated genes. Science **244:** 463–465.

Goto, S., and S. Hayashi, 1997 Specification of the embryonic limb primordium by graded activity of Decapentaplegic. Development **124:** 125–132.

Grether, M. E., J. M. Abrams, J. Agapite, K. White and H. Steller, 1995 The *head involution defective* gene of *Drosophila melanogaster* functions in programmed cell death. Genes Dev. **9:** 1694–1708.

Guillemin, K., J. Groppe, K. Ducker, R. Treisman, E. Hafen *et al.*, 1996 The *pruned* gene encodes the Drosophila serum response factor and regulates cytoplasmic outgrowth during terminal branching of the tracheal system. Development **122:** 1353–1362.

Hacker, U., X. Lin and N. Perrimon, 1997 The Drosophila *sugarless* gene modulates Wingless signaling and encodes an enzyme involved in polysaccharide biosynthesis. Development **124:** 3565–3573.

Hahn, M., and H. Jackle, 1996 Drosophila *goosecoid* participates in neural development but not in body axis formation. EMBO J. **15:** 3077–3084.

Hamilton, B. A., M. J. Palazzolo, J. H. Chang, K. VijayRaghavan, C. A. Mayeda *et al.*, 1991 Large scale screen for transposon insertions into cloned genes. Proc. Natl. Acad. Sci. USA **88:** 2731–2735.

Harvie, P. D., M. Filippova and P. J. Bryant, 1998 Genes expressed in the ring gland, the major endocrine organ of *Drosophila melanogaster*. Genetics **149:** 217–231.

Hassan, B. A., S. N. Prokopenko, S. Breuer, B. Zhang, A. Paululat *et al.*, 1998 *skittles*, a Drosophila phosphatidylinositol 4-phosphate 5-kinase, is required for cell viability, germline development and bristle morphology, but not for neurotransmitter release. Genetics **150:** 1527–1537.

Hay, B. A., D. A. Wassarman and G. M. Rubin, 1995 Drosophila homologs of baculovirus inhibitor of apoptosis proteins function to block cell death. Cell **83:** 1253–1262.

Heberlein, U., T. Wolff and G. M. Rubin, 1993 The TGFbeta homolog *dpp* and the segment polarity gene *hedgehog* are required for propagation of a morphogenetic wave in the Drosophila retina. Cell **75:** 913–926.

Heck, M. M. S., A. Pereira, P. Pesavento, Y. Yannoni, A. C. Spradling *et al.*, 1993 The kinesin-like protein KLP61F is essential for mitosis in Drosophila. J. Cell Biol. **123:** 665–679.

Hong, C. C., and C. Hashimoto, 1995 An unusual mosaic protein with a protease domain, encoded by the *nudel* gene, is involved in defining embryonic dorsoventral polarity in Drosophila. Cell **82:** 785–794.

Horowitz, H., and C. A. Berg, 1996 The Drosophila *pipsqueak* gene encodes a nuclear BTB-domain-containing protein required early in oogenesis. Development **122:** 1859–1871.

Hoshizaki, D. K., 1994 *Kruppel* expression during postembryonic development of Drosophila. Dev. Biol. **163:** 133–140.

Hou, X. S., M. B. Melnick and N. Perrimon, 1996 *marelle* acts downstream of the *Drosophila* HOP/JAK kinase and encodes a protein similar to the mammalian STATs. Cell **84:** 411–419.

Iliopoulos, I., I. Torok and B. M. Mechler, 1997 The *DnaJ60* gene of *Drosophila melanogaster* encodes a new member of the DnaJ family of proteins. Biol. Chem. Hoppe-Seyler **378:** 1177–1181.

Isaac, D. D., and D. J. Andrew, 1996 Tubulogenesis in Drosophila: a requirement for the *trachealess* gene product. Genes Dev. **10:** 103–117.

Ito, H., K. Fujitani, K. Usui, K. Shimizu-Nishikawa, S. Tanaka *et*

al., 1996 Sexual orientation in Drosophila is altered by the satori mutation in the sex-determination gene *fruitless* that encodes a zinc finger protein with a BTB domain. Proc. Natl. Acad. Sci. USA **93:** 9687–9692.

Jaenisch, R., 1988 Transgenic animals. Science **240:** 1468–1474.

Jiang, J., and G. Struhl, 1998 Regulation of Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein slimb. Nature **391:** 493–496.

Justice, R. W., O. Zilian, D. F. Woods, M. Noll and P. J. Bryant, 1995 The Drosophila tumor suppressor gene *warts* encodes a homolog of human myotonic dystrophy kinase and is required for the control of cell shape and proliferation. Genes Dev. **9:** 534–546.

Kania, A., A. Salzberg, M. Bhat, D. D'Evelyn, Y. He *et al.*, 1995 P-element mutations affecting embryonic peripheral nervous system development in *Drosophila melanogaster*. Genetics **139:** 1663–1678.

Karpen, G. H., and A. C. Spradling, 1992 Analysis of subtelomeric heterochromatin in a Drosophila minichromosome by single *P* element insertional mutagenesis. Genetics **132:** 737–753.

Karsch-Mizrachi, I., and S. R. Haynes, 1993 The Rb97D gene encodes a potential RNA-binding protein required for spermatogenesis in Drosophila. Nucleic Acids Res. **21:** 2229–2235.

Kauffmann, R. C., S. Li, P. A. Gallagher, J. Zhang and R. W. Carthew, 1996 Ras1 signaling and transcriptional competence in the R7 cell of Drosophila. Genes Dev. **10:** 2167–2178.

Kelley, R. L., 1993 Initial organization of the Drosophila dorsoventral axis depends on an RNA-binding protein encoded by the *squid* gene. Genes Dev. **7:** 948–960.

Keyes, L. N., and A. C. Spradling, 1997 The Drosophila gene *fs(2)cup* interacts with *otu* to define a cytoplasmic pathway required for the structure and function of germ-line chromosomes. Development **124:** 1419–1431.

Kidwell, M., 1986 P-M mutagenesis, pp. 59–81 in *Drosophila: A Practical Approach*, edited by D. B. Roberts. IRL, Oxford.

Kleckner, N., J. Roth and D. Botstein, 1977 Genetic engineering in vivo using translocatable drug-resistance elements. New methods in bacterial genetics. J. Mol. Biol. **116:** 125–159.

Kloss, B., J. L. Price, L. Saez, J. Blau, A. Rothenfluh *et al.*, 1998 The Drosophila clock gene double-time encodes a protein closely related to human casein kinase Iepsilon. Cell **94:** 97–107.

Knoblich, J. A., K. Sauer, L. Jones, H. Richardson, R. Saint *et al.*, 1994 Cyclin E controls S phase progression and its down-regulation during Drosophila embryogenesis is required for the arrest of cell proliferation. Cell **77:** 107–120.

Kockel, L., G. Vorbruggen, H. Jackle, M. Mlodzik and D. Bohmann, 1997 Requirement for Drosophila *14-3-3 zeta* in Raf-dependent photoreceptor development. Genes Dev. **11:** 1140–1147.

Kolhekar, A. S., M. S. Roberts, N. Jiang, R. C. Johnson, R. E. Mains *et al.*, 1997 Neuropeptide amidation in Drosophila: separate genes encode the two enzymes catalyzing amidation. J. Neurosci. **17:** 1363–1376.

Kolodkin, A. L., D. J. Matthes and C. S. Goodman, 1993 The *semaphorin* genes encode a family of transmembrane and secreted growth cone guidance molecules. Cell **75:** 1389–1399.

Korswagen, H. C., R. M. Durbin, M. T. Smits and R. H. Plasterk, 1996 Transposon Tc1-derived, sequence-tagged sites in *Caenorhabditis elegans* as markers for gene mapping. Proc. Natl. Acad. Sci. USA **93:** 14680–14685.

Kozlova, T., G. V. Pokholkova, G. Tzertzinis, J. D. Sutherland, I. F. Zhimulev *et al.*, 1998 Drosophila *hormone receptor 38* functions in metamorphosis: a role in adult cuticle formation. Genetics **149:** 1465–1475.

Kussel, P., and M. Frasch, 1995 Pendulin, a Drosophila protein with cell cycle-dependent nuclear localization, is required for normal cell proliferation. J. Cell Biol. **129:** 1491–1507.

Lantz, V., L. Ambrosio and P. Schedl, 1992 The Drosophila *orb* gene is predicted to encode sex-specific germline RNA-binding proteins and has localized transcripts in ovaries and early embryos. Development **115:** 75–88.

Lee, J. J., D. Von Kessler, S. Parks and P. A. Beachy, 1992 Secretion and localized transcription suggest a role in positional signaling for products of the segmentation gene *hedgehog*. Cell **71:** 33–50.

Lehner, C. F., and P. H. O'Farrell, 1989 Expression and function of Drosophila cyclin A during embryonic cell cycle progression. Cell **56:** 957–968.

Lekven, A. C., U. Tepass, M. Keshmeshian and V. Hartenstein, 1998 *faint sausage* encodes a novel extracellular protein of the immunoglobulin superfamily required for cell migration and the establishment of normal axonal pathways in the Drosophila nervous system. Development **125:** 2747–2758.

Lepage, T., S. M. Cohen, F. J. Diaz-Benjumea and S. M. Parkhurst, 1995 Signal transduction by cAMP-dependent protein kinase A in Drosophila limb patterning. Nature **373:** 711–715.

Letsou, A., K. Arora, J. L. Wrana, K. Simin, V. Twombly *et al.*, 1995 Drosophila Dpp signaling is mediated by the *punt* gene product: a dual ligand-binding type II receptor of the TGF beta receptor family. Cell **80:** 899–908.

Levine, A., A. Bashan-Ahrend, O. Budai-Hadrian, D. Gartenberg, S. Menasherow *et al.*, 1994 *Odd Oz*: a novel Drosophila pair rule gene. Cell **77:** 587–598.

Lin, H., and A. C. Spradling, 1993 Germline stem cell division and egg chamber development in transplanted Drosophila germaria. Dev. Biol. **159:** 140–152.

Lin, H., and A. C. Spradling, 1997 A novel group of *pumilio* mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. Development **124:** 2463–2476.

Lin, W. H., L. H. Huang, J. Y. Yeh, J. Hoheisel, H. Lehrach *et al.*, 1995 Expression of a Drosophila GATA transcription factor in multiple tissues in the developing embryos: identification of homozygous lethal mutants with P-element insertion at the promoter region. J. Biol. Chem. **270:** 25150–25158.

Lo, P. C. H., and M. Frasch, 1998 *bagpipe*-dependent expression of *vimar*, a novel armadillo-repeats gene, in Drosophila visceral mesoderm. Mech. Dev. **72:** 65–75.

Mancebo, R., P. C. Lo and S. M. Mount, 1990 Structure and expression of the *Drosophila melanogaster* gene for the U1 small nuclear ribonucleoprotein particle 70K protein. Mol. Cell. Biol. **10:** 2492–2502.

Mardon, G., N. M. Solomon and G. M. Rubin, 1994 *dachshund* encodes a nuclear protein required for normal eye and leg development in Drosophila. Development **120:** 3473–3486.

Martin-Blanco, E., and A. A. Garcia-Bellido, 1996 Mutations in the *rotated abdomen* locus affect muscle development and reveal an intrinsic asymmetry in Drosophila. Proc. Natl. Acad. Sci. USA **93:** 6048–6052.

Mathies, L. D., S. Kerridge and M. P. Scott, 1994 Role of the *teashirt* gene in Drosophila midgut morphogenesis: secreted proteins mediate the action of homeotic genes. Development **120:** 2799–2809.

Meinke, D. W., J. M. Cherry, C. Dean, S. D. Rounsley and M. Koornneef, 1998 *Arabidopsis thaliana*: a model plant for genome analysis. Science **282:** 662–682.

Mellerick, D. M., J. A. Kassis, S. D. Zhang and W. F. Odenwald, 1992 *Castor* encodes a novel zinc finger protein required for the development of a subset of CNS neurons in Drosophila. Neuron **9:** 789–803.

McKearin, D., and A. C. Spradling, 1990 *Bag-of-marbles*: a Drosophila gene required to initiate both male and female gametogenesis. Genes Dev. **4:** 2242–2251.

McKim, K. S., and A. Hayashi-Hagihara, 1998 *mei-W68* in *Drosophila melanogaster* encodes a Spo11 homolog: evidence that the mechanism for initiating meiotic recombination is conserved. Genes Dev. **12:** 2932–2942.

Miklos, G. L., and G. M. Rubin, 1996 The role of the genome project in determining gene function: insights from model organisms. Cell **86:** 521–529.

Mlodzik, M., Y. Hiromi, U. Weber, C. S. Goodman and G. M. Rubin, 1990 The Drosophila *seven-up* gene, a member of the steroid receptor gene superfamily, controls photoreceptor cell fates. Cell **60:** 211–224.

Mohler, J., J. W. Mahaffey, E. Deutsch and K. Vani, 1995 Control of Drosophila head segment identity by the bZIP homeotic gene cnc. Development **121:** 237–247.

Montell, D. J., P. Rørth and A. C. Spradling, 1992 *slow border cells*, a locus required for a developmentally regulated cell migration during oogenesis, encodes Drosophila C/EBP. Cell **71:** 51–62.

Murphy, A. M., T. Lee, C. M. Andrews, B. Z. Shilo and D. J. Montell,

1995 The breathless FGF receptor homolog, a downstream target of Drosophila C/EBP in the developmental control of cell migration. Development **121:** 2255–2263.

Nakato, H., T. A. Futch and S. B. Selleck, 1995 The *division abnormally delayed, dally*, gene: a putative integral membrane proteoglycan required for cell division patterning during postembryonic development of the nervous system in Drosophila. Development **121:** 3687–3702.

Neufeld, T. P., and G. M. Rubin, 1994 The Drosophila *peanut* gene is required for cytokinesis and encodes a protein similar to yeast putative bud neck filament proteins. Cell **77:** 371–379.

Neufeld, T. P., A. H. Tang and G. M. Rubin, 1998 A genetic screen to identify components of the *sina* signaling pathway in Drosophila eye development. Genetics **148:** 277–286.

Nibu, Y., H. Zhang and M. Levine, 1998 Interaction of short-range repressors with Drosophila CtBP in the embryo. Science **280:** 101–104.

Nozaki, M., Y. Onishi, S. Togashi and H. Miyamoto, 1996 Molecular characterization of the Drosophila *Mo25* gene, which is conserved among Drosophila, mouse, and yeast. DNA Cell Biol. **15:** 505–509.

Okano, H., S. Hayashi, T. Tanimura, K. Sawamoto, S. Yoshikawa *et al.*, 1992 Regulation of Drosophila neural development by a putative secreted protein. Differentiation **52:** 1–11.

O'Neill, E. M., I. Rebay, R. Tjian and G. M. Rubin, 1994 The activities of two Ets-related transcription factors required for Drosophila eye development are modulated by the Ras/MAPK pathway. Cell **78:** 137–147.

Park, M., X. Wu, K. Golden, J. D. Axelrod and R. Bodmer, 1996 The *wingless* signaling pathway is directly involved in Drosophila heart development. Dev. Biol. **177:** 104–116.

Perrimon, N., A. Lanjuin, C. Arnold and E. Noll, 1996 Zygotic lethal mutations with maternal effect phenotypes in *Drosophila melanogaster*. Part II. Loci on the second and third chromosomes identified by P-element-induced mutations. Genetics **144:** 1681–1692.

Petersen, S. A., R. D. Fetter, J. N. Noordermeer, C. S. Goodman and A. DiAntonio, 1997 Genetic analysis of glutamate receptors in Drosophila reveals a retrograde signal regulating presynaptic transmitter release. Neuron **19:** 1237–1248.

Plasterk, R. H., 1993 Reverse genetics of *Caenorhabditis elegans*. Bioessays **14:** 629–633.

Rauskolb, C., K. M. Smith, M. Peifer and E. Wieschaus, 1995 Extradenticle determines segmental identities throughout Drosophila development. Development **121:** 3663–3673.

Rehorn, K. P., H. Thelen, A. M. Michelson and R. Reuter, 1996 A molecular aspect of hematopoiesis and endoderm development common to vertebrates and Drosophila. Development **122:** 4023–4031.

Rieckhof, G. E., F. Casares, H. D. Ryoo, M. Abu-Shaar and R. S. Mann, 1997 Nuclear translocation of extradenticle requires homothorax, which encodes an extradenticle-related homeodomain protein. Cell **91:** 171–183.

Rittenhouse, K. R., and C. A. Berg, 1995 Mutations in the Drosophila gene *bullwinkle* cause the formation of abnormal eggshell structures and bicaudal embryos. Development **121:** 3023–3033.

Roch, F., F. Serras, F. J. Cifuentes, M. Corominas, B. Alsina *et al.*, 1998 Screening of larval/pupal P-element induced lethals on the second chromosome in *Drosophila melanogaster*: clonal analysis and morphology of imaginal discs. Mol. Gen. Genet. **257:** 103–112.

Rodriguez, A., Z. Zhou, M. L. Tang, S. Meller, J. Chen *et al.*, 1996 Identification of immune system and response genes, and novel mutations causing melanotic tumor formation in *Drosophila melanogaster*. Genetics **143:** 929–940.

Rooke, J., D. Pan, T. Xu and G. M. Rubin, 1996 KUZ, a conserved metalloprotease-disintegrin protein with two roles in Drosophila neurogenesis. Science **273:** 1227–1231.

Rørth, P., 1996 A modular misexpression screen in Drosophila detecting tissue-specific phenotypes. Proc. Natl. Acad. Sci. USA **93:** 12418–12422.

Rørth, P., K. Szabo, A. Bailey, T. Laverty, J. Rehm *et al.*, 1998 Systematic gain-of-function genetics in Drosophila. Development **125:** 1049–1057.

Rottgen, G., T. Wagner and G. Hinz, 1998 A genetic screen for elements of the network that regulates neurogenesis in Drosophila. Mol. Gen. Genet. **257:** 442–451.

Roulier, E. M., S. Panzer and S. K. Beckendorf, 1998 The Tec29 tyrosine kinase is required during Drosophila embryogenesis and interacts with Src64 in ring canal development. Mol. Cell **6:** 819–829.

Ruberte, E., T. Marty, D. Nellen, M. Affolter and K. Basler, 1995 An absolute requirement for both the Type II and Type I receptors, punt and thick veins, for Dpp signaling in vivo. Cell **80:** 889–897.

Ruden, D. M., W. Cui, V. Sollars and M. A. Alterman, 1997 Drosophila kinesin-like protein (Klp38B) functions during meiosis, mitosis, and segmentation. Dev. Biol. **191:** 284–296.

Rudner, D. Z., R. Kanaar, K. S. Breger and D. C. Rio, 1996 Mutations in the small subunit of the Drosophila U2AF splicing factor cause lethality and developmental defects. Proc. Natl. Acad. Sci. USA **93:** 10333–10337.

Russell, S. R. H., G. Heimbeck, C. M. Goddard, A. T. C. Carpenter and M. Ashburner, 1996 The Drosophila *Eip78C* gene is not vital but has a role in regulating chromosome puffs. Genetics **144:** 159–170.

Russell, M. A., L. Ostafichuk and S. Scanga, 1998 Lethal P-lacZ insertion lines expressed during pattern respecification in the imaginal discs of Drosophila. Genome **41:** 7–13.

Salzberg, A., K. Golden, R. Bodmer and H. J. Bellen, 1996 *gutfeeling*, a Drosophila gene encoding an antizyme-like protein, is required for late differentiation of neurons and muscles. Genetics **144:** 183–196.

Samakovlis, C., G. Manning, P. Steneberg, N. Hacohen, R. Cantera *et al.*, 1996 Genetic control of epithelial tube fusion during Drosophila tracheal development. Development **122:** 3531–3536.

Savant-Bhonsale, S., and D. J. Montell, 1993 *Torso-like* encodes the localized determinant of Drosophila terminal pattern formation. Genes Dev. **7:** 2548–2555.

Schmucker, D., H. Jackle and U. Gaul, 1997 Genetic analysis of the larval optic nerve projection in Drosophila. Development **124:** 937–948.

Schneider, L. E., and A. C. Spradling, 1997 The Drosophila G-protein-coupled receptor kinase homologue Gprk2 is required for egg morphogenesis. Development **124:** 2591–2602.

Schnorr, J. D., and C. A. Berg, 1996 Differential activity of Ras1 during patterning of the Drosophila dorsoventral axis. Genetics **144:** 1545–1557.

Schulze, K. L., K. Broadie, M. S. Perin and H. J. Bellen, 1995 Genetic and electrophysiological studies of Drosophila Syntaxin-1A demonstrate its role in nonneuronal secretion and neurotransmission. Cell **80:** 311–320.

Siegel, V., T. A. Jongens, L. Y. Jan and Y. N. Jan, 1993 *pipsqueak*, an early acting member of the posterior group of genes, affects vasa level and germ cell-somatic cell interaction in the developing egg chamber. Development **119:** 1187–1202.

Simon, M. A., G. S. Dodson and G. M. Rubin, 1993 An SH3-SH2-SH3 protein is required for p21Ras1 activation and binds to sevenless and Sos proteins in vitro. Cell **73:** 169–177.

Sisson, J. C., K. S. Ho, K. Suyama and M. P. Scott, 1997 Costal2, a novel kinesin-related protein in the Hedgehog signaling pathway. Cell **90:** 235–245.

Sliter, T. J., V. C. Henrich, R. L. Tucker and L. I. Gilbert, 1989 The genetics of the *Dras3-Roughened-ecdysoneless* chromosomal region (62B3-4 to 62D3-4) in *Drosophila melanogaster*: analysis of recessive lethal mutations. Genetics **123:** 327–336.

Smith, D., Y. Yanai, Y. G. Lie, S. Ishiguro, K. Okada *et al.*, 1996 Characterization and mapping of Ds-GUS-T-DNA lines for targeted insertional mutagenesis. Plant J. **10:** 721–732.

Song, Z., K. McCall and H. Steller, 1997 DCP-1, a Drosophila cell death protease essential for development. Science **275:** 536–540.

Sotillos, S., F. Roch and S. Campuzano, 1997 The metalloprotease-disintegrin Kuzbanian participates in Notch activation during growth and patterning of Drosophila imaginal discs. Development **124:** 4769–4779.

Sozen, M. A., J. D. Armstrong, M. Yang, K. Kaiser and J. A. T. Dow, 1997 Functional domains are specified to single-cell resolution

in a Drosophila epithelium. Proc. Natl. Acad. Sci. USA **94:** 5207–5212.

Spradling, A. C., D. M. Stern, I. Kiss, J. Foote, T. Laverty *et al.*, 1995 Gene disruptions using P transposable elements: an integral component of the Drosophila genome project. Proc. Natl. Acad. Sci. USA **92:** 10824–10830.

Stark, K. A., G. H. Yee, C. E. Roote, E. L. Williams, S. Zusman *et al.*, 1997 A novel alpha integrin subunit associates with betaPS and functions in tissue morphogenesis and movement during Drosophila development. Development **124:** 4583–4594.

Steller, H., and V. Pirrotta, 1986 P transposons controlled by the heat shock promoter. Mol. Cell. Biol. **6:** 1640–1649.

Stroumbakis, N. D., Z. Li and P. P. Tolias, 1996 A homolog of human transcription factor NF-X1 encoded by the Drosophila *shuttle craft* gene is required in the embryonic central nervous system. Mol. Cell. Biol. **16:** 192–201.

Strutt, D. I., U. Weber and M. Mlodzik, 1997 The role of RhoA in tissue polarity and Frizzled signalling. Nature **387:** 292–295.

Sullivan, W., P. Fogarty and W. E. Theurkauf, 1993 Mutations affecting the cytoskeletal organization of syncytial Drosophila embryos. Development **118:** 1245–1254.

Sundaresan, V., P. Springer, T. Volpe, S. Haward, J. D. G. Jones, C. Dean, H. Ma and R. Martienssen, 1995 Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. Genes Dev. **9:** 1797–1810.

Taylor, C. A., K. N. Stanley and A. D. Shirras, 1997 The Orct gene of Drosophila melanogaster codes for a putative organic cation transporter with six or 12 transmembrane domains. Gene **201:** 69–74.

Tepass, U., E. Gruszynski-Defeo, T. A. Haag, L. Omatyar, T. Torok *et al.*, 1996 *shotgun* encodes Drosophila E-cadherin and is preferentially required during cell rearrangement in the neurectoderm and other morphogenetically active epithelia. Genes Dev. **10:** 672–685.

Tetzlaff, M. T., H. Jaeckle and M. J. Pankratz, 1996 Lack of Drosophila cytoskeletal tropomyosin effects head morphogenesis and the accumulation of oskar mRNA required to germ cell formation. EMBO J. **15:** 1247–1254.

Therrien, M., H. C. Y. Chang, N. M. Solomon, F. D. Karim, D. A. Wassarman *et al.*, 1995 KSR, a novel protein kinase required for RAS signal transduction. Cell **83:** 879–888.

Therrien, M., A. M. Wong and G. M. Rubin, 1998 CNK, a FAF-binding multidomain protein required for RAS signaling. Cell **95:** 343–353.

Török, T., T. Tick, M. Alvarado and I. Kiss, 1993 *P-lacW* insertional mutagenesis on the second chromosome of *Drosophila melanogaster*: isolation of lethals with different overgrowth phenotypes. Genetics **135:** 71–80.

Török, T., P. D. Harvie, M. Buratovich and P. J. Bryant, 1997 The product of *proliferation disrupter* is concentrated at centromeres and required for mitotic chromosome condensation and cell proliferation in Drosophila. Genes Dev. **11:** 213–225.

Treisman, J. E., and G. M. Rubin, 1996 Targets of glass regulation in the Drosophila eye disc. Mech. Dev. **56:** 17–24.

Treisman, J. E., P. J. Follette, P. H. O'Farrell and G. M. Rubin, 1995a Cell proliferation and DNA replication defects in a Drosophila MCM2 mutant. Genes Dev. **9:** 1709–1715.

Treisman, J. E., Z. C. Lai and G. M. Rubin, 1995b Shortsighted acts in the *decapentaplegic* pathway in Drosophila eye development and has homology to a mouse TGF-beta-responsive gene. Development **121:** 2835–2845.

Treisman, J. E., N. Ito and G. M. Rubin, 1997a *misshapen* encodes a protein kinase involved in cell shape control in Drosophila. Gene **186:** 119–125.

Treisman, J. E., A. Luk, G. M. Rubin and U. Heberlein, 1997b *eyelid* antagonizes *wingless* signaling during Drosophila development and has homology to the Bright family of DNA-binding proteins. Genes Dev. **11:** 1949–1962.

Tsuneizumi, K., T. Nakayama, Y. Kamoshida, T. B. Kornberg, J. L. Christian *et al.*, 1997 *Daughters against dpp* modulates *dpp* organizing activity in Drosophila wing development. Nature **389:** 627–631.

Twombly, V., R. K. Blackman, H. Jin, J. M. Graff, R. W. Padgett *et al.*, 1996 The TGF-beta signaling pathway is essential for Drosophila oogenesis. Development **122:** 1555–1565.

Uemura, T., S. Shepherd, L. Ackerman, L. Y. Jan and Y. N. Jan, 1989 *numb*, a gene required in determination of cell fate during sensory organ formation in Drosophila embryos. Cell **58:** 349–360.

Uemura, T., K. Shiomi, S. Togashi and M. Takeichi, 1993 Mutation of *twins* encoding a regulator of protein Phosphatase 2A leads to pattern duplication in Drosophila imaginal discs. Genes Dev. **7:** 429–440.

Uemura, T., H. Oda, R. Kraut, S. Hayashi, Y. Kotaoka and M. Takeichi, 1996 Zygotic Drosophila E-cadherin expression is required for processes of dynamic epithelial cell rearrangements in the Drosophila embryo. Genes Dev. **10:** 659–671.

van der Straten, A., C. Rommel, B. Dickson and E. Hafen, 1997 The heat shock protein 83 (Hsp83) is required for Raf-mediated signalling in Drosophila. EMBO J. **16:** 1961–1969.

Vankatesh, K., and G. Hasan, 1997 Disruption of the IP3 receptor gene of Drosophila affects larval metamorphosis and ecdysone release. Curr. Biol. **7:** 500–509.

Venter, J. C., M. D. Adams, G. G. Sutton, A. R. Kerlavage, H. O. Smith *et al.*, 1998 Shotgun sequencing of the human genome. Science **280:** 1540–1542.

Verheyen, E. M., K. J. Purcell, M. E. Fortini and S. Artavanis-Tsakonas, 1996 Analysis of dominant enhancers and suppressors of activated Notch in Drosophila. Genetics **144:** 1127–1141.

Wassarman, D. A., N. M. Solomon, H. C. Y. Chang, F. D. Karim, M. Therrien *et al.*, 1996 Protein phosphatase 2A positively and negatively regulates Ras1-mediated photoreceptor development in Drosophila. Genes Dev. **10:** 272–278.

Wilk, R., I. Weizman and B. Z. Shilo, 1996 *trachealess* encodes a bHLH-PAS protein that is an inducer of tracheal cell fates in Drosophila. Genes Dev. **10:** 93–102.

Wurst, W., J. Rossant, V. Prideaux, M. Kownacka, A. Joyner *et al.*, 1995 A large-scale gene-trap screen for insertional mutations in developmentally regulated genes in mice. Genetics **139:** 889–899.

Xiong, W. C., and C. Montell, 1993 *tramtrack* is a transcriptional repressor required for cell fate determination in the Drosophila eye. Genes Dev. **7:** 1085–1096.

Xiong, W. C., H. Okano, N. H. Patel, J. A. Blendy and C. Montell, 1994 *Repo* encodes a glial-specific homeo domain protein required in the Drosophila nervous system. Genes Dev. **8:** 981–994.

Xue, F., and L. Cooley, 1993 *kelch* encodes a component of intercellular bridges in Drosophila egg chambers. Cell **72:** 681–693.

Yang, X., S. Bahri, T. Klein and W. Chia, 1997 Klumpfuss, a putative Drosophila zinc finger transcription factor, acts to differentiate between the identities of two secondary precursor cells within one neuroblast lineage. Genes Dev. **11:** 1396–1408.

Yarnitzky, T., L. Min and T. Volk, 1997 The Drosophila neuregulin homolog vein mediates inductive interactions between myotubes and their epidermal attachment cells. Genes Dev. **11:** 2691–2700.

Yasothornsrikul, S., W. J. Davis, G. Cramer, D. A. Kimbrell and C. R. Dearolf, 1997 viking: dentification and characterization of a second type IV collagen in Drosophila. Gene **198:** 17–25.

Yu, Y., W. Li, K. Su, M. Yussa, W. Han *et al.*, 1997 The nuclear hormone receptor Ftz-F1 is a cofactor for the Drosophila homeodomain protein Ftz. Nature **385:** 552–555.

Yu, H. H., H. H. Araj, S A. Ralls and A. L. Kolodkin, 1998 The transmembrane Semaphorin Sema I is required in Drosophila for embryonic motor and CNS axon guidance. Neuron **20:** 207–220.

Yue, L., and A. C. Spradling, 1992 *hu-li tai shao*, a gene required for ring canal formation during Drosophila oogenesis, encodes a homolog of adducin. Genes Dev. **6:** 2443–2454.

Zambrowicz, B. P., G. A. Friedrich, E. C. Buxton, S. L. Lilleberg, C. Person *et al.*, 1998 Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. Nature **392:** 608–611.

Zhang, P., and A. C. Spradling, 1995 The Drosophila salivary gland chromocenter contains highly polytenized subdomains of mitotic heterochromatin. Genetics **139:** 659–670.

Zhang, N., J. Zhang, K. J. Purcell, Y. Cheng and K. Howard, 1997 The Drosophila protein Wunen repels migrating germ cells. Nature **385:** 64–67.

Zhang, B., Y. H. Koh, R. B. Beckstead, V. Budnik, B. Ganetzky *et al.*, 1998 Synaptic vesicle size and number are regulated by a clathrin adaptor protein required for endocytosis. Neuron **21:** 1465–1475.

Zhou, Q., and D. S. Haymer, 1997 Molecular structure of yoyo, a gypsy-like retrotransposon from the mediterranean fruit fly, *Ceratitis capitata.* Genetica **101:** 167–178.

zur Lage, P., A. D. Shrimpton, A. J. Flavell, T. F. Mackay and A. J. Brown, 1997 Genetic and molecular analysis of *smooth*, a quantitative trait locus affecting bristle number in *Drosophila melanogaster.* Genetics **146:** 607–618.

Communicating editor: R. S. Hawley