

The ribosomal protein genes and *Minute* loci of *Drosophila melanogaster*

Steven J Marygold^{*}, John Roote[†], Gunter Reuter[‡], Andrew Lambertsson[§], Michael Ashburner[†], Gillian H Millburn[†], Paul M Harrison[¶], Zhan Yu[¶], Naoya Kenmochi[¥], Thomas C Kaufman[#], Sally J Leever^{*} and Kevin R Cook[#]

Addresses: ^{*}Growth Regulation Laboratory, Cancer Research UK London Research Institute, Lincoln's Inn Fields, London WC2A 3PX, UK. [†]Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK. [‡]Institute of Genetics, Biologikum, Martin Luther University Halle-Wittenberg, Weinbergweg, Halle D-06108, Germany. [§]Institute of Molecular Biosciences, University of Oslo, Blindern, Oslo N-0316, Norway. [¶]Department of Biology, McGill University, Dr Penfield Ave, Montreal, Quebec H3A 1B1, Canada. [¥]Frontier Science Research Center, University of Miyazaki, 5200 Kihara, Kiyotake, Miyazaki 889-1692, Japan. [#]Department of Biology, Indiana University, E. Third Street, Bloomington, IN 47405-7005, USA.

Correspondence: Steven J Marygold. Email: s.marygold@gen.cam.ac.uk. Kevin R Cook. Email: kcook@bio.indiana.edu

Published: 10 October 2007

Genome Biology 2007, **8**:R216 (doi:10.1186/gb-2007-8-10-r216)

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2007/8/10/R216>

Received: 17 June 2007

Revised: 10 October 2007

Accepted: 10 October 2007

© 2007 Marygold et al.; licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Mutations in genes encoding ribosomal proteins (RPs) have been shown to cause an array of cellular and developmental defects in a variety of organisms. In *Drosophila melanogaster*, disruption of RP genes can result in the 'Minute' syndrome of dominant, haploinsufficient phenotypes, which include prolonged development, short and thin bristles, and poor fertility and viability. While more than 50 *Minute* loci have been defined genetically, only 15 have so far been characterized molecularly and shown to correspond to RP genes.

Results: We combined bioinformatic and genetic approaches to conduct a systematic analysis of the relationship between RP genes and *Minute* loci. First, we identified 88 genes encoding 79 different cytoplasmic RPs (CRPs) and 75 genes encoding distinct mitochondrial RPs (MRPs). Interestingly, nine CRP genes are present as duplicates and, while all appear to be functional, one member of each gene pair has relatively limited expression. Next, we defined 65 discrete *Minute* loci by genetic criteria. Of these, 64 correspond to, or very likely correspond to, CRP genes; the single non-CRP-encoding *Minute* gene encodes a translation initiation factor subunit. Significantly, MRP genes and more than 20 CRP genes do not correspond to *Minute* loci.

Conclusion: This work answers a longstanding question about the molecular nature of *Minute* loci and suggests that Minute phenotypes arise from suboptimal protein synthesis resulting from reduced levels of cytoribosomes. Furthermore, by identifying the majority of haplolethal and haplosterile loci at the molecular level, our data will directly benefit efforts to attain complete deletion coverage of the *D. melanogaster* genome.

Background

Ribosomes are sophisticated macromolecular machines that catalyze cellular protein synthesis in all cells of all organisms. They have an ancient evolutionary origin and are essential for cell growth, proliferation and viability. Though larger and more complex in higher organisms, both the structure and function of ribosomes have been conserved throughout evolution. Genetic approaches in *Drosophila melanogaster* have shown that disrupting ribosome function can result in an array of fascinating dominant phenotypes [1,2]. Despite this, there has so far been no comprehensive inventory of genes encoding ribosome components in this organism, nor any systematic effort to determine their mutant phenotypes.

All ribosomes comprise a set of ribosomal proteins (RPs) surrounding a catalytic core of ribosomal RNA (rRNA). Bacteria possess a single type of ribosome composed of three rRNA molecules and typically 54 RPs. All eukaryotic cells, in contrast, contain at least two distinct types of ribosomes: cytoplasmic ribosomes (cytoribosomes) and mitochondrial ribosomes (mitoribosomes). Cytoribosomes are found on the endoplasmic reticulum and in the aqueous cytoplasm. They translate all mRNAs produced from nuclear genes and perform the vast majority of cellular protein synthesis. Each cytoribosome contains four different rRNAs and 78-80 cytoplasmic RPs (CRPs). Mitoribosomes consist of only two rRNA molecules and up to 80 mitochondrial RPs (MRPs). They are located in the mitochondrial matrix and synthesize proteins involved in oxidative phosphorylation encoded by those few genes retained in the mitochondrial genome. A third unique type of eukaryotic ribosome is found within the plastids (for example, chloroplasts) of plant and various algal cells. In all cases, distinct small and large ribosomal subunits exist that join together during the translation initiation process to form mature ribosomes capable of protein synthesis. (See references [3-6] for general reviews of ribosomal structure and function.)

The protein components of ribosomes are interesting from several points of view. First, and most obviously, RPs play critical roles in ribosome assembly and function [7]. Second, several RPs perform important extra-ribosomal functions, including roles in DNA repair, transcriptional regulation and apoptosis [6,8]. Third, misexpression of human CRP and MRP genes has been implicated in a wide spectrum of human syndromes and diseases, including Diamond-Blackfan anaemia [9], Turner syndrome [10], hearing loss [11] and cancer [12]. Fourth, mutations in the CRP genes of *D. melanogaster* are important tools for the study of growth, development and cell competition [2]. Finally, many RPs are conserved from bacteria to humans, so their peptide and nucleotide sequences are useful for studying phylogenetic relationships [13].

The first eukaryotic CRPs characterized in detail were isolated from the rat cytoribosome [3]. Individual proteins were

separated by two-dimensional gel electrophoresis and named from their origin in the small (S) or large (L) subunit and their relative electrophoretic migration positions, for example, RPS9 or RPL28. Subsequent studies revealed that some protein spots contained non-ribosomal proteins or chemically modified versions of another CRP, and that some spots contained two co-migrating CRPs [3,5]. Consequently, the nomenclature system used today contains numerical gaps as well as 'A' suffixes for those additional CRPs not resolved by the original electrophoresis (for example, RPL36A). Seventy-nine distinct mammalian CRPs are now acknowledged and their amino acid sequences and biochemical properties have been described [5,14]. With the exception of RPLP1 and RPLP2, each of which forms homodimers in the cytoribosomal large subunit [15], all CRPs are present as single molecules in each cytoribosome [3].

Seventy-eight different mammalian MRPs have been described [6] and their individual amino acid sequences and biochemical properties have been determined [16,17]. Although the nomenclature of MRPs was originally based on electrophoretic properties, the current system reflects homology between mammalian MRPs and their bacterial orthologs [18]. Thus, MRPS1 through MRPS21 are orthologous to *Escherichia coli* RPs S1-S21, while higher numbers have been assigned to the MRPs not found in bacteria. Gaps also exist in MRP numbering because a gap occurs in the bacterial enumeration or because there is no mammalian ortholog.

The RPs of *D. melanogaster* were first studied in the 1970s and early 1980s. Up to 78 individual CRPs were observed on two-dimensional gels [19-31] and about 30 were purified and analyzed biochemically [32,33]. A more recent characterization used mass spectrometry to identify 52 *D. melanogaster* CRPs [34], all of which are orthologous to known mammalian CRPs. The protein composition of *Drosophila* mitoribosomes has not been characterized biochemically to date.

CRPs and MRPs are encoded by the nuclear genome. Knowledge of the primary sequences of rat CRPs and bovine MRPs has led to the identification and mapping of the RP-encoding genes in many eukaryotic species [14]. Indeed, systematic analyses of whole RP gene sets have been described for several organisms, including *Saccharomyces cerevisiae* [35], *Arabidopsis thaliana* [36] and humans [37-40]. However, the complete set of *D. melanogaster* CRP and MRP genes has not been previously documented or characterized.

Several *D. melanogaster* RP genes were initially identified by virtue of their dominant 'Minute' mutant phenotypes [2], which include prolonged development, low fertility and viability, altered body size and abnormally short, thin bristles on the adult body. All of these phenotypes may be explained by a cell-autonomous defect in protein biosynthesis: the production of each bristle, for example, requires a very high rate of

protein synthesis in a single cell during a short developmental period. Merriam and colleagues reported the first unequivocal molecular link between a *Minute* locus and a CRP gene in 1985 [41]. Since then, 14 additional *Minute* loci have been definitively linked with distinct CRP genes [2,42-53]. However, there are at least 35 genetically validated *Minute* loci that have not yet been associated with a specific gene and there may be additional *Minute* genes to be discovered. Several investigators have hypothesized that all *Minute* loci encode protein components of ribosomes (reviewed in [2]). Whether this is truly the case and whether both CRP and MRP genes are associated with *Minute* phenotypes are open and intriguing questions.

Many *Minute* loci were originally identified from the phenotypes of flies heterozygous for a chromosomal deletion [54,55] and all *Minute* point mutations studied in depth have been found to be loss-of-function alleles [2]. This indicates that *Minute* phenotypes can be attributed to genetic haploinsufficiency; that is, a single gene copy is not sufficient for normal development. (Note that *X*-linked mutations that cause *Minute* phenotypes in heterozygous females are lethal in hemizygous males.) The most popular explanation for the haploinsufficiency of *Minute* loci is that they correspond to RP genes and that RPs are required in equimolar amounts: halving the copy number of a single RP gene limits the availability of the encoded RP, thereby reducing the number of functional ribosomes that are assembled in the cell and impairing protein synthesis [2]. While this idea is consistent with the available data, there may be other explanations.

The reduced fertility and viability associated with many *Minute* loci makes the recovery of deletions uncovering them rather difficult - the mutant strains are too weak to maintain as stable heterozygous stocks. In fact, some *Minute* loci are known only from the phenotypes of transient aneuploids [54,56]. This means that several chromosomal regions containing a *Minute* locus are not uncovered by current deletion collections [57]. This is frustrating for researchers because deletions are basic tools for mutational analysis and are widely used for mapping new mutations and identifying genetic modifiers. Efforts to maximize deletion coverage of the *D. melanogaster* genome would benefit from a systematic assessment of the relationship between RP genes and *Minute* loci. It would allow the isolation of deletions that flank haploinsufficient RP genes as closely as possible, or the design of transgenic constructs or chromosomal duplications to rescue the haploinsufficiency of deletions uncovering *Minute* genes.

Here, we report the systematic identification, naming and characterization of all the CRP and MRP genes of *D. melanogaster*. We have used this information, together with phenotypic data obtained from examining mutation and deficiency strains, to assess the correspondence between RP genes and *Minute* loci. We find that 66 of the 88 CRP genes identified are, or are very likely to be, haploinsufficient and

associated with a *Minute* phenotype, whereas MRP genes and the remaining 22 CRP genes are not. Significantly, we show that all but one of the known *Minute* loci in the genome correspond to CRP genes - the single exception encodes a subunit of an essential translation initiation factor. Together, these results identify the majority of haploinsufficient loci in the *D. melanogaster* genome that significantly affect viability, fertility and/or external morphology, and also provide a mechanistic framework for understanding the *Minute* syndrome and the phenotypic effects of aneuploidy.

Results

Identification of *D. melanogaster* ribosomal protein genes

In order to conduct an exhaustive survey of *Drosophila* CRP and MRP genes, we first performed a series of BLAST searches using human RP sequences as queries, because both CRPs and MRPs have been well-characterized in humans [5,6]. Tables 1 and 2 list the genes we identified together with their cytological locations. Where necessary, *D. melanogaster* genes were named or renamed according to the standard metazoan RP gene nomenclature proposed by Wool and colleagues [5,58,59] and approved by the HUGO Gene Nomenclature Committee [18], whilst still conforming to FlyBase [60] conventions - that is, CRP genes are given an '*Rp*' prefix and MRP genes have an '*mRp*' prefix. The seven exceptions to this standard RP nomenclature are mostly genes originally named to reflect a mutant phenotype, for example, the *string of pearls* (*sop*) gene encodes RpS2 [61] and *bonsai* encodes mRpS15 [62,63]. In these cases, the original gene symbol has been preserved, with the apposite RP symbol given as a synonym.

Cytoplasmic ribosomal protein genes

We identified 88 genes that encode a total of 79 different CRPs (Table 1). Thus, the *D. melanogaster* proteome contains orthologs of all 79 mammalian CRPs (32 small subunit and 47 large subunit proteins). While the majority of CRPs are encoded by single genes, nine are encoded by two distinct genes. In addition, we identified another five genes predicted to encode proteins with significantly lower similarity to human CRPs, which we term 'CRP-like' genes. Two fragments of the *RpS6* gene were also identified. (The list of 88 CRP genes presented by Cherry *et al.* [64] originated from an earlier report of our results to FlyBase (MA and SJM, FBrf0178764). These authors also list five CRP-like genes from our original report, but two of these have been eliminated and two additional CRP-like genes have been added in the current analysis.)

The deduced characteristics of *D. melanogaster* and human CRPs are compared in Additional data file 1. As might be expected, the amino acid identity between the CRPs of the two species is very high (average of 69% with a range of 27-98%, excluding the CRP-like proteins) and the predicted molecular

Table 1**The CRP genes of *D. melanogaster***

Human CRP	<i>D. melanogaster</i> gene			
	Symbol*	CG number	Location†	BLAST E value‡
RPSA	<i>sta/RpSA</i>	CG14792	X: 2B1	3e-94
RPS2	<i>sop/RpS2</i>	CG5920	2L: 30E1	e-118
RPS3	<i>RpS3</i>	CG6779	3R: 94E13	e-106
RPS3A	<i>RpS3A</i>	CG2168	4: 101F1	e-100
RPS4	<i>RpS4</i>	CG11276	3L: 69F6	e-121
RPS5	<i>RpS5a</i>	CG8922	X: 15E5-7	1e-98
	RpS5b	CG7014	3R: 88D6	3e-96
RPS6	<i>RpS6</i>	CG10944	X: 7C2	e-103
	CG11386	CG11386	X: 7C2	2e-15
	CG33222	CG33222	X: 7C2	2e-15
RPS7	<i>RpS7</i>	CG1883	3R: 99E2	1e-74
RPS8	<i>RpS8</i>	CG7808	3R: 99C4	1e-82
RPS9	<i>RpS9</i>	CG3395	3L: 67B11	8e-92
RPS10	RpS10a	CG12275	3R: 98A14	4e-43
	<i>RpS10b</i>	CG14206	X: 18D3	8e-52
RPS11	<i>RpS11</i>	CG8857	2R: 48E8-9	1e-58
RPS12	<i>RpS12</i>	CG11271	3L: 69F5	2e-43
RPS13	<i>RpS13</i>	CG13389	2L: 29B2	2e-74
RPS14	<i>RpS14a</i>	CG1524	X: 7C6-7	3e-71
	RpS14b	CG1527	X: 7C8	3e-71
RPS15	<i>RpS15</i>	CG8332	2R: 53C8	6e-62
RPS15A	<i>RpS15Aa</i>	CG2033	X: 11E11-12	2e-65
	RpS15Ab	CG12324	2R: 47C1	6e-65
RPS16	<i>RpS16</i>	CG4046	2R: 58F1	2e-69
RPS17	<i>RpS17</i>	CG3922	3L: 67B5	5e-52
RPS18	<i>RpS18</i>	CG8900	2R: 56F11	2e-69
RPS19	<i>RpS19a</i>	CG4464	X: 14F4	5e-48
	RpS19b	CG5338	3R: 95C13	1e-43
RPS20	<i>RpS20</i>	CG15693	3R: 93A1	7e-50
RPS21	<i>oho23B/RpS21</i>	CG2986	2L: 23B6	3e-30
RPS23	<i>RpS23</i>	CG8415	2R: 50E4	8e-70
RPS24	<i>RpS24</i>	CG3751	2R: 58F3	7e-55
RPS25	<i>RpS25</i>	CG6684	3R: 86D8	1e-38
RPS26	<i>RpS26</i>	CG10305	2L: 36F4	3e-47
RPS27	<i>RpS27</i>	CG10423	3R: 96C8	2e-39
RPS27A	<i>RpS27A</i>	CG5271	2L: 31E1	8e-80
RPS28	RpS28a	CG15527	3R: 99D2	1e-21
	<i>RpS28b</i>	CG2998	X: 8E7	2e-23
	RpS28-like	CG34182	2L: 30B3	1e-07
RPS29	<i>RpS29</i>	CG8495	3R: 85E8	5e-23
RPS30	<i>RpS30</i>	CG15697	3R: 93A2	6e-32
RPLP0	<i>RpLp0</i>	CG7490	3L: 79B2	e-122

Table 1 (Continued)**The CRP genes of *D. melanogaster***

	RpLP0-like	<i>CG1381</i>	2R: 46E5-6	3e-10
RPLP1	<i>RpLP1</i>	<i>CG4087</i>	2L: 21C2	3e-34
RPLP2	<i>RpLP2</i>	<i>CG4918</i>	2R: 53C9	5e-32
RPL3	<i>RpL3</i>	<i>CG4863</i>	3R: 86D8	0.0
RPL4	<i>RpL4</i>	<i>CG5502</i>	3R: 98B6	e-141
RPL5	<i>RpL5</i>	<i>CG17489</i>	2L: h35/40B	e-120
RPL6	<i>RpL6</i>	<i>CG11522</i>	3R: 100C7	8e-58
RPL7	<i>RpL7</i>	<i>CG4897</i>	2L: 31B1	2e-75
	RpL7-like	<i>CG5317</i>	2L: 33C1	3e-32
RPL7A	<i>RpL7A</i>	<i>CG3314</i>	X: 6B1	e-102
RPL8	<i>RpL8</i>	<i>CG1263</i>	3L: 62E7	e-119
RPL9	<i>RpL9</i>	<i>CG6141</i>	2L: 32C1	1e-66
RPL10	<i>Qm/RpL10</i>	<i>CG17521</i>	3L: h47/80A	7e-98
RPL10A	RpL10Aa	<i>CG3843</i>	3R: 88D10	3e-65
	<i>RpL10Ab</i>	<i>CG7283</i>	3L: 68E1	3e-95
RPL11	<i>RpL11</i>	<i>CG7726</i>	2R: 56D7	6e-83
RPL12	<i>RpL12</i>	<i>CG3195</i>	2R: 60B7	7e-75
RPL13	<i>RpL13</i>	<i>CG4651</i>	2L: 30F3	1e-68
RPL13A	<i>RpL13A</i>	<i>CG1475</i>	3R: 83B6-7	3e-68
RPL14	<i>RpL14</i>	<i>CG6253</i>	3L: 66D8	2e-30
RPL15	<i>RpL15</i>	<i>CG17420</i>	3L: h50-52/80F	5e-90
RPL17	<i>RpL17</i>	<i>CG3203</i>	X: 6C10	5e-72
RPL18	<i>RpL18</i>	<i>CG8615</i>	3L: 65E9	3e-71
RPL18A	<i>RpL18A</i>	<i>CG6510</i>	2R: 54C3	3e-68
RPL19	<i>RpL19</i>	<i>CG2746</i>	2R: 60E11	4e-83
RPL21	<i>RpL21</i>	<i>CG12775</i>	2L: 40A-B	4e-66
RPL22	<i>RpL22</i>	<i>CG7434</i>	X: 1C4	4e-40
	RpL22-like	<i>CG9871</i>	2R: 59D3	2e-24
RPL23	<i>RpL23</i>	<i>CG3661</i>	2R: 59B3	1e-68
RPL23A	<i>RpL23A</i>	<i>CG7977</i>	3L: 62A10	7e-52
RPL24	<i>RpL24</i>	<i>CG9282</i>	2L: 34B10	7e-55
	RpL24-like	<i>CG6764</i>	3R: 86E5	8e-14
RPL26	<i>RpL26</i>	<i>CG6846</i>	3L: 75E4	1e-59
RPL27	<i>RpL27</i>	<i>CG4759</i>	3R: 96E9-10	1e-43
RPL27A	<i>RpL27A</i>	<i>CG15442</i>	2L: 24F3	5e-60
RPL28	<i>RpL28</i>	<i>CG12740</i>	3L: 63B14	3e-31
RPL29	<i>RpL29</i>	<i>CG10071</i>	2R: 57D8	8e-14
RPL30	<i>RpL30</i>	<i>CG10652</i>	2L: 37B9	3e-46
RPL31	<i>RpL31</i>	<i>CG1821</i>	2R: 45F5	5e-47
RPL32	<i>RpL32</i>	<i>CG7939</i>	3R: 99D3	6e-58
RPL34	RpL34a	<i>CG6090</i>	3R: 96F10	2e-29
	<i>RpL34b</i>	<i>CG9354</i>	3R: 85D15	1e-29
RPL35	<i>RpL35</i>	<i>CG4111</i>	X: 5A11	2e-38
RPL35A	<i>RpL35A</i>	<i>CG2099</i>	3R: 83A4	1e-35
RPL36	<i>RpL36</i>	<i>CG7622</i>	X: 1B12	3e-33

Table 1 (Continued)**The CRP genes of *D. melanogaster***

RPL36A	<i>RpL36A</i>	CG7424	2L: 28D3	3e-43
RPL37	<i>RpL37a</i>	CG9091	X: 13B1	2e-39
	<i>RpL37b</i>	CG9873	2R: 59C4	1e-31
RPL37A	<i>RpL37A</i>	CG5827	2L: 25C4	2e-38
RPL38	<i>RpL38</i>	CG18001	2R: h46/41C-E	2e-25
RPL39	<i>RpL39</i>	CG3997	2R: 60B7	3e-18
RPL40	<i>RpL40</i>	CG2960	2L: 24E1	2e-69
RPL41	<i>RpL41</i>	CG30425	2R: 60E5	8e-08

*Additional gene synonyms exist in most cases [60]. Bold font indicates CRP-like genes, putative pseudogenic fragments (CG11386 and CG33222) or the member of a duplicate gene pair that is expressed in a small number of tissues and/or at relatively low levels. †Computed cytological position is given for euchromatic genes (Genome Release 5 [60]). The cytological and h-band locations for heterochromatic genes are based on data in reference [151] or estimated from images of *in situ* hybridizations of BACs to polytene chromosomes (*RpL5* and *RpL21*) [152]. The h-band location of *RpL15* was provided by B Honda (personal communication). ‡Expect (E) value obtained from a BLASTp search of the *D. melanogaster* annotated proteome (Genome Release 5.1) with human RefSeq CRP sequences. (E values corresponding to *RpL15* and *RpS28-like* were obtained from a BLAST search using Release 5.3.) Where multiple protein isoforms exist, the highest scoring hit is given.

weights and isoelectric points of the homologous proteins are very similar. However, several *D. melanogaster* proteins (*RpL14*, *RpL22*, *RpL23A*, *RpL29*, *RpL34a*, *RpL34b*, *RpL35A*) have significantly lower overall identity and different molecular weights owing to terminal deletions or extensions (data not shown; also see [65]). (If these seven proteins are discounted, the average identity of fly and human CRPs increases to 72% with a range of 43-98%.) Similar to humans and other species, there are very few acidic CRPs in *D. melanogaster*: only six proteins (*RpSA*, *RpS12*, *RpS21*, *RpLP0*, *RpLP1* and *RpLP2*) have isoelectric points less than pH 7. (Note that *RpS21* is an acidic protein, whereas its human counterpart is basic.) As in other eukaryotes, *RpS27A* and *RpL40* are carboxyl extensions of ubiquitin [66-69], and, as in other animals, *RpS30* is fused to a ubiquitin-like sequence. From these gross characterizations of component proteins, it appears that the fly cytoribosome differs only slightly from its human counterpart and is essentially the same as other eukaryotic cytoribosomes.

Previous biochemical analyses estimated that the *D. melanogaster* cytoribosome contains up to 78 CRPs [29]. This figure compares very well to the 79 different CRPs predicted by our orthology analysis (Table 1). Unfortunately, very few of the CRPs identified in the 1970s and 1980s were characterized to the level of amino acid sequence, so their correspondences to CRP genes are generally unknown, though there are a few exceptions (see references [70-73]). We have been unable, therefore, to correlate the CRPs identified in these earlier studies with those encoded by the CRP genes identified in this study. In contrast, our CRP inventory certainly does contain all 52 CRPs identified by the recent biochemical analysis of *D. melanogaster* cytoribosomes by Alonso and Santarén [34].

Mitochondrial ribosomal protein genes

We identified 75 *D. melanogaster* genes encoding proteins of the mitoribosome (28 in the small subunit and 47 in the large

subunit) by orthology to human MRPs (Table 2). These data complement and extend previous analyses of homology between human and *D. melanogaster* MRPs [16,17]. As in these previous studies, genes encoding orthologs of three human MRPs (MRPS27, MRPS36 and LACTB/MRPL56) were not found.

The MRPs of humans and *D. melanogaster* are much more divergent than are their CRPs: MRPs have an average identity of only 34% (with a range of 15-57%) and several homologous pairs differ markedly in their sizes and isoelectric points (Additional data file 2). Indeed, it is known that the mitoribosome is a rapidly evolving structure whose composition varies among eukaryotic organisms [6]. It is quite possible that there are proteins in *Drosophila* mitoribosomes that are not found in their human counterparts and these will have been missed by our orthology analysis - a definitive inventory will require biochemical characterization of the fly mitoribosome. As in mammals, three distinct genes encode three different isoforms of MRPS18 (Table 2); it is thought that each mitoribosome contains a single MRPS18 protein and that mitoribosomes may, therefore, be heterogeneous in composition [6].

Duplicate cytoplasmic ribosomal protein genes

Of the 79 different CRPs of *D. melanogaster*, 9 are encoded by two distinct genes (Table 1). These are distinguished by a lowercase 'a' or 'b' suffix to the gene symbol. (The lowercase 'a' should not be confused with the uppercase 'A' suffix used in the standard CRP nomenclature; for example, *RpL37a* and *RpL37A* are different genes that encode different proteins.) Six of these gene pairs encode proteins of the small ribosomal subunit and the other three encode large subunit proteins. In humans, each CRP is typically encoded by a single, functional gene [37,74], but thousands of nonfunctional CRP pseudogenes are known to exist [75]. We therefore investigated the evolutionary origin, sequence conservation and expression

Table 2**The MRP genes of *D. melanogaster***

Human MRP	<i>D. melanogaster</i> gene			
	Symbol*	CG number	Location†	BLAST E value‡
MRPS2	<i>mRpS2</i>	CG2937	2L: 25B1	1e-69
MRPS5	<i>mRpS5</i>	CG40049	3L: h47/80A-B	4e-63
MRPS6	<i>mRpS6</i>	CG15016	3L: 64B2	2e-20
MRPS7	<i>mRpS7</i>	CG5108	2L: 31D11	6e-38
MRPS9	<i>mRpS9</i>	CG2957	3R: 84E4	8e-81
MRPS10	<i>mRpS10</i>	CG4247	3R: 88E3	1e-33
MRPS11	<i>mRpS11</i>	CG5184	3R: 89E11	8e-26
MRPS12	<i>tko/mRpS12</i>	CG7925	X: 3A3	1e-33
MRPS14	<i>mRpS14</i>	CG32531	X: 18C7	1e-35
MRPS15	<i>bonsai/mRpS15</i>	CG4207	2R: 58F3	1e-15
MRPS16	<i>mRpS16</i>	CG8338	2R: 50E1	2e-24
MRPS17	<i>mRpS17</i>	CG4326	2R: 60C1	2e-14
MRPS18A	<i>mRpS18A</i>	CG31450	3R: 85A3	8e-13
MRPS18B	<i>mRpS18B</i>	CG10757	2L: 38B6	4e-33
MRPS18C	<i>mRpS18C</i>	CG9688	3R: 99F4	7e-23
MRPS21	<i>mRpS21</i>	CG32854	3R: 87E8	3e-22
MRPS22	<i>mRpS22</i>	CG12261	3R: 98B6	3e-38
MRPS23	<i>mRpS23</i>	CG31842	2L: 34D6	2e-20
MRPS24	<i>mRpS24</i>	CG13608	3R: 95E6	2e-31
MRPS25	<i>mRpS25</i>	CG14413	X: 12F1	1e-49
MRPS26	<i>mRpS26</i>	CG7354	3L: 75B9	3e-11
MRPS27	NA	NA	NA	No hit
MRPS28	<i>mRpS28</i>	CG5497	2R: 55E2	1e-27
DAP3/MRPS29	<i>mRpS29</i>	CG3633	2R: 58E1	7e-72
MRPS30	<i>mRpS30</i>	CG8470	X: 13E18	3e-21
MRPS31	<i>mRpS31</i>	CG5904	3L: 72C2	5e-35
MRPS33	<i>mRpS33</i>	CG10406	3R: 89B16	3e-30
MRPS34	<i>mRpS34</i>	CG13037	3L: 72E1-2	8e-06
MRPS35	<i>mRpS35</i>	CG2101	3L: 62F4	2e-72
MRPS36	NA	NA	NA	No hit
MRPL1	<i>mRpL1</i>	CG7494	3R: 84F9-10	8e-25
MRPL2	<i>mRpL2</i>	CG7636	3L: 68A7	2e-56
MRPL3	<i>mRpL3</i>	CG8288	X: 13E14	5e-52
MRPL4	<i>mRpL4</i>	CG5818	2L: 35F1	6e-70
MRPL9	<i>mRpL9</i>	CG31478	3R: 88F1	6e-17
MRPL10	<i>mRpL10</i>	CG11488	2L: 21B4	8e-20
MRPL11	<i>mRpL11</i>	CG3351	3R: 88C3	4e-29
MRPL12	<i>mRpL12</i>	CG5012	3L: 66E5	2e-13
MRPL13	<i>mRpL13</i>	CG10603	2L: 37B1	2e-47
MRPL14	<i>mRpL14</i>	CG14048	X: 3A1	3e-31
MRPL15	<i>mRpL15</i>	CG5219	3L: 77C3	4e-74
MRPL16	<i>mRpL16</i>	CG3109	X: 2B14	3e-58
MRPL17	<i>mRpL17</i>	CG13880	3L: 61B3	8e-21
MRPL18	<i>mRpL18</i>	CG12373	2R: 49C2	3e-22
MRPL19	<i>mRpL19</i>	CG8039	3R: 85A5	1e-56
MRPL20	<i>mRpL20</i>	CG11258	3L: 69F5	8e-21

Table 2 (Continued)**The MRP genes of *D. melanogaster***

MRPL21	<i>mRpL21</i>	CG9730	3L: 76A3	2e-19
MRPL22	<i>mRpL22</i>	CG4742	X: 15A7-8	4e-41
MRPL23	<i>mRpL23</i>	CG1320	3L: 62D7	4e-28
MRPL24	<i>mRpL24</i>	CG8849	2L: 25B4	2e-45
MRPL27	<i>mRpL27</i>	CG33002	2L: 24F3	2e-10
MRPL28	<i>mRpL28</i>	CG3782	2L: 25B5	5e-27
MRPL30	<i>mRpL30</i>	CG7038	X: 4C11	3e-15
MRPL32	<i>mRpL32</i>	CG12220	3R: 100B8	6e-15
MRPL33	<i>mRpL33</i>	CG3712	X: 4B6	5e-08
MRPL34	<i>mRpL34</i>	CG34147	2R: 52E4	1e-07
MRPL35	<i>mRpL35</i>	CG13410	3R: 94A1	4e-25
MRPL36	<i>mRpL36</i>	CG18767	3L: 66B7	2e-09
MRPL37	<i>mRpL37</i>	CG6547	3R: 86C6	4e-16
MRPL38	<i>mRpL38</i>	CG15871	X: 12E5	2e-61
MRPL39	<i>mRpL39</i>	CG17166	3L: 71B1	2e-57
MRPL40	<i>mRpL40</i>	CG5242	3R: 86E4	6e-13
MRPL41	<i>mRpL41</i>	CG12954	2R: 51E7	2e-11
MRPL42	<i>mRpL42</i>	CG12921	2R: 46E1	1e-11
MRPL43	<i>mRpL43</i>	CG5479	2R: 59F6	1e-25
MRPL44	<i>mRpL44</i>	CG2109	3R: 83A4	6e-39
MRPL45	<i>mRpL45</i>	CG6949	3R: 94B6	4e-61
MRPL46	<i>mRpL46</i>	CG13922	3L: 62B4	5e-39
MRPL47	<i>Ric1/mRpL47</i>	CG9378	3R: 85D19	1e-35
MRPL48	<i>mRpL48</i>	CG17642	2L: 22B1	4e-15
MRPL49	<i>mRpL49</i>	CG4647	X: 11D1	9e-22
MRPL50	<i>mRpL50</i>	CG8612	3L: 65E9	6e-09
MRPL51	<i>mRpL51</i>	CG13098	2L: 29D4	2e-13
MRPL52	<i>mRpL52</i>	CG1577	2R: 43E9	2e-14
MRPL53	<i>mRpL53</i>	CG30481	2R: 50C16	3e-06
MRPL54	<i>mRpL54</i>	CG9353	2R: 57B16	2e-18
MRPL55	<i>mRpL55</i>	CG14283	3R: 91F1	3e-17
LACTB/MRPL56	NA	NA	NA	No hit

*Additional gene synonyms exist in many cases [60]. †Computed cytological position is given for euchromatic genes (Genome Release 5 [60]). *mRpS5* h-band location provided by C Smith (DHGP, personal communication) and cytological position inferred from reference [151]. ‡Expect (E) value obtained from a BLASTp search of the *D. melanogaster* annotated proteome (Genome Release 5.1) with human RefSeq MRP sequences. Where multiple protein isoforms exist, the highest scoring hit is given.

profile of the duplicate *D. melanogaster* CRP genes in order to assess whether both members of each pair are likely to be functional (Table 3 and Figure 1).

In five cases, one member of the gene pair lacks introns (*RpS10a*, *RpS15Ab*, *RpS28a*, *RpL10Aa* and *RpL37b*) while the other member does not. These five intronless genes are likely to have arisen by retrotransposition; that is, generated via reverse transcription of mRNA from the precursor gene followed by insertion into a new genomic location. In contrast, the *RpS5*, *RpS19* and *RpL34* duplicates arose through gene transposition events as both members of each pair retain introns. The *RpL34* duplication occurred through an intrachromosomal transposition on chromosome arm 3R, and *RpL34a* and *RpL34b* have retained almost identical gene

structures. In contrast, the *RpS5* and *RpS19* duplications involved interchromosomal transposition events that must have been followed by extensive gene remodeling as the intron-exon structures differ within each pair. Finally, *RpS14a* and *RpS14b* probably arose via unequal exchange: these paralogs are situated adjacent to each other as a tandem duplication on the X chromosome, share identical intron-exon structures and encode identical proteins [76]. All nine duplicate genes appear to have arisen within the Drosophilidae, albeit at different stages in the lineage leading to *D. melanogaster* (Figure 1).

Neither member of these 9 CRP gene pairs contains a nonsense mutation in the protein-coding region (data not shown), indicating that all 18 genes are potentially functional.

Table 3**Analysis of duplicate CRP genes and CRP-like genes**

	Gene			K_A/K_S^b		cDNA clones		
	Symbol ^a	Location	Comments	Pair-wise ^c	Branch-specific ^d	Total ^e	% testis ^f	% Amino acid identity ^g
RpS5	<i>RpS5a</i>	X: 15E5-7	-	0.07	0.09	133	1	78
	RpS5b	3R: 88D6	-			60	7	
RpS6	<i>RpS6</i>	X: 7C2	-	NA	NA	150	5	33 ^h
	CG11386	X: 7C2	<i>CG11386</i> and <i>CG33222</i> are tandem repeats of the third exon and flanking regions of <i>RpS6</i>			0	0	
	CG33222	X: 7C2	-			3	0	
RpS10	RpS10a	3R: 98A14	Lacks introns; likely retrogene	0.10	0.15	1	0	73
	<i>RpS10b</i>	X: 18D3	-			156	0	
RpS14	<i>RpS14a</i>	X: 7C6-7	<i>RpS14a</i> and <i>RpS14b</i> share identical gene structures	NA	NA	154	2	100
	RpS14b	X: 7C8				12	0	
RpS15A	<i>RpS15Aa</i>	X: 11E11-12	-	0.16	0.00	150	1	98
	RpS15Ab	2R: 47C1	Lacks introns; likely retrogene			67	1	
RpS19	<i>RpS19a</i>	X: 14F4	-	0.09	0.01	134	1	65
	RpS19b	3R: 95C13	-			1	100	
RpS28	RpS28a	3R: 99D2	Lacks introns; likely retrogene	0.05	0.06	0	0	82
	<i>RpS28b</i>	X: 8E7	-			131	1	
	RpS28-like	2L: 30B3	-			NA ⁱ	NA	
RpLP0	<i>RpLP0</i>	3L: 79B2	-	NA	NA	157	2	18
	RpLP0-like	2R: 46E5-6	Present in all eukaryotes			11	0	
RpL7	<i>RpL7</i>	2L: 31B1	-	NA	NA	168	4	30
	RpL7-like	2L: 33C1	-			32	0	
RpL10A	RpL10Aa	3R: 88D10	Lacks introns; likely retrogene	0.05	0.06	7	71	64
	<i>RpL10Ab</i>	3L: 68E1	-			160	4	
RpL22	<i>RpL22</i>	X: 1C4	-	NA	NA	145	2	38
	RpL22-like	2R: 59D3	-			5	60	
RpL24	<i>RpL24</i>	2L: 34B10	-	NA	NA	162	3	23
	RpL24-like	3R: 86E5	Present in all eukaryotes			34	3	
RpL34	RpL34a	3R: 96F10	<i>RpL34a</i> and <i>RpL34b</i> share identical gene structures	0.04	0.09	56	4	78
	<i>RpL34b</i>	3R: 85D15				132	2	
RpL37	<i>RpL37a</i>	X: 13B1	-	0.01	NA	158	3	72
	RpL37b	2R: 59C4	Lacks introns; likely retrogene			3	0 ^k	

^aBold font indicates CRP-like genes, putative pseudogenic fragments (*CG11386* and *CG33222*) or the member of a duplicate gene pair that is expressed in a small number of tissues and/or at relatively low levels. ^b K_A/K_S calculations are not applicable (NA) to highly diverged sequences or cases where the numbers of both synonymous and nonsynonymous substitutions are very small (<5). ^cCalculated for each *D. melanogaster* CRP gene pair using maximum likelihood analysis. Values for pairwise comparisons are shown on the first row of each pair. ^dBranch-specific score in a three-way maximum likelihood tree including *D. pseudoobscura* orthologs. A four-way tree was used for *RpS19* sequences. ^eTotal number of cDNA clones (excluding those from cultured cell lines) given in FlyBase [60] (April 2007). *RpS28-like* cDNA evidence from L Crosby (personal communication). ^fPercentage of cDNA clones from adult testis cDNA libraries (AT, UT and BS), rounded to the nearest integer. ^gIdentity between proteins across their whole length. Values for pairwise comparisons are shown on the first row of each pair. ^hIdentity between *RpS6* and the *CG11386* or *CG33222* protein. If *CG11386* or *CG33222* were used as alternative third exons of *RpS6*, the protein encoded would be 60% identical to the conventional *RpS6* (see text for details). ⁱ*RpS28-like* is too highly diverged from both *RpS28a* and *RpS28b* for a pair-wise K_A/K_S calculation to be applicable. ^jIdentity between the *RpS28-like* protein and *RpS28a/RpS28b*. ^kThere is experimental evidence that *RpL37b* expression is enriched in adult testis [79].

Moreover, the low ratio of nonsynonymous to synonymous substitutions (K_A/K_S) between the members of each gene pair suggests that there are selective constraints on their protein-coding regions (Table 3; a K_A/K_S ratio significantly lower than

0.5 indicates functional constraints on both genes). Branch-specific K_A/K_S values further indicate that the putatively retrotransposed genes have been under overall purifying selection since their formation. Together, these data argue that

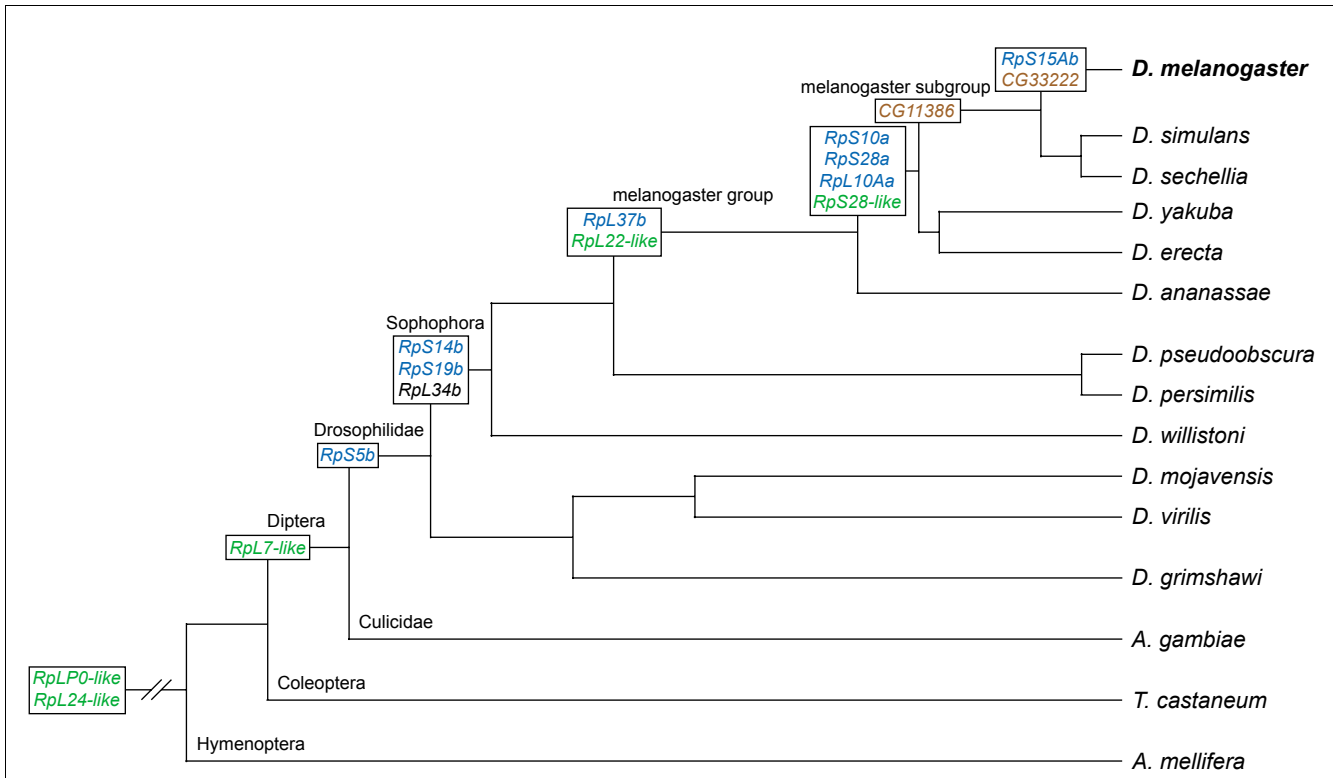


Figure 1
 Evolution of *D. melanogaster* CRP gene duplicates and CRP-like genes. The likely pattern of emergence of CRP duplicate genes with restricted expression (blue), CRP-like genes (green) and CRP pseudogenic fragments (brown) in the lineage leading to *D. melanogaster* is shown. *RpL34b* is shown in black text: this is the only case where the newly emerged duplicate gene (*RpL34b*), rather than precursor gene (*RpL34a*), acts as the principal gene copy. The relative placement of *CG11386* and *CG33222* is consistent with the model presented by Stewart and Denell [86]. The dendrogram is based on that given in reference [140], in which the relationships among the Drosophilidae are taken from [149]; note that the branch lengths do not accurately reflect evolutionary time.

none of these duplicate genes are nonfunctional pseudogenes, which is consistent with a previous analysis [77]. Indeed, the recovery of multiple cDNA clones for the majority (15/18) of these duplicate genes supports their expression *in vivo* (Table 3).

Although none of these CRP gene duplicates appear to be pseudogenes, it is evident that one member of each pair - the one with higher similarity to its human ortholog, where this difference exists (Table 1 and Additional data file 1) - is expressed at a significantly higher level and, in some cases, in a wider array of tissues than the other. This suggests that one gene of the pair produces the majority of each CRP in most cells, while the other gene has a more restricted expression pattern and, perhaps, a specialized function (indicated by bold font in Tables 1 and 3). In eight of the nine duplication events, the 'younger' gene copy has adopted the lower expression level or more restricted expression pattern; the *RpL34* gene pair is exceptional in this regard (Figure 1 and Table 3). The expression of *RpS5b*, *RpS19b*, *RpL10Aa* and *RpL37b* appears enriched in the adult testis, suggesting the existence of testis-specific CRPs and a testis-specific cytoribosome (Table 3). Significantly, three of these genes (*RpS5b*, *RpS19b*

and *RpL37b*), together with *RpS10a*, *RpS15Ab* and *RpS28a*, are autosomal copies of *X*-linked genes. These duplication events are consistent with previous studies reporting that genes with male-biased expression are predominantly autosomal [78], and that retrotransposed genes in *D. melanogaster* have preferentially retrotransposed from the *X* chromosome onto autosomes [79]. It is possible that these autosomal duplicates enable CRP expression in male germline cells, where it is hypothesized that *X* chromosome inactivation occurs during spermatogenesis [80]. Similarly, in humans, *RPS4Y* is a *Y*-linked duplicate of the *X*-linked *RPS4* gene [10] and *RPL10L*, *RPL36AL* and *RPL39L* are autosomal retrogene copies of *X*-linked progenitors [74]. It is worth noting that expression of *D. melanogaster* *RpS5b*, *RpS10a* and *RpS19b* is also enriched in the germline cells of embryonic gonads [81] and/or stem cells of adult ovaries [82]. These findings suggest a germline-specific role, rather than a testis-specific role, for these CRP gene duplicates.

To conclude, the 'principal' CRPs of *D. melanogaster* - those that are expressed at high levels in most cells - are each encoded by single genes.

Cytoplasmic ribosomal protein-like genes

We identified five *D. melanogaster* 'CRP-like' genes that encode proteins with significantly lower identity to human CRPs than those described above. These are *RpS28-like*, *RpLPO-like*, *RpL7-like*, *RpL22-like* and *RpL24-like* (shown in bold font in Tables 1 and 3). Of these, *RpLPO-like* and *RpL24-like* show the most divergence from their cognate proteins, *RpLPO* and *RpL24*. Consistent with this, *RpLPO-like* and *RpL24-like* have ancient evolutionary origins, while *RpL7-like*, *RpL22-like* and *RpS28-like* arose more recently within the Diptera (Figure 1).

cDNA evidence indicates that all five of these CRP-like genes are expressed *in vivo*, albeit at far lower levels than their cognate genes (Table 3). The evolutionary conservation of *RpLPO-like* and *RpL24-like* suggests they have important cellular functions. Indeed, the yeast ortholog of *RpL24-like* is found in pre-ribosomal complexes where it is thought to function in large subunit biogenesis [83]. It remains to be seen whether the other CRP-like proteins have similar functions. Interestingly, the *RpL22* gene is X-linked and expressed ubiquitously, whereas *RpL22-like* is an autosomal gene that is expressed predominantly in germline cells [81,82,84,85]. This suggests that *RpL22-like* may have a specialized role in the germline, and perhaps within germline-specific cytoribosomes, as proposed above for some of the CRP duplicates.

CG11386 and *CG33222* are 99% identical in DNA sequence and are tandem repeats of the third exon and flanking regions of the *RpS6* gene. They likely arose via two sequential unequal crossover events [86]; the first occurring after the evolutionary split of the *D. melanogaster* subgroup, and the second being specific to *D. melanogaster* (Figure 1). Gene prediction algorithms suggest that *CG11386* and *CG33222* are distinct genes encoding identical amino-terminally truncated versions of *RpS6* [87]; however, such proteins would lack critical functional domains and would probably be nonfunctional. In a different scenario, *CG11386* and/or *CG33222* could serve as alternative third exons of the *RpS6* gene: the proteins produced would be full-length, but would differ substantially in their carboxy-terminal two-thirds from the *RpS6* generated by using the conventional third exon [86]. There is, however, no direct evidence that such alternative transcripts are made. Indeed, only three cDNA clones suggest that *CG11386* or *CG33222* are expressed at all (Table 3). We have tentatively classified *CG11386* and *CG33222* as nonfunctional pseudogenic fragments.

Chromosomal distribution of ribosomal protein genes

As has been found for other eukaryotes [35,36,38,39], the RP genes of *D. melanogaster* are distributed across the entire genome (Figure 2). Some RP genes are tightly linked to other RP genes and, while this posed challenges for determining the phenotypes associated with individual genes (see below), we have no evidence that this distribution has functional consequences or that closely linked RP genes are transcriptionally co-regulated. Five RP genes (*RpL5*, *Qm/RpL10*, *RpL15*, *RpL38*, and *mRpS5*) are located within heterochromatic regions, as are certain human MRP genes [38] and some *Arabidopsis thaliana* CRP genes [36]. As heterochromatin is generally associated with the silencing of gene expression [88], the regulation of these genes must have adapted to the heterochromatic environment in order for the encoded proteins to be expressed at sufficiently high levels to meet the demand for ribosome synthesis in the cell [89].

Ribosomal protein gene haploinsufficiency and the Minute syndrome

Classical genetic studies have defined more than fifty regions of the *D. melanogaster* genome that are haploinsufficient and associated with the dominant phenotypes of prolonged development and short, thin bristles - the *Minute* loci [2] (Figure 3). To date, only fifteen *Minute* loci have been tied unequivocally to molecularly defined genes and all of these encode RPs (reviewed in reference [2]; also see references [48-53]). It has not been clear, however, if all *Minute* loci correspond to RP genes, or whether *Minute* loci may correspond to both CRP and MRP genes. We have conducted a new survey of *Minute* loci in the *D. melanogaster* genome which, combined with our RP gene inventory, has now allowed us to assess these relationships systematically.

Recent large-scale projects have provided a wealth of new genetic reagents that enable the mapping of *Minute* loci with a precision unavailable only a few years ago. Hundreds of new deletions with molecularly defined breakpoints have been provided by the efforts of the DrosDel consortium [90,91], Exelixis, Inc. [92], and the Bloomington *Drosophila* Stock Center [92]. When combined with older deletions characterized primarily through polytene chromosome cytology, these deletions have increased euchromatic genome coverage to 96-97%. In addition, transposable element insertions now exist within 0.5 kb of 57% of all genes (R Levis, personal communication), largely through the efforts of the *Drosophila* Gene Disruption Project [93] and Exelixis, Inc. [94]. We used

Figure 2 (see following page)

Chromosomal map of the RP genes of *D. melanogaster*. RP genes are depicted on a physical map of the genome (Release 5) [60]. Genes encoded on the positive and negative strands are shown above and below the chromosome, respectively. (The orientation of *RpL15* is not known and its position below the chromosome is arbitrary.) Chromosomes are divided into cytological bands as determined from sequence-to-cytogenetic band correspondence tables [150]. *Minute* genes are boxed as described in the key.

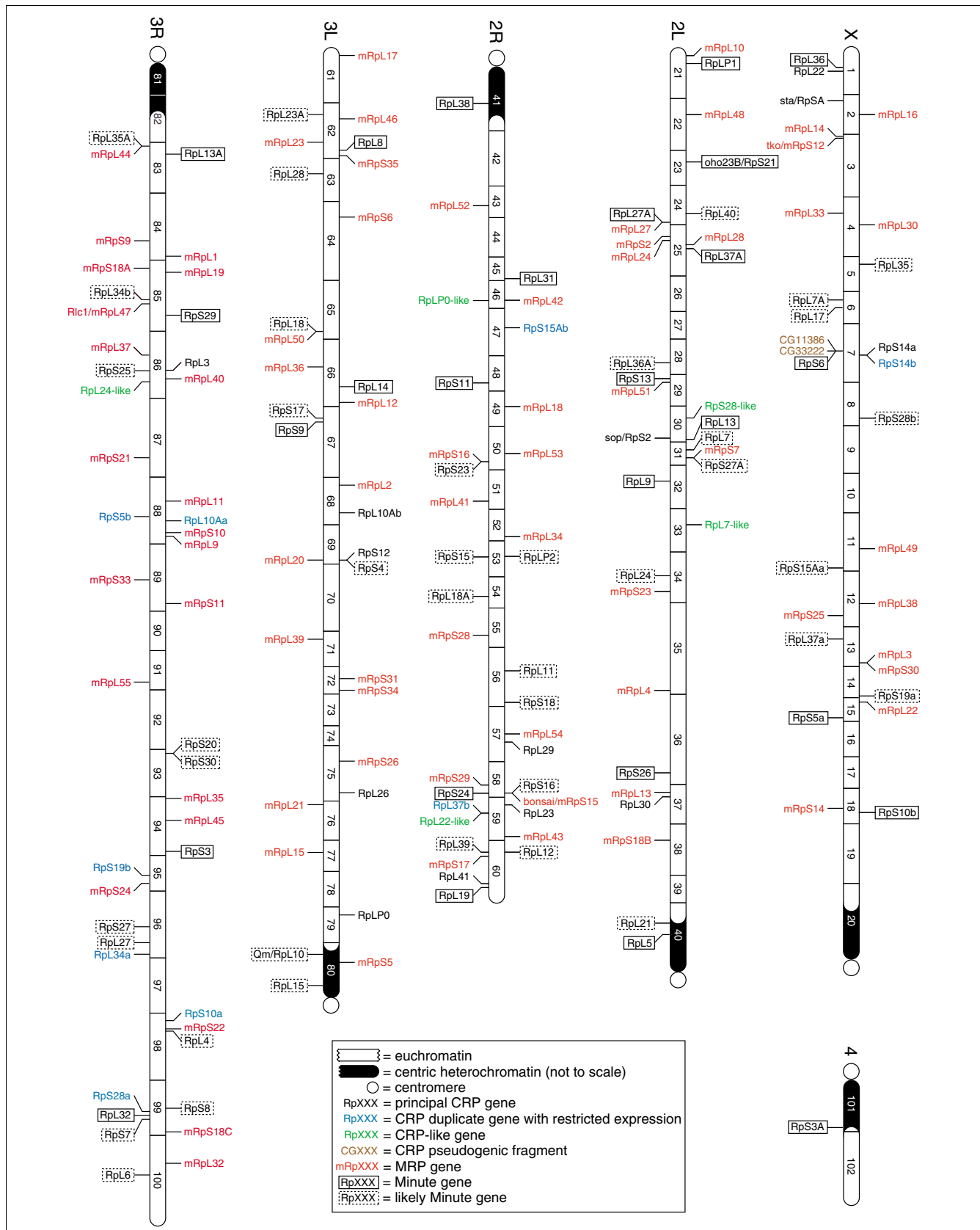


Figure 2 (see legend on previous page)

these resources to conduct a genome-wide search for *Minute* loci. In so doing, we considered the characteristic Minute bristle phenotype (Figure 3) to be diagnostic of the Minute syndrome; we did not methodically evaluate more subtle Minute traits, such as slower development, or traits observed in only a subset of *Minute* mutants, such as impaired fecundity, reduced viability or altered body size. By combining our observations with information gleaned from published studies, we have identified 61 distinct *Minute* loci. Many of these correlate with *Minute* loci described previously (Additional data file 3), though our work has often refined their map positions. Significantly, six *Minute* loci (*M(2)31E*, *M(2)34BC*, *M(2)45F*, *M(2)50E*, *M(3)93A* and *M(3)98B*) are reported here for the first time. We also found four instances (*M(2)31A*, *M(2)53*, *M(2)58F* and *M(3)67C*) where a single *Minute* locus characterized by previous aneuploidy analyses actually comprises two separable, closely linked *Minute* genes. As we have inferred the existence of four additional *Minute* loci from patterns of deletion coverage (described below), we conclude that there are 65 distinct *Minute* loci in the *D. melanogaster* genome.

We were able to demonstrate definitively that a particular *Minute* locus corresponds to a specific RP gene when a Minute bristle phenotype was observed in one or more of the following situations: flies heterozygous for a molecularly characterized mutation in a RP gene (for example, *M(2)36F/RpS26*); flies heterozygous for a chromosomal deletion when the Minute phenotype could be mapped unambiguously to a single RP gene with deletion breakpoints (for example, *M(2)25C/RpL37A*); or flies heterozygous for a chromosomal deletion when the Minute phenotype could be rescued by a specific RP transgene (for example, *M(3)99D/RpL32*). We found that there are 26 unequivocally *Minute* CRP genes by these criteria (Additional data file 4; summarized in Table 4). In contrast, no MRP or CRP-like genes were definitively demonstrated to be *Minute* genes.

These 26 cases of proven CRP gene-*Minute* locus correspondences provide a strong precedent for expecting that other CRP genes are also *Minute* genes. Although existing reagents do not allow us to demonstrate the correspondences definitively, we judged that a CRP gene very likely corresponds to a genetically defined *Minute* locus when one or more of the following criteria are fulfilled: a Minute phenotype is seen for a heterozygous multi-gene deletion that uncovers a single CRP gene (for example, *M(3)63B/RpL28*); a CRP gene lies in a gap in deletion coverage and a molecularly uncharacterized *Minute* mutation maps to the same region (for example, *M(1)8F/RpS28b*); or a CRP gene lies in a gap in deletion coverage and previous studies of transient aneuploids document the presence of a *Minute* locus in the same region (for example, *M(3)99E/RpS7*). In this way, we identified an additional 36 CRP genes that likely correspond to 34 genetically defined *Minute* loci (Additional data file 4; summarized in Table 4). Closely linked pairs of CRP genes map to the same regions as

M(2)60B and *M(3)93A* and, as it was impossible to determine whether one or both genes of each pair are haploinsufficient, we have classified all four CRP genes as likely *Minute* genes. No CRP-like genes mapped to the regions of proven *Minute* loci. Although five MRP genes map to regions containing *Minute* loci, it is unlikely that any of them are haploinsufficient: MRP genes are not associated with Minute phenotypes in any other situation, and each of these five MRP genes is closely linked to a CRP gene (Additional data file 4).

We concluded that a further four CRP genes (*RpL17*, *RpL18A*, *RpL34b* and *RpL35A*) are likely to be *Minute* genes despite no Minute phenotype having been associated with the genomic region in which they reside. In each of these cases, the CRP gene lies in a gap in deletion coverage (Table 4, Additional data file 4), suggesting that it is a *Minute* associated with strongly reduced fertility and/or viability, which prevents the establishment of stable deletion stocks (in the absence of a corresponding duplication). Supporting this view, such severe haploinsufficiency also appears to be associated with 15 other CRP genes - all these CRP genes lie in gaps in deletion coverage and they are only considered *Minute* genes here because they have point or transposon insertion (likely hypomorphic) mutations that cause Minute phenotypes, or because they lie in regions known to harbour *Minute* loci from the phenotypes of transient aneuploids (Table 4, Additional data file 4).

For all of the 40 CRP genes classified as 'likely *Minute* genes' (through correlation with genetically proven *Minute* loci or gaps in deletion coverage), we determined the maximum number of candidate genes that could possibly account for the haploinsufficiency. We used deletions to define the smallest chromosomal interval containing the *Minute* and then eliminated genes known not to be associated with a Minute phenotype from previous studies or from our own examinations of mutant fly strains. (This task benefited greatly from the recent work of the Bloomington *Drosophila* Stock Center which, in its efforts to maximize genomic deletion coverage, has systematically generated deletions flanking haploinsufficient loci.) The number of candidate genes defined in this way was always small, ranging from 2 to 33 genes with a median of 8.5 candidate genes per *Minute* locus (Table 4, Additional data file 4). These data increase our confidence in the likely correspondences between these *Minute* loci and CRP genes.

The results presented above indicate that 66 CRP genes are, or are likely to be, *Minute* genes, whereas the remaining 22 CRP genes are not (Table 4 and Additional data files 4 and 5; summarized in Figure 4). CRPs of the large and small ribosomal subunit are encoded by both *Minute* and non-*Minute* genes, with no apparent bias. Notably, none of the nine duplicate CRP genes with relatively restricted expression is a *Minute*, whereas seven of the more highly and widely expressed gene pair members are *Minute* genes. This is consistent with the idea that only one member of each of these

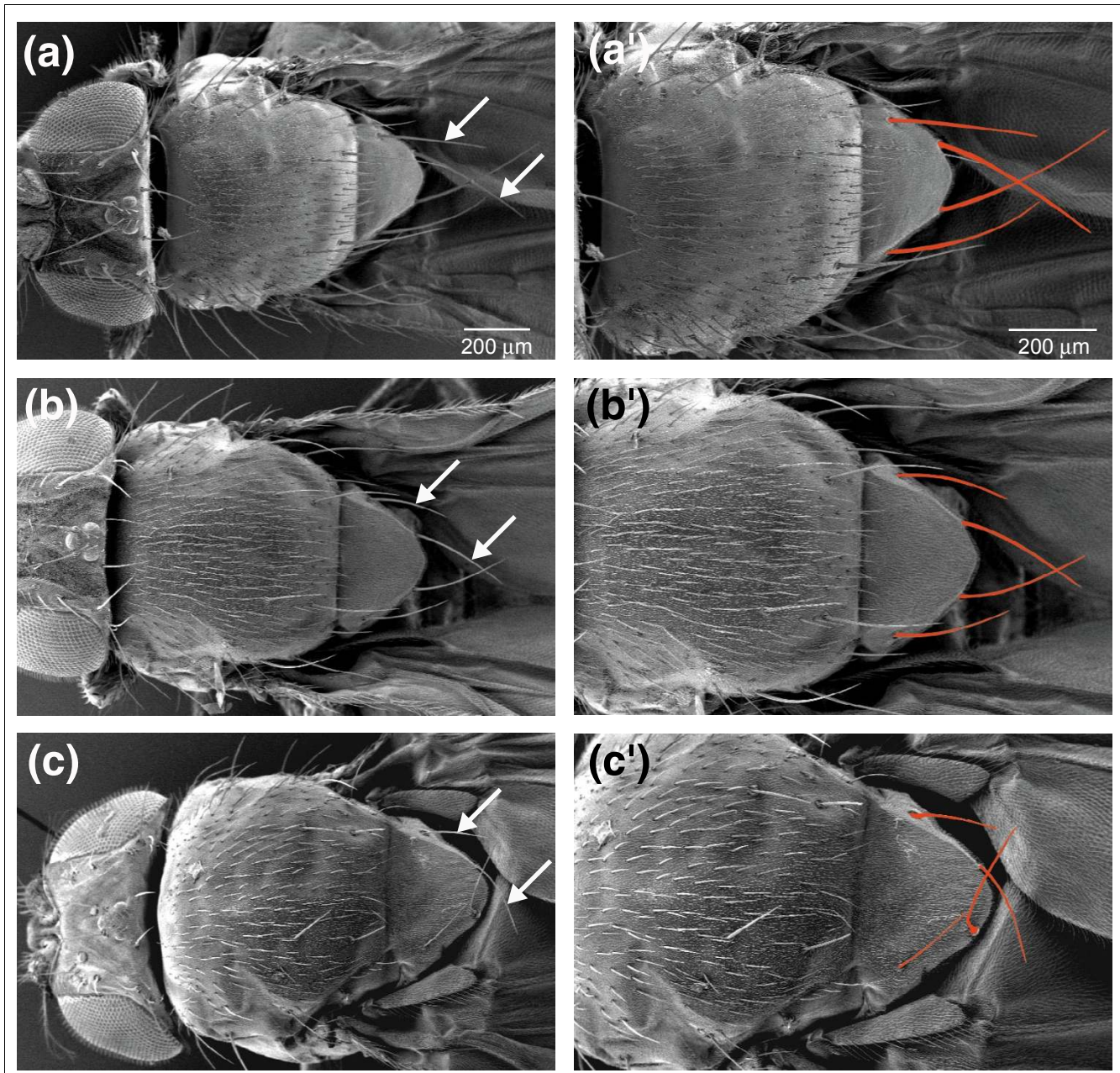


Figure 3
The Minute bristle phenotype. Minute flies have shorter and thinner bristles than wild type flies. This is most clearly seen by comparing the scutellar bristles, indicated here by the arrows and pseudocoloring. (a, a') Wild type. (b, b') *RpS13*¹ heterozygotes. (c, c') *RpL14*¹ heterozygotes.

gene pairs contributes significantly to cytoribosomal function in the majority of cells, while the one with restricted expression encodes a component of qualitatively distinct cytoribosomes in certain cell types. As the 'principal' copy of *RpS14* or *RpL10A* is not a *Minute*, it is unsurprising that the simultaneous heterozygous deletion of both *RpS14* genes [95] or both *RpL10A* genes (in flies with genotypes *Df(3L)ED4475/Df(3R)ED10556* or *Df(3L)ED4475/Df(3R)ED5660*; our observations) does not produce flies with Minute phenotypes.

Other possible reasons for different dosage sensitivities among CRP genes are discussed below.

The one verified *Minute* locus that does not correspond to a CRP gene is *M(1)14C*. We mapped this *Minute* gene to region 14C6 by showing that the deletions *Df(1)ED7364* (14A8;14C6) and *Df(1)FDD-0230908* (14C6;14E1) are each associated with a Minute phenotype. Moreover, we could rescue these phenotypes, as well as the Minute phenotype associated with

Table 4**CRP gene haploinsufficiency**

Symbol	CRP gene ^a Location	Genetic analysis ^b			Is the CRP gene a <i>Minute</i> ? ^f	No. of candidate <i>Minute</i> genes ^g	<i>Minute</i> synonym ^h	Reference ⁱ
		CRP gene mutations ^{c, d}	Deletions removing a single CRP gene ^{c, e}	Other evidence				
X chromosome								
<i>RpL36</i>	1B12		M		Yes		<i>M(1)1B</i>	[53]
<i>RpL22</i>	1C4	+	+		No			
<i>sta/RpSA</i>	2B1	+	+		No			
<i>RpL35</i>	5A11			Lies in a gap in deletion coverage. 5A6-13 aneuploids were Minute (Merriam et al. [55])	Likely	33	<i>M(1)5A</i>	
<i>RpL7A</i>	6B1		M		Likely	2	<i>M(1)5D6A</i>	
<i>RpL17</i>	6C10			Lies in a gap in deletion coverage	Likely	12	New?	
<i>RpS6</i>	7C2	M	M		Yes		<i>M(1)7BC</i> and <i>M(1)7C</i>	[2]
<i>RpS14a</i>	7C6-7			A deletion that removes <i>RpS14a</i> and <i>RpS14b</i> is not Minute [95]	No			
<i>RpS14b</i>	7C8			A deletion that removes <i>RpS14a</i> and <i>RpS14b</i> is not Minute [95]	No			
<i>RpS28b</i>	8E7			Lies in a gap in deletion coverage. A <i>Minute</i> mutation was mapped to 8D8-9A2 [153]	Likely	17	<i>M(1)8F</i>	
<i>RpS15Aa</i>	11E11-12		M		Likely	8	<i>M(1)11F</i>	
<i>RpL37a</i>	13B1			Lies in a gap in deletion coverage. 12F6-13B6 aneuploids were Minute (Merriam et al. [55])	Likely	2	<i>M(1)13A</i>	
<i>RpS19a</i>	14F4		M		Likely	14	<i>M(1)14E</i>	
<i>RpS5a</i>	15E5-7	M		Lies in a gap in deletion coverage	Yes		<i>M(1)15D</i>	[44]
<i>RpS10b</i>	18D3	M	M		Yes		<i>M(1)18C</i>	
Chromosome arm 2L								
<i>RpLP1</i>	21C2	M	M		Yes		<i>M(2)21C</i>	[50]
<i>oho23B/RpS21</i>	23B6	M	M		Yes		<i>M(2)23B</i>	[49]
<i>RpL40</i>	24E1		M		Likely	4	<i>M(2)24D</i>	
<i>RpL27A</i>	24F3	M	M		Yes		<i>M(2)24F</i>	
<i>RpL37A</i>	25C4		M	The interval between flanking non-Minute deletions contains only <i>RpL37A</i>	Yes		<i>M(2)25C</i>	
<i>RpL36A</i>	28D3		M		Likely	5	<i>M(2)28DE*</i>	
<i>RpS13</i>	29B2	M	M		Yes		<i>M(2)29B*</i>	[45]
<i>sop/RpS2</i>	30E1	+	+		No			
<i>RpL13</i>	30F3		M	The interval between flanking non-Minute deletions contains only <i>RpL13</i>	Yes		<i>M(2)31A</i>	
<i>RpL7</i>	31B1		M		Likely	2	<i>M(2)31A</i>	
<i>RpS27A</i>	31E1		M		Likely	6	<i>M(2)31E*</i>	

Table 4 (Continued)

CRP gene haploinsufficiency							
<i>RpL9</i>	32C1	M	M		Yes		<i>M(2)32D</i> [46]
<i>RpL24</i>	34B10		M		Likely	8	<i>M(2)34BC*</i>
<i>RpS26</i>	36F4	M	M		Yes		<i>M(2)36F</i>
<i>RpL30</i>	37B9	+	+		No		
<i>RpL21</i>	40A-B		M		Likely	10	<i>M(2)39F</i>
<i>RpL5</i>	40B	M	M		Yes		<i>M(2)40B*</i> [51]
Chromosome arm 2R							
<i>RpL38</i>	41C-E	M	M		Yes		<i>M(2)41A</i> [51]
<i>RpL31</i>	45F5	M		Lies in a gap in deletion coverage	Yes		<i>M(2)45F*</i>
<i>RpS15Ab</i>	47C1		+		No		
<i>RpS11</i>	48E8-9	M		Lies in a gap in deletion coverage	Yes		<i>M(2)48E*</i>
<i>RpS23</i>	50E4		M		Likely	2	<i>M(2)50E*</i>
<i>RpS15</i>	53C8			<i>M(2)53¹</i> Minute phenotype rescued by duplication of 51F-54A [154] but not by a <i>RpLP2</i> transgene [155]	Likely	16	<i>M(2)53</i>
<i>RpLP2</i>	53C9		M		Likely	5	<i>M(2)53</i>
<i>RpL18A</i>	54C3			Lies in a gap in deletion coverage	Likely	16	New?
<i>RpL11</i>	56D7			Lies in a gap in deletion coverage. 56C-D aneuploids were Minute [55]	Likely	3	<i>M(2)56CD</i>
<i>RpS18</i>	56F11		M		Likely	2	<i>M(2)56F</i>
<i>RpL29</i>	57D8		+		No		
<i>RpS16</i>	58F1			Deletions that remove both <i>RpS16</i> and <i>RpS24</i> are Minute. The <i>Minute</i> mutations <i>M(2)58F¹</i> and <i>RpS24^{SH2053}</i> complement, suggesting <i>RpS16</i> is a <i>Minute</i> gene	Likely	25	<i>M(2)58F</i>
<i>RpS24</i>	58F3	M		Deletions that remove both <i>RpS16</i> and <i>RpS24</i> are Minute	Yes		<i>M(2)58F</i>
<i>RpL23</i>	59B3		+		No		
<i>RpL37b</i>	59C4		+		No		
<i>RpL12</i>	60B7			<i>RpL12</i> and <i>RpL39</i> lie in the same gap in deletion coverage. 60B3-10 aneuploids were Minute [156]	Likely	9	<i>M(2)60B</i>
<i>RpL39</i>	60B7			<i>RpL12</i> and <i>RpL39</i> lie in the same gap in deletion coverage. 60B3-10 aneuploids were Minute [156]	Likely	9	<i>M(2)60B</i>
<i>RpL41</i>	60E5	+	+		No		
<i>RpL19</i>	60E11	M		A deletion that removes <i>RpL19</i> and <i>RpL41</i> is Minute	Yes		<i>M(2)60E</i> [42]
Chromosome arm 3L							
<i>RpL23A</i>	62A10		M		Likely	8	<i>M(3)62A</i>

Table 4 (Continued)**CRP gene haploinsufficiency**

<i>RpL8</i>	62E7			Lies in a gap in deletion coverage containing only <i>RpL8</i> . 62E-63A aneuploids were Minute [54]	Yes		<i>M(3)62F</i>		
<i>RpL28</i>	63B14		M		Likely	10	<i>M(3)63B</i>		
<i>RpL18</i>	65E9		M		Likely	8	<i>M(3)65F</i>		
<i>RpL14</i>	66D8	M		Lies in a gap in deletion coverage	Yes		<i>M(3)66D</i>	[47]	
<i>RpS17</i>	67B5			A deletion removing both <i>RpS17</i> and <i>RpS9</i> is Minute. The unsequenced Minute mutations <i>RpS17</i> ⁴ and <i>RpS17</i> ⁶ complement the Minute mutation <i>RpS9</i> ^{EP3299} , suggesting <i>RpS17</i> is a Minute gene	Likely	13	<i>M(3)67C</i>		
<i>RpS9</i>	67B11	M		A deletion removing both <i>RpS17</i> and <i>RpS9</i> is Minute	Yes		<i>M(3)67C</i>		
<i>RpL10Ab</i>	68E1		+		No				
<i>RpS12</i>	69F5		+	A deletion that removes <i>RpS12</i> and <i>RpS4</i> is Minute	No				
<i>RpS4</i>	69F6			A deletion that removes <i>RpS12</i> and <i>RpS4</i> is Minute	Likely	2	<i>M(3)69E</i>		
<i>RpL26</i>	75E4		+		No				
<i>RpLP0</i>	79B2	+	+		No				
<i>Qm/RpL10</i>	80A		M		Likely	23	<i>M(3)80</i>		
<i>RpL15</i>	80F		M		Likely	= 11	<i>M(3)80F*</i>		
Chromosome arm 3R									
<i>RpL35A</i>	83A4			Lies in a gap in deletion coverage	Likely	3	New?		
<i>RpL13A</i>	83B6-7	M		Lies in a gap in deletion coverage	Yes		<i>M(3)83B*</i>	[52]	
<i>RpL34b</i>	85D15			Lies in a gap in deletion coverage	Likely	3	New?		
<i>RpS29</i>	85E8	M	M		Yes		<i>M(3)85E</i>		
<i>RpS25</i>	86D8			A deletion that removes <i>RpS25</i> and <i>RpL3</i> is Minute	Likely	14	<i>M(3)86D</i>		
<i>RpL3</i>	86D8	+			No				
<i>RpS5b</i>	88D6	+	+		No				
<i>RpL10Aa</i>	88D10		+		No				
<i>RpS20</i>	93A1			A deletion that removes <i>RpS20</i> and <i>RpS30</i> is Minute	Likely	21	<i>M(3)93A*</i>		
<i>RpS30</i>	93A2			A deletion that removes <i>RpS20</i> and <i>RpS30</i> is Minute	Likely	21	<i>M(3)93A*</i>		
<i>RpS3</i>	94E13	M	M		Yes		<i>M(3)95A</i>	[43]	
<i>RpS19b</i>	95C13		+		No				
<i>RpS27</i>	96C8		M		Likely	5	<i>M(3)96C</i>		
<i>RpL27</i>	96E9-10		M		Likely	6	<i>M(3)96CF</i>		
<i>RpL34a</i>	96F10		+		No				
<i>RpS10a</i>	98A14		+		No				
<i>RpL4</i>	98B6		M		Likely	6	<i>M(3)98B*</i>		
<i>RpS8</i>	99C4			Lies in a gap in deletion coverage. 99B aneuploids were Minute [157]	Likely	11	<i>M(3)99B</i>		

Table 4 (Continued)

CRP gene haploinsufficiency							
RpS28a	99D2				No		
		The Minute phenotype of a deletion removing <i>RpS28a</i> and <i>RpL32</i> is rescued by a <i>RpL32</i> transgene [41]					
<i>RpL32</i>	99D3				Yes	<i>M(3)99D</i>	[41]
		The Minute phenotype of a deletion removing <i>RpS28a</i> and <i>RpL32</i> is rescued by a <i>RpL32</i> transgene [41]					
<i>RpS7</i>	99E2				Likely	11	<i>M(3)99E</i>
		Lies in a gap in deletion coverage. 99E-F aneuploids were Minute [157]					
<i>RpL6</i>	100C7				Likely	16	<i>M(3)100CF</i>
		Lies in a gap in deletion coverage. 100C-F aneuploids were Minute [157]					
Chromosome 4							
<i>RpS3A</i>	101F1	M	M		Yes	<i>M(4)101</i>	[48]

^aBold font indicates the member of a duplicate gene pair that is expressed in a small number of tissues and/or at relatively low levels. ^bComplete details are given in Additional data file 4. ^c'M' indicates that mutation or deletion heterozygotes display a Minute bristle phenotype; '+' indicates they are wild type; a blank indicates the absence of appropriate mutations or deletions. ^dMutations mapped molecularly to a single CRP gene. ^eDeletions removing several genes, but only a single CRP gene. ^fJudged according to evidence summarized in previous three columns and presented in detail in Additional data file 4. ^gThe maximum number of genes that could correspond to the *Minute*; defined as the number of genes between the relevant deletion breakpoints minus the number of genes with non-*Minute* mutations. ^h*Minute* synonyms from literature sources (see Additional data files 3 and 4). Asterisks indicate new synonyms assigned in this study. ⁱReference demonstrating definite correspondence between a *Minute* locus and CRP gene. Where no reference is given, the correspondence is shown for the first time in this study.

the *M(1)14C⁸¹⁵⁻²⁹* point mutation [44], using the small tandem duplication *Dp(1;1)FDDP-0024486* (14C4;14D1). The *Minute* region defined by these experiments contains only two genes: *CG4420* and *eIF-2α*. Significantly, flies heterozygous for *P{RS5}eIF-2α^{5-HA-1790}*, an insertion in the 5' untranslated region (UTR) of *eIF-2α* that creates a likely hypomorphic allele, show a discernable, albeit weak Minute phenotype (our observations). This identifies *eIF-2α* as *M(1)14C*. Consistent with this conclusion, flies expressing a dominant-negative *eIF-2α* protein grow slowly and attain a small body size [96], phenotypes that are typical of the Minute syndrome. *eIF-2α* is one of the three subunits that constitute eIF2, a key translation initiation factor that delivers the methionine-loaded initiator tRNA to the ribosome by transiently associating with the small cytoribosomal subunit [97]. Although *eIF-2α* is not a component of cytoribosome complexes isolated by standard biochemical preparations, a reduction in *eIF-2α* gene dosage might still be expected to adversely affect cytoribosomal function and decrease overall rates of protein synthesis by specifically impairing translation initiation.

Interestingly, the gene encoding *eIF-2γ*, another subunit of the eIF2 translation initiation factor, is also haploinsufficient. Transcripts from the *Su(var)3-9* gene are alternatively spliced to produce two different proteins with distinct functions: one protein is the *eIF-2γ* translation factor, the other is responsible for suppression of position effect variegation [98]. Mutations that specifically eliminate the suppressor

protein are homozygous viable and are not associated with Minute phenotypes [99], but deletions of the entire gene are haplolethal in the absence of *P{(ry⁺), 11. 5kb}*, a transgenic construct carrying the complete *Su(var)3-9* genomic region (our observations; Additional data file 6). These data indicate that the regions of the *Su(var)3-9* transcription unit encoding *eIF-2γ* are haplolethal. Moreover, it is possible that this haplolethality actually represents an extreme Minute phenotype associated with the *eIF-2γ*-coding regions; hypomorphic *eIF-2γ* mutants, if isolated, may show less severe Minute phenotypes.

To assess the possibility that other translation factor genes might also be haploinsufficient/*Minute*, we examined the heterozygous loss-of-function phenotypes of 68 translation factor genes we identified from BLAST searches and/or Gene Ontology classification (Additional data file 6). We identified no other cases of haploinsufficiency, though five genes could not be assessed with existing deletion and mutation strains. In contrast to the genes encoding the other two subunits of eIF2, the *eIF-2β* gene is not haploinsufficient.

As mentioned above, we have compared our inventory of *Minute* genes with the *Minute* loci defined and named from previous genetic analyses (Additional data file 3). In so doing, we failed to validate the existence of several *Minute* loci described in the past, namely *M(1)3E* [55], *M(1)4BC* [55], *M(2)21AB* [100-102], *M(2)44C* [55], *M(3)76A* [55], *M(3)82BC* [55] and *M(3)96A* [103]. The existence of some of

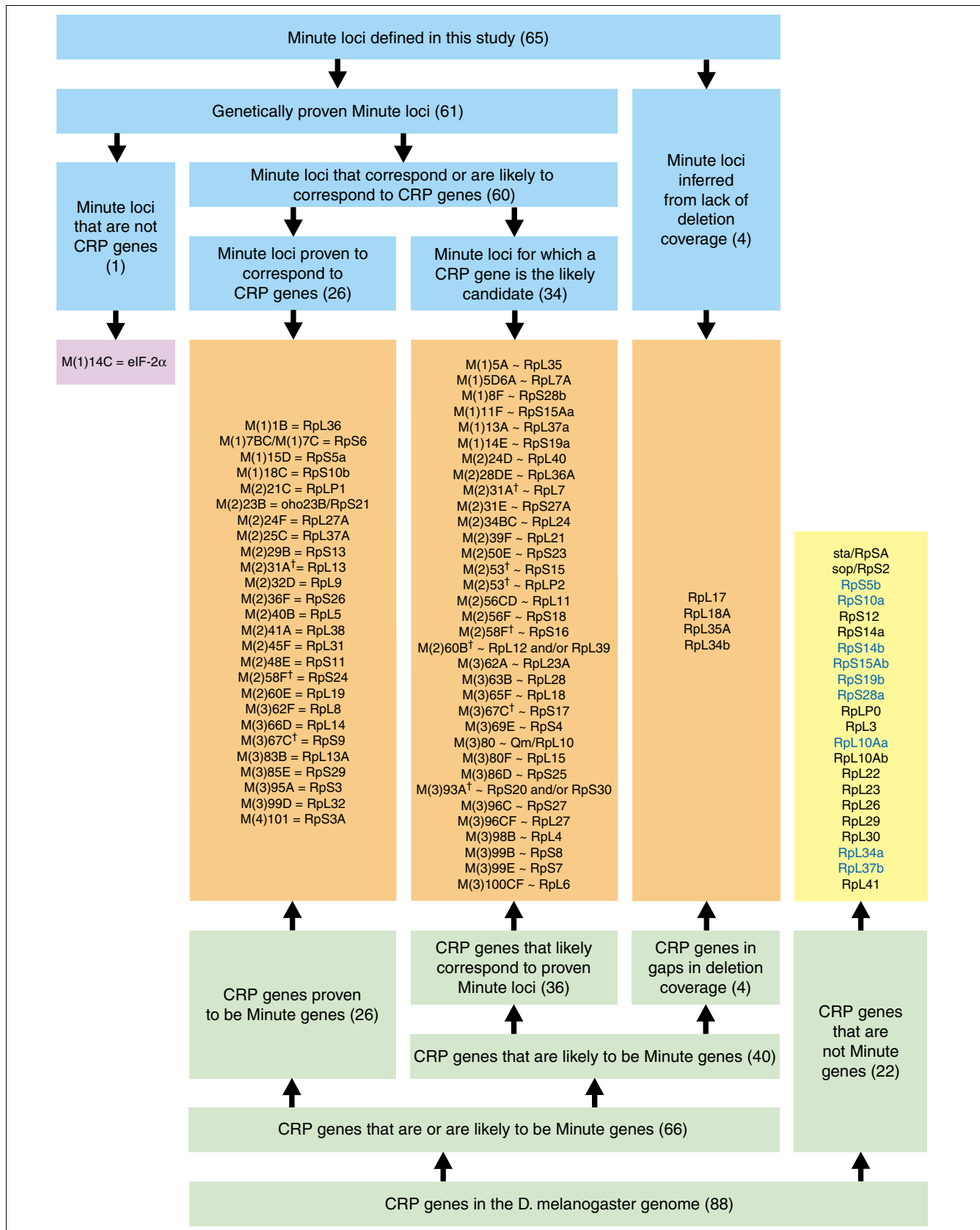


Figure 4 (see legend on next page)

Figure 4 (see previous page)

Summary of *Minute* locus - CRP gene correspondences. This figure shows the relationship between *Minute* loci defined by genetic criteria and CRP genes identified using bioinformatics. '=' indicates definite correspondence, '~' indicates probable correspondence. Daggers mark *Minute* loci that we know or strongly suspect correspond to two CRP genes (as detailed in Table 4 and Additional data file 4).

these loci has been questioned previously and many cases appear to have involved chromosomal aberrations that were unusually complex or point mutations that were mismapped. Our failure to observe a *Minute* phenotype for deletions of *S-adenosylmethionine synthetase* (*Sam-S*), also known as *M(2)21AB*, is consistent with the phenotypic instability of dominant *Sam-S* mutations documented previously [100-102]. This suggests that mutations in *Sam-S* can phenocopy *Minute* mutations under certain conditions, but that *Sam-S* is not a typical *Minute* gene.

In summary, CRP genes are likely to correspond to all but one of the 65 *Minute* loci defined in this study, with the sole exception encoding a translation initiation factor subunit (Figure 4). No MRP or CRP-like genes are unequivocally associated with a *Minute* phenotype, indicating that the *Minute* syndrome is specifically related to the function of the cytoribosomes responsible for the majority of cellular protein translation, rather than the function of specialized cytoribosomal variants. Twenty-five percent of CRP genes are not associated with an obvious haploinsufficient phenotype, clearly reinforcing previous findings that not all CRP genes are *Minute* genes [61,95,104].

Discussion

CRP gene haploinsufficiency and the cytoribosome

When one examines the phenotypes of flies carrying chromosomal deletions, one is struck by the remarkable tolerance of *Drosophila* to aneuploidy: flies heterozygous for deletions of hundreds of kilobases of DNA usually have no obvious dominant phenotypes. For this reason, the haploinsufficiency of single genes is all the more remarkable. It is even more striking that the vast majority of these haploinsufficient genes encode proteins of the cytoribosome and that haploinsufficiency is not apparent for genes encoding components of equally elaborate cellular complexes, such as mitoribosomes or spliceosomes. What accounts for the exquisite dosage sensitivity of CRP genes?

The primary cause of CRP gene haploinsufficiency is reasonably clear: halving the copy number of a CRP gene results in reduced mRNA expression of that CRP [105]. Similarly, depleting CRP mRNAs through antisense- or RNA interference (RNAi)-mediated approaches can also produce *Minute* phenotypes [106,107] (SJM and SJL, unpublished data). As there appear to be no compensatory increases in transcription [105], reducing dosages of CRP genes must result in reduced CRP protein levels in the absence of dramatic changes in mRNA stability or CRP protein turnover. How

then does the reduction in the level of a single CRP result in impaired cytoribosomal function and reduced general protein synthesis, and how is this manifested as the *Minute* phenotype?

One possibility, termed the 'balance hypothesis' [108,109], is linked to the multisubunit nature of the cytoribosome. It posits that an imbalance in the concentrations of CRPs results in the assembly of incomplete and non-functional ribosomal subunits. Indeed, it is known that depletion of individual CRPs in *Saccharomyces cerevisiae* causes inefficient ribosomal subunit assembly and/or function [110,111]. Nevertheless, the balance hypothesis also predicts that overexpression of individual CRPs should cause stoichiometric imbalances and phenotypes similar to those produced by underexpression. This prediction is not upheld in either *S. cerevisiae* [112] or *D. melanogaster* [113]. Consequently, imbalance *per se* cannot account for the haploinsufficiency of CRP genes.

A simpler explanation of CRP haploinsufficiency is that a high concentration of cytoribosomes is required for proper cellular functions and that the cytoribosome population decreases sharply when the level of a single CRP is reduced. Cytoribosomes and their components do appear to be required in unusually high quantities: CRP mRNAs are among the most abundant cellular transcripts both in yeast [114] and in flies [60], and can account for 50% of all RNA polymerase II-mediated transcription [89]. What seems critical, however, to the high-level production of fully formed cytoribosomes is that the concentration of each and every CRP never falls below a minimal level. In other words, cytoribosomal assembly is strictly limited by the availability of the least abundant component. This is probably not just a matter of simple self-assembly kinetics as improperly assembled ribosomal subunits and excess CRPs are actively degraded in yeast [111,115]. Similarly, RNAi-mediated depletion of single CRPs leads to the depletion of other CRPs in flies [64], suggesting the existence of similar degradation processes. Consequently, halving the supply of a single, limiting CRP is expected to halve the number of functional cytoribosomes. This may be tolerated by many cellular processes but will have severe effects wherever high protein synthesis is required, such as bristle formation and oocyte production in flies, or growth of *S. cerevisiae* in rich media [116]. It appears, therefore, that it is the combination of high demand for cytoribosomes and an assembly mechanism that assures that the level of the least abundant CRP determines the final concentration of cytoribosomes that makes CRP genes so exquisitely and specifically dosage sensitive. This perspective also provides a context for understanding the non-additivity of *Minute* mutations [117], where

combinations of *Minute* mutations usually do not have a cumulative effect, but rather result in a phenotype similar to that of the most severe individual *Minute* mutant.

If adequate levels of cytoribosomes depend not so much on precise equimolar CRP concentrations as a minimal concentration of each and every CRP, then we should expect that variation in the expression of different CRPs (above the minimum level) might normally be tolerated *in vivo*. Such variation could be the result of differences in rates of gene transcription, mRNA translation, or mRNA/protein stability. This view provides a framework for understanding the spectrum of haploinsufficient phenotypes associated with CRP genes, which ranges from no obvious phenotypes, through bristle defects and reduced fecundity and viability, to dominant sterility or haplolethality in the most severe cases. That is, the severity of *Minute* phenotypes may be related to the rates at which individual CRPs are normally produced [105,106].

In reality, the explanation of CRP gene haploinsufficiency is probably more complicated than cytoribosome assembly relying simply on minimal CRP concentrations. The exact function, position or stoichiometry of CRPs within the cytoribosome may determine whether its gene is haploinsufficient and the severity of the *Minute* phenotype. For instance, our finding that the gene encoding the eIF-2 α translation initiation factor subunit is a *Minute* could indicate that haploinsufficient CRP genes encode ribosomal components involved specifically in translation initiation. As another example, RpLP1 and RpLP2 are the only CRPs required in two copies per cytoribosome [15] and, consequently, they must be produced at twice the level of all other CRPs. It is perhaps not surprising, therefore, that both *RpLP1* and *RpLP2* are haploinsufficient (Table 4) [50]. It may even be the case that the haploinsufficiency of some CRP genes arises by less conventional mechanisms. For example, the introns of 27 CRP genes host genes for small nucleolar RNAs (snoRNAs) [118-122], a class of non-coding RNAs that guide post-transcriptional modifications of rRNA molecules necessary for the maturation and incorporation of rRNA into ribosomes [123]. The expression of intronic snoRNAs depends upon the expression and processing of mRNAs from the host gene [124]. Consequently, mutations that reduce expression or splicing of CRP transcripts harboring snoRNAs will simultaneously deplete the cell of both a CRP and properly processed rRNA molecules, thereby impairing cytoribosome biogenesis in two different ways. Although 21 of the 27 CRP genes that carry snoRNA genes within their introns are *Minute* or likely *Minute* genes (data not shown), the presence of intronic snoRNA genes cannot be the sole factor determining CRP gene haploinsufficiency. Indeed, we have not found any single factor that clearly determines the degree of dosage sensitivity exhibited by different CRP genes.

Regardless of the exact causes and mechanisms of haploinsufficiency, it is pertinent to ask why the majority of CRPs are expressed so close to the level of sufficiency, such that loss of one gene copy is debilitating, rather than being synthesized in excess? One possibility concerns economics: cytoribosomal synthesis is an incredibly costly affair [89] and excessive CRP production would both be wasteful and monopolize the limited resources of the cell. A second possibility is that CRP levels are normally constrained to guard against inappropriate activation of cell growth, proliferation or apoptosis - processes in which CRPs and cytoribosomes have been postulated to play direct roles [125]. A final and intriguing possibility is that the barely sufficient expression levels of some CRP genes may have evolved as a viral defense mechanism. Cherry *et al.* [64] found that reducing the levels of 64 of the 79 principal CRPs by RNAi inhibits the propagation of *Drosophila* C virus in *Drosophila* adults and cultured cells. Because this virus requires high concentrations of cytoribosomes in its host cell to undergo efficient translation, tightly controlled expression of CRP genes at levels just sufficient for normal growth and development may protect against viral infection and provide a selective advantage. Interestingly, we found a modest correlation between a CRP gene being a *Minute* and it being able to inhibit virus replication in this assay. Clearly, further work will be required to test whether there is truly a relationship between normal CRP gene expression levels and susceptibility to viruses.

Minute mutations attracted the attention of early geneticists because they were isolated so often in *D. melanogaster*. In fact, Schultz said in 1929, "...so many have been found that this mutant type is one of the most frequent in *Drosophila*" [117]. As a considerable number of *Minute* mutants have also been isolated in other *Drosophila* species [60], one might be justified in thinking *Drosophila* are unusually sensitive to CRP gene haploinsufficiency. On the other hand, the phenotypic consequences of CRP gene haploinsufficiency may simply be more noticeable in flies because they include conspicuous changes in external morphology. In fact, 'Minute-ness' may be a widespread phenomenon that is under-recognized because CRP gene haploinsufficiency has different and varied phenotypic consequences in other organisms. Recent research suggests this is the case [116,126-132]. For example, *RPS5* haploinsufficiency disrupts cell division and causes developmental and growth phenotypes in *Arabidopsis* [126]; several CRP genes are haploinsufficient for suppression of nerve sheath tumors in zebrafish [127]; and *RPS19* haploinsufficiency is a causative factor of Diamond-Blackfan anemia in humans [128,130]. In fact, our reliance on the obvious bristle phenotype to distinguish *Minute* from non-*Minute* loci may present a biased assessment of CRP gene haploinsufficiency in the fly: it is quite possible that the 22 CRP genes classified as non-*Minute* in this study are associated with more subtle haploinsufficient phenotypes. How reduced CRP expression gives rise to diverse phenotypes is a mystery that, at least in part, reflects our current ignorance of

the regulation and roles of CRPs in different cell types. This is certainly a topic worthy of more research.

Conclusion

We have assessed an idea that has been discussed for more than three decades; namely, that the haploinsufficient *Minute* loci of *Drosophila* correspond to the genes encoding protein components of ribosomes [2,133]. Our results confirm this idea and add important details. We have shown that *Minute* genes encode proteins of cytoplasmic ribosomes and not mitochondrial ribosomes, and we have defined the subset of CRP genes that are haploinsufficient. While duplicate genes encoding tissue-specific CRPs are not associated with *Minute* phenotypes, it is not otherwise clear what distinguishes the CRP genes that are haploinsufficient from those that are not. We identified a single *Minute* gene encoding a different kind of protein, a cytoplasmic translation initiation factor subunit. This hints that haploinsufficient CRP genes may encode proteins specifically involved in translation initiation, although further work is obviously needed to test this idea.

Minute genes account for the vast majority of the haploinsufficient genes in the *D. melanogaster* genome with effects on fertility and viability strong enough to prevent the recovery of chromosomal deletions in the absence of corresponding duplications. Indeed, there are very few additional genes (for example, *dpp* [134], *Abd-B* [135]) or chromosomal regions (for example, *Tpl* [136], *wupA* [137], *Fs(1)10A* [138]) unequivocally associated with haplolethality or haplosterility. (A few other regions have been reported but not investigated in detail.) The most immediate practical use for our data will be in systematic efforts to maximize genome deletion coverage. Knowing which specific genes are haploinsufficient will make it feasible to flank each one as closely as possible with pairs of deletions, or to delete these genes in the presence of duplications or transgenic rescue constructs. Further improvements in deletion coverage will undoubtedly identify and map the remaining haplolethal or haplosterile loci.

Collectively, our inventories of the RP genes and *Minute* loci of *D. melanogaster* provide a solid foundation for further studies of RPs, ribosomes, and the causes and consequences of haploinsufficiency in flies and other organisms.

Materials and methods

Bioinformatics

RefSeq human RP sequences were obtained from the National Center for Biotechnology Information [139]. The FlyBase BLASTp service [140] was used to identify high scoring hits from among the annotated proteins of *D. melanogaster*; tBLASTn was used when orthologs were not identified by a BLASTp search. The ExPASy proteomics server [141] was used to compute the average pI and molecular weight of the RPs. The percentage identity between

human and *D. melanogaster* RP sequences or between *D. melanogaster* RP pairs was calculated using the NPS@ ClustalW alignment tool at the Pôle Bioinformatique Lyonnais using default parameters [142,143]. K_A/K_S values were estimated using the program package PAML [144]. cDNA clone data were obtained from FlyBase [60].

The identification of CRP gene orthologs and the plotting of their evolutionary emergence (Figure 1) were achieved using a combinatorial approach. First, sequences corresponding to the relevant CRP genes from *D. melanogaster* and *Homo sapiens* were used as queries in BLASTn, BLASTp, tBLASTn and BLAT searches of the genomes of other Drosophilid and insect species using the FlyBase BLAST server [140] and the UCSC Genome Bioinformatics BLAT server [145,146]. High scoring matches were judged to be potential orthologs and were analyzed further using the FlyBase OrthoView tool [87]. Second, the coding sequences (CDS) of relevant CRP genes from *D. melanogaster* and *H. sapiens* were used as queries in BLASTn searches of the GLEANR CDS prediction sets of other Drosophilid species [140]; phylogenetic trees were then generated from the high scoring matches, with the CDS of *S. cerevisiae* CRPs as roots, using the ClustalW tools at EMBL-EBI [147]. Third, NCBI Homologene [148] was searched for any relevant homology calls: *RpLPO-like*, *RpL7-like* and *RpL24-like* were found in HomoloGene clusters 102093, 64526 and 9462, respectively. The results of all these analyses were then compared, with the most parsimonious interpretations being used to annotate the dendrogram shown in Figure 1.

Assessing *Minute* phenotypes and mapping *Minute* mutations

In order to compile a list of all genetically defined *Minute* loci, we first catalogued all the *Minute* loci described in the fly literature [2,54,55,60]. We then inspected deletions for all genomic regions having deletion coverage to confirm or refute the existence of these *Minute* loci and to identify any new *Minute* loci that had previously gone undetected. *Minute* phenotypes were scored primarily by visual inspection of bristle length, although body size, developmental timing, fertility and viability were considered when information was available. Deletion-bearing flies were outcrossed to Oregon-R or Canton-S wild-type strains whenever we could not unambiguously score *Minute* phenotypes in stocks. The phenotypic effects of deleting or disrupting X-linked genes were assessed only in heterozygous females. Finally, the cytological locations of all verified *Minute* loci were correlated with the positions of RP genes to identify candidate genes.

To assess RP gene haploinsufficiency directly, we inspected flies heterozygous for deletions and/or mutations of molecularly identified RP genes. *Minute* phenotypes were scored as described above. For some deletions that had not been characterized molecularly, it was necessary to refine the mapping of breakpoints with complementation tests against

molecularly mapped mutations or with polytene chromosome preparations to determine whether RP genes were deleted. (In a few cases, RP genes were classified as lacking deletion coverage when the only existing deletions were not useful in a practical sense owing to their associated chromosomal rearrangements or extremely large size.) A transposable element insertion was judged to disrupt a RP gene if it failed to complement other mutations in the gene, if we saw a Minute phenotype in the insertion strain, or if the transposon is inserted in the protein-coding region or 5' UTR of the gene based on FlyBase annotations [87] or our own BLAST analyses. (By these criteria, many nearby insertions, intronic insertions, and insertions in 3' UTRs were not used in our analysis.) We included molecularly characterized point mutations in our analyses for the few RP genes where they were available. Molecularly uncharacterized *Minute* point mutations from the Bloomington Stock Center were complementation tested against mutations and deletions known to disrupt or delete specific RP genes.

Fly strains were obtained from the Bloomington, Szeged, Kyoto and Harvard *Drosophila* Stock Centers. Helene Doerflinger and Daniel St Johnston provided *Df(3R)IR16* and *Df(3R)MR22* stocks, and Yuri Sedkov and Alexander Mazo provided *mRpL16^A*, *mRpL16^B* and *mRpL16^C* stocks.

Abbreviations

CDS, coding sequence; CRP, cytoplasmic ribosomal protein; MRP, mitochondrial ribosomal protein; RNAi, RNA interference; RP, ribosomal protein; rRNA, ribosomal RNA; snoRNA, small nucleolar RNA; UTR, untranslated region.

Authors' contributions

SJM, JR, GR, AL and KRC conceived, designed and performed the experiments. SJM, MA, PH, ZY and KRC analyzed the data. JR, GR and KRC contributed fly strains. MA, GM and NK helped with gene classification and nomenclature. TCK and SJL provided funds, lab space and general support. SJM and KRC wrote the paper.

Additional data files

The following additional data are available with the online version of this paper. Additional data file 1 is a table comparing the physical characteristics of *D. melanogaster* and human CRPs, together with their RefSeq accession numbers. Additional data file 2 is a similar table comparing *D. melanogaster* and human MRPs. Additional data file 3 is a table listing all the *Minute* loci in a historical context. Additional data file 4 is a table showing our comprehensive genetic analyses of ribosomal protein gene haploinsufficiency. Additional data file 5 is a table listing CRP gene-*Minute* locus correspondences arranged in alpha-numerical order by RP gene

symbol. Additional data file 6 is a table showing our genetic analyses of translation factor gene haploinsufficiency.

Acknowledgements

We would like to thank Bernard Mechler, Erika Viragh, Janos Szidonya, Chris Smith, Barry Honda, Don Sinclair, Kathleen Fitzpatrick, Mirelle Galloni, Lynn Crosby, Helene Doerflinger, Yuri Sedkov, Kathy Matthews and Phil East for data, assistance, helpful discussions or fly strains, and the Bloomington *Drosophila* Stock Center, the Szeged *Drosophila* Stock Centre, the *Drosophila* Genetic Resource Center at the Kyoto Institute of Technology, and the Exelixis Stock Collection at Harvard Medical School for providing the hundreds of strains used in this paper. Special thanks to Rachel Andrade, Kevin Bogart, Stacey Christensen, Jennifer Deal, Megan Deal and Jill Gresens in Bloomington who generated many of the deletions that helped us map haploinsufficient loci with such precision. We also wish to thank Yuk Sang Chan and Terri Morley at the University of Cambridge for help with cytology and fly crosses, and Torill Rolfsen at the University of Oslo for the SEM pictures. SJM and SJL were funded by Cancer Research UK during this project. KRC and TCK were supported by a National Center for Research Resources grant (RR014106) and the Indiana Genomics Initiative. JR and MA were supported by an MRC Programme Grant (G8225539) and a European Union grant (QLRT-1999-30915). GR was supported by grants from the Deutsche Forschungsgemeinschaft (DFG Re911/3-2 and Re911/5-3) and the European Union (QLRI-CT-2000-00915 and LSHG-CT 2000-503433). GM was supported by National Human Genome Research Institute grant HG000739. NK was funded by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology and the Japan Society for the Promotion of Science. PH and ZY were supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada.

References

- Kay MA, Jacobs-Lorena M: **Developmental genetics of ribosome synthesis in *Drosophila***. *Trends Genet* 1987, **3**:347-351.
- Lambertsson A: **The *Minute* genes in *Drosophila* and their molecular functions**. *Adv Genet* 1998, **38**:69-134.
- Wool IG: **The structure and function of eukaryotic ribosomes**. *Annu Rev Biochem* 1979, **48**:719-754.
- Harris EH, Boynton JE, Gillham NW: **Chloroplast ribosomes and protein synthesis**. *Microbiol Rev* 1994, **58**:700-754.
- Wool IG, Chan YL, Gluck A: **Structure and evolution of mammalian ribosomal proteins**. *Biochem Cell Biol* 1995, **73**:933-947.
- O'Brien TW: **Properties of human mitochondrial ribosomes**. *IUBMB Life* 2003, **55**:505-513.
- Brodersen DE, Nissen P: **The social life of ribosomal proteins**. *FEBS J* 2005, **272**:2098-2108.
- Wool IG: **Extraribosomal functions of ribosomal proteins**. *Trends Biochem Sci* 1996, **21**:164-165.
- Draptchinskaia N, Gustavsson P, Andersson B, Pettersson M, Willig TN, Dianzani I, Ball S, Tchernia G, Klar J, Mattsson H, et al.: **The gene encoding ribosomal protein S19 is mutated in Diamond-Blackfan anaemia**. *Nat Genet* 1999, **21**:169-175.
- Fisher EM, Beer-Romero P, Brown LG, Ridley A, McNeil JA, Lawrence JB, Willard HF, Bieber FR, Page DC: **Homologous ribosomal protein genes on the human X and Y chromosomes: escape from X inactivation and possible implications for Turner syndrome**. *Cell* 1990, **63**:1205-1218.
- O'Brien TW, O'Brien BJ, Norman RA: **Nuclear MRP genes and mitochondrial disease**. *Gene* 2005, **354**:147-151.
- Ruggero D, Pandolfi PP: **Does the ribosome translate cancer?** *Nat Rev Cancer* 2003, **3**:179-192.
- Veuthey AL, Bittar G: **Phylogenetic relationships of fungi, plants, and animalia inferred from homologous comparison of ribosomal proteins**. *J Mol Evol* 1998, **47**:81-92.
- Nakao A, Yoshihama M, Kenmochi N: **RPG: the Ribosomal Protein Gene database**. *Nucleic Acids Res* 2004, **32**:D168-170.
- Wahl MC, Moller W: **Structure and function of the acidic ribosomal stalk proteins**. *Curr Protein Pept Sci* 2002, **3**:93-106.
- Koc EC, Burkhart W, Blackburn K, Moseley A, Spremulli LL: **The small subunit of the mammalian mitochondrial ribosome. Identification of the full complement of ribosomal proteins present**. *J Biol Chem* 2001, **276**:19363-19374.
- Koc EC, Burkhart W, Blackburn K, Moyer MB, Schlatter DM, Moseley

- A, Spremulli LL: **The large subunit of the mammalian mitochondrial ribosome. Analysis of the complement of ribosomal proteins present.** *J Biol Chem* 2001, **276**:43958-43969.
18. **HUGO Gene Nomenclature Committee** [<http://www.genenames.org/>]
 19. Lambertsson AG, Rasmuson SB, Bloom GD: **The ribosomal proteins of *Drosophila melanogaster*. I. Characterization in polyacrylamide gel of proteins from larval, adult, and ammonium chloride-treated ribosomes.** *Mol Gen Genet* 1970, **108**:349-357.
 20. Lambertsson AG: **The ribosomal proteins of *Drosophila melanogaster*. II. Comparison of protein patterns of ribosomes from larvae, pupae and adult flies by two-dimensional polyacrylamide gel electrophoresis.** *Mol Gen Genet* 1972, **118**:215-222.
 21. Berger E: **The ribosomes of *Drosophila*. I. Subunit and protein composition.** *Mol Gen Genet* 1974, **128**:1-9.
 22. Berger EM, Weber L: **The ribosomes of *Drosophila*. II. Studies on intraspecific variation.** *Genetics* 1974, **78**:1173-1183.
 23. Lambertsson AG: **The ribosomal proteins of *Drosophila melanogaster*. 3. Further studies on the ribosomal protein composition during development.** *Mol Gen Genet* 1974, **128**:241-247.
 24. Lambertsson AG: **The ribosomal proteins of *Drosophila melanogaster*. IV. Characterization by two-dimensional gel electrophoresis of the ribosomal proteins from nine postembryonic developmental stages.** *Mol Gen Genet* 1975, **139**:133-144.
 25. Vaslet C, Berger E: **The ribosomes of *Drosophila*. IV. Electrophoretic identify among ribosomal subunit proteins from wild type and mutant *D. melanogaster* and *D. simulans*.** *Mol Gen Genet* 1976, **147**:189-194.
 26. Weber L, Berger E, Vaslet C, Yedvobnick B: **The ribosomes of *Drosophila*. III. RNA and protein homology between *D. melanogaster* and *D. virilis*.** *Genetics* 1976, **84**:573-585.
 27. Berger E: **The ribosomes of *Drosophila*. Normal and defective ribosome biosynthesis in *Drosophila* cell cultures.** *Mol Gen Genet* 1977, **155**:35-40.
 28. Fekete E, Lambertsson A: **Imaginal disc ribosomal proteins of *D. melanogaster*.** *Mol Gen Genet* 1978, **159**:85-87.
 29. Chooi WY, Sabatini LM, Macklin M, Fraser D: **Group fractionation and determination of the number of ribosomal subunit proteins from *Drosophila melanogaster* embryos.** *Biochemistry* 1980, **19**:1425-1433.
 30. Chooi WY: **A comparison of the ribosomal proteins of *Drosophila* ovary, adult, and embryo.** *Mol Gen Genet* 1981, **184**:342-346.
 31. Kay MA, Jacobs-Lorena M: **Selective translational regulation of ribosomal protein gene expression during early development of *Drosophila melanogaster*.** *Mol Cell Biol* 1985, **5**:3583-3592.
 32. Chooi WY: **Purification of *Drosophila* ribosomal proteins. Isolation of proteins S8, S13, S14, S16, S19, S20/L24, S22/L26, S24, S25/S27, S26, S29, L4, L10/L11, L12, L13, L16, L18, L19, L27, I, 7/8, 9, and 11.** *Biochemistry* 1980, **19**:3469-3476.
 33. Chooi WY, Macklin MD, Leiby KR, Hong TH, Scofield SR, Sabatini LM, Burns DK: **Purification of *Drosophila* acidic ribosomal proteins.** *Eur J Biochem* 1982, **127**:199-205.
 34. Alonso J, Santaren JF: **Characterization of the *Drosophila melanogaster* ribosomal proteome.** *J Proteome Res* 2006, **5**:2025-2032.
 35. Planta RJ, Mager WH: **The list of cytoplasmic ribosomal proteins of *Saccharomyces cerevisiae*.** *Yeast* 1998, **14**:471-477.
 36. Barakat A, Szick-Miranda K, Chang IF, Guyot R, Blanc C, Cooke R, Delsenay M, Bailey-Serres J: **The organization of cytoplasmic ribosomal protein genes in the *Arabidopsis* genome.** *Plant Physiol* 2001, **127**:398-415.
 37. Kenmochi N, Kawaguchi T, Rozen S, Davis E, Goodman N, Hudson TJ, Tanaka T, Page DC: **A map of 75 human ribosomal protein genes.** *Genome Res* 1998, **8**:509-523.
 38. Kenmochi N, Suzuki T, Uechi T, Magoori M, Kuniba M, Higa S, Watanabe K, Tanaka T: **The human mitochondrial ribosomal protein genes: mapping of 54 genes to the chromosomes and implications for human disorders.** *Genomics* 2001, **77**:65-70.
 39. Uechi T, Tanaka T, Kenmochi N: **A complete map of the human ribosomal protein genes: assignment of 80 genes to the cytogenetic map and implications for human disorders.** *Genomics* 2001, **72**:223-230.
 40. Yoshihama M, Uechi T, Asakawa S, Kawasaki K, Kato S, Higa S, Maeda N, Minoshima S, Tanaka T, Shimizu N, Kenmochi N: **The human ribosomal protein genes: sequencing and comparative analysis of 73 genes.** *Genome Res* 2002, **12**:379-390.
 41. Kongsuwan K, Yu Q, Vincent A, Frisardi MC, Rosbash M, Lengyel JA, Merriam J: ***Drosophila Minute* gene encodes a ribosomal protein.** *Nature* 1985, **317**:555-558.
 42. Hart K, Klein T, Wilcox M: ***A Minute* encoding a ribosomal protein enhances wing morphogenesis mutants.** *Mech Dev* 1993, **43**:101-110.
 43. Andersson S, Saeboe-Larssen S, Lambertsson A, Merriam J, Jacobs-Lorena M: ***A Drosophila* third chromosome *Minute* locus encodes a ribosomal protein.** *Genetics* 1994, **137**:513-520.
 44. McKim KS, Dahmus JB, Hawley RS: **Cloning of the *Drosophila melanogaster* meiotic recombination gene *mei-218*: a genetic and molecular analysis of interval 15E.** *Genetics* 1996, **144**:215-228.
 45. Saeboe-Larssen S, Lambertsson A: **A novel *Drosophila Minute* locus encodes ribosomal protein S13.** *Genetics* 1996, **143**:877-885.
 46. Schmidt A, Hollmann M, Schafer U: **A newly identified *Minute* locus, *M(2)32D*, encodes the ribosomal protein L9 in *Drosophila melanogaster*.** *Mol Gen Genet* 1996, **251**:381-387.
 47. Saeboe-Larssen S, Urbanczyk Mohebi B, Lambertsson A: **The *Drosophila* ribosomal protein L14-encoding gene, identified by a novel *Minute* mutation in a dense cluster of previously undescribed genes in cytogenetic region 66D.** *Mol Gen Genet* 1997, **255**:141-151.
 48. van Beest M, Mortin M, Clevers H: ***Drosophila Rps3a*, a novel *Minute* gene situated between the segment polarity genes *cubitus interruptus* and *dTCF*.** *Nucleic Acids Res* 1998, **26**:4471-4475.
 49. Torok I, Herrmann-Horle D, Kiss I, Tick G, Speer G, Schmitt R, Mechler BM: **Down-regulation of *RpS21*, a putative translation initiation factor interacting with P40, produces viable *Minute* imagos and larval lethality with overgrown hematopoietic organs and imaginal discs.** *Mol Cell Biol* 1999, **19**:2308-2321.
 50. Fauvarque MO, Laurenti P, Boivin A, Bloyer S, Griffin-Shea R, Bourbon HM, Dura JM: **Dominant modifiers of the *polyhomeotic* extra-sex-combs phenotype induced by marked P element insertional mutagenesis in *Drosophila*.** *Genet Res* 2001, **78**:137-148.
 51. Marygold SJ, Coelho CM, Leever SJ: **Genetic analysis of *RpL38* and *RpL5*, two *Minute* genes located in the centric heterochromatin of chromosome 2 of *Drosophila melanogaster*.** *Genetics* 2005, **169**:683-695.
 52. Alexander SJ, Woodling NS, Yedvobnick B: **Insertional inactivation of the *L13a* ribosomal protein gene of *Drosophila melanogaster* identifies a new *Minute* locus.** *Gene* 2006, **368**:46-52.
 53. Tyler DM, Li W, Zhuo N, Pellock B, Baker NE: **Genes affecting cell competition in *Drosophila*.** *Genetics* 2007, **175**:643-657.
 54. Lindsley DL, Sandler L, Baker BS, Carpenter AT, Denell RE, Hall JC, Jacobs PA, Miklos GL, Davis BK, Gethmann RC, et al.: **Segmental aneuploidy and the genetic gross structure of the *Drosophila* genome.** *Genetics* 1972, **71**:157-184.
 55. Lindsley DL, Zimm GG: *The Genome of Drosophila melanogaster* San Diego, California: Academic Press; 1992.
 56. Stewart B, Merriam JR: **Segmental aneuploidy of the X-chromosome.** *Drosophila Information Service* 1973, **50**:167-170.
 57. **The Bloomington *Drosophila* Stock Center: General Information About the Bloomington Deficiency Kit** [<http://flystocks.bio.indiana.edu/Browse/df-dp/dfkit-info.htm>]
 58. McConkey EH, Bielka H, Gordon J, Lastick SM, Lin A, Ogata K, Reboud JP, Traugh JA, Traut RR, Warner JR, et al.: **Proposed uniform nomenclature for mammalian ribosomal proteins.** *Mol Gen Genet* 1979, **169**:1-6.
 59. Wool IG, Chan YL, Gluck A, Suzuki K: **The primary structure of rat ribosomal proteins P0, P1, and P2 and a proposal for a uniform nomenclature for mammalian and yeast ribosomal proteins.** *Biochimie* 1991, **73**:861-870.
 60. **FlyBase: A Database of *Drosophila* Genes and Genomes** [<http://flybase.org/>]
 61. Cramton SE, Laski FA: ***string of pearls* encodes *Drosophila* ribosomal protein S2, has *Minute*-like characteristics, and is required during oogenesis.** *Genetics* 1994, **137**:1039-1048.
 62. Galloni M, Edgar BA: **Cell-autonomous and non-autonomous growth-defective mutants of *Drosophila melanogaster*.** *Development* 1999, **126**:2365-2375.
 63. Galloni M: **Bonsai, a ribosomal protein S15 homolog, involved in gut mitochondrial activity and systemic growth.** *Dev Biol* 2003, **264**:482-494.
 64. Cherry S, Doukas T, Armknecht S, Whelan S, Wang H, Sarnow P,

- Perrimon N: **Genome-wide RNAi screen reveals a specific sensitivity of IRES-containing RNA viruses to host translation inhibition.** *Genes Dev* 2005, **19**:445-452.
65. Koyama Y, Katagiri S, Hanai S, Uchida K, Miwa M: **Poly(ADP-ribose) polymerase interacts with novel *Drosophila* ribosomal proteins, L22 and L23a, with unique histone-like amino-terminal extensions.** *Gene* 1999, **226**:339-345.
 66. Lee HS, Simon JA, Lis JT: **Structure and expression of ubiquitin genes of *Drosophila melanogaster*.** *Mol Cell Biol* 1988, **8**:4727-4735.
 67. Cabrera y, Poch HL, Arribas C, Izquierdo M: **Sequence of a *Drosophila* cDNA encoding a ubiquitin gene fusion to a 52-aa ribosomal protein tail.** *Nucleic Acids Res* 1990, **18**:3994.
 68. Cabrera HL, Barrio R, Arribas C: **Structure and expression of the *Drosophila* ubiquitin-52-amino-acid fusion-protein gene.** *Biochem J* 1992, **286**:281-288.
 69. Barrio R, del Arco A, Cabrera HL, Arribas C: **Structure and expression of the *Drosophila* ubiquitin-80-amino-acid fusion-protein gene.** *Biochem J* 1994, **302**:237-244.
 70. Vaslet CA, O'Connell P, Izquierdo M, Rosbash M: **Isolation and mapping of a cloned ribosomal protein gene of *Drosophila melanogaster*.** *Nature* 1980, **285**:674-676.
 71. Burns DK, Stark BC, Macklin MD, Chooi WY: **Isolation and characterization of cloned DNA sequences containing ribosomal protein genes of *Drosophila melanogaster*.** *Mol Cell Biol* 1984, **4**:2643-2652.
 72. Qian S, Zhang JY, Kay MA, Jacobs-Lorena M: **Structural analysis of the *Drosophila* *rpA1* gene, a member of the eucaryotic 'A' type ribosomal protein family.** *Nucleic Acids Res* 1987, **15**:987-1003.
 73. Kay MA, Zhang JY, Jacobs-Lorena M: **Identification and germline transformation of the ribosomal protein *rp21* gene of *Drosophila*: complementation analysis with the *Minute QIII* locus reveals nonidentity.** *Mol Gen Genet* 1988, **213**:354-358.
 74. Uechi T, Maeda N, Tanaka T, Kenmochi N: **Functional second genes generated by retrotransposition of the X-linked ribosomal protein genes.** *Nucleic Acids Res* 2002, **30**:5369-5375.
 75. Zhang Z, Harrison P, Gerstein M: **Identification and analysis of over 2000 ribosomal protein pseudogenes in the human genome.** *Genome Res* 2002, **12**:1466-1482.
 76. Brown SJ, Rhoads DD, Stewart MJ, Van Slyke B, Chen IT, Johnson TK, Denell RE, Roufa DJ: **Ribosomal protein S14 is encoded by a pair of highly conserved, adjacent genes on the X chromosome of *Drosophila melanogaster*.** *Mol Cell Biol* 1988, **8**:4314-4321.
 77. Harrison PM, Milburn D, Zhang Z, Bertone P, Gerstein M: **Identification of pseudogenes in the *Drosophila melanogaster* genome.** *Nucleic Acids Res* 2003, **31**:1033-1037.
 78. Parisi M, Nuttall R, Naiman D, Bouffard G, Malley J, Andrews J, Eastman S, Oliver B: **Paucity of genes on the *Drosophila* X chromosome showing male-biased expression.** *Science* 2003, **299**:697-700.
 79. Betran E, Thornton K, Long M: **Retroposed new genes out of the X in *Drosophila*.** *Genome Res* 2002, **12**:1854-1859.
 80. Lifschytz E, Lindsley DL: **The role of X-chromosome inactivation during spermatogenesis.** *Proc Natl Acad Sci USA* 1972, **69**:182-186.
 81. Shigenobu S, Kitadate Y, Noda C, Kobayashi S: **Molecular characterization of embryonic gonads by gene expression profiling in *Drosophila melanogaster*.** *Proc Natl Acad Sci USA* 2006, **103**:13728-13733.
 82. Kai T, Williams D, Spradling AC: **The expression profile of purified *Drosophila* germline stem cells.** *Dev Biol* 2005, **283**:486-502.
 83. Saveanu C, Bienvenu D, Namane A, Gleizes PE, Gas N, Jacquier A, Fromont-Racine M: **Nog2p, a putative GTPase associated with pre-60S subunits and required for late 60S maturation steps.** *EMBO J* 2001, **20**:6475-6484.
 84. Zhao W, Bidwai AP, Glover CV: **Interaction of casein kinase II with ribosomal protein L22 of *Drosophila melanogaster*.** *Biochem Biophys Res Commun* 2002, **298**:60-66.
 85. Shigenobu S, Arita K, Kitadate Y, Noda C, Kobayashi S: **Isolation of germline cells from *Drosophila* embryos by flow cytometry.** *Dev Growth Differentiation* 2006, **48**:49-57.
 86. Stewart MJ, Denell R: **The *Drosophila* ribosomal protein S6 gene includes a 3' triplication that arose by unequal crossing-over.** *Mol Biol Evol* 1993, **10**:1041-1047.
 87. **FlyBase: GBrowse - *D. melanogaster*** [<http://flybase.org/cgi-bin/gbrowse/dmel/>]
 88. Elgin SC, Grewal SI: **Heterochromatin: silence is golden.** *Curr Biol* 2003, **13**:R895-898.
 89. Warner JR: **The economics of ribosome biosynthesis in yeast.** *Trends Biochem Sci* 1999, **24**:437-440.
 90. Ryder E, Blows F, Ashburner M, Bautista-Llacer R, Coulson D, Drummond J, Webster J, Gubb D, Gunton N, Johnson G, et al.: **The DrosDel collection: a set of P-element insertions for generating custom chromosomal aberrations in *Drosophila melanogaster*.** *Genetics* 2004, **167**:797-813.
 91. Ryder E, Ashburner M, Bautista-Llacer R, Drummond J, Webster J, Johnson G, Morley T, Chan YS, Blows F, Coulson D, et al.: **The DrosDel Deletion Collection: A *Drosophila* Genomewide Chromosomal Deficiency Resource.** *Genetics* 2007, **177**:615-629.
 92. Parks AL, Cook KR, Belvin M, Dompe NA, Fawcett R, Huppert K, Tan LR, Winter CG, Bogart KP, Deal JE, et al.: **Systematic generation of high-resolution deletion coverage of the *Drosophila melanogaster* genome.** *Nat Genet* 2004, **36**:288-292.
 93. Bellen HJ, Levis RW, Liao G, He Y, Carlson JW, Tsang G, Evans-Holm M, Hiesinger PR, Schulze KL, Rubin GM, et al.: **The BDGP gene disruption project: single transposon insertions associated with 40% of *Drosophila* genes.** *Genetics* 2004, **167**:761-781.
 94. Thibault ST, Singer MA, Miyazaki WY, Milash B, Dompe NA, Singh CM, Buchholz R, Demsky M, Fawcett R, Francis-Lang HL, et al.: **A complementary transposon tool kit for *Drosophila melanogaster* using P and piggyBac.** *Nat Genet* 2004, **36**:283-287.
 95. Dorer DR, Anane-Firempong A, Christensen AC: **Ribosomal protein S14 is not responsible for the Minute phenotype associated with the *M(1)7C* locus in *Drosophila melanogaster*.** *Mol Gen Genet* 1991, **230**:8-11.
 96. Qu S, Perlaky SE, Organ EL, Crawford D, Cavener DR: **Mutations at the Ser50 residue of translation factor eIF-2alpha dominantly affect developmental rate, body weight, and viability of *Drosophila melanogaster*.** *Gene Expression* 1997, **6**:349-360.
 97. Kimball SR: **Eukaryotic initiation factor eIF2.** *Int J Biochem Cell Biol* 1999, **31**:25-29.
 98. Krauss V, Reuter G: **Two genes become one: the genes encoding heterochromatin protein Su(var)3-9 and translation initiation factor subunit eIF-2gamma are joined to a dicistronic unit in holometabolous insects.** *Genetics* 2000, **156**:1157-1167.
 99. Schotta G, Ebert A, Krauss V, Fischer A, Hoffmann J, Rea S, Jenuwein T, Dorn R, Reuter G: **Central role of *Drosophila* SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing.** *EMBO J* 2002, **21**:1121-1131.
 100. Persson K: **A Minute mutant with suppressor effect on the eye-colour gene *zeste* in *Drosophila melanogaster*.** *Hereditas* 1976, **82**:57-62.
 101. Larsson J, Zhang J, Rasmuson-Lestander A: **Mutations in the *Drosophila melanogaster* gene encoding S-adenosylmethionine synthetase suppress position-effect variegation.** *Genetics* 1996, **143**:887-896.
 102. Larsson J, Rasmuson-Lestander A: **Somatic and germline clone analysis in mutants of the S-adenosylmethionine synthetase encoding gene in *Drosophila melanogaster*.** *FEBS Lett* 1998, **427**:119-123.
 103. Gonzalez C, Molina I, Casal J, Ripoll P: **Gross genetic dissection and interaction of the chromosomal region 95E;96F of *Drosophila melanogaster*.** *Genetics* 1989, **123**:371-377.
 104. Frolov MV, Birchler JA: **Mutation in P0, a dual function ribosomal protein/apurinic/apyrimidinic endonuclease, modifies gene expression and position effect variegation in *Drosophila*.** *Genetics* 1998, **150**:1487-1495.
 105. Saeboe-Larssen S, Lyamouri M, Merriam J, Oksvold MP, Lambertsson A: **Ribosomal protein insufficiency and the Minute syndrome in *Drosophila*: a dose-response relationship.** *Genetics* 1998, **148**:1215-1224.
 106. Patel R, Jacobs-Lorena M: **Generation of Minute phenotypes by a transformed antisense ribosomal protein gene.** *Dev Genet* 1992, **13**:256-263.
 107. Enerly E, Larsson J, Lambertsson A: **Silencing the *Drosophila* ribosomal protein L14 gene using targeted RNA interference causes distinct somatic anomalies.** *Gene* 2003, **320**:41-48.
 108. Veitia RA: **Exploring the etiology of haploinsufficiency.** *Bioessays* 2002, **24**:175-184.
 109. Papp B, Pal C, Hurst LD: **Dosage sensitivity and the evolution of gene families in yeast.** *Nature* 2003, **424**:194-197.
 110. Baronas-Lowell DM, Warner JR: **Ribosomal protein L30 is dispensable in the yeast *Saccharomyces cerevisiae*.** *Mol Cell Biol*

- 1990, **10**:5235-5243.
111. Zhao Y, Sohn JH, Warner JR: **Autoregulation in the biosynthesis of ribosomes.** *Mol Cell Biol* 2003, **23**:699-707.
 112. Warner JR, Mitra G, Schwindinger WF, Studeny M, Fried HM: **Saccharomyces cerevisiae coordinates accumulation of yeast ribosomal proteins by modulating mRNA splicing, translational initiation, and protein turnover.** *Mol Cell Biol* 1985, **5**:1512-1521.
 113. Tamate HB, Patel RC, Riedl AE, Jacobs-Lorena M: **Overproduction and translational regulation of rp49 ribosomal protein mRNA in transgenic Drosophila carrying extra copies of the gene.** *Mol Gen Genet* 1990, **221**:171-175.
 114. Velculescu VE, Zhang L, Zhou W, Vogelstein J, Basrai MA, Bassett DE Jr, Hieter P, Vogelstein B, Kinzler KW: **Characterization of the yeast transcriptome.** *Cell* 1997, **88**:243-251.
 115. Tsay YF, Thompson JR, Rotenberg MO, Larkin JC, Woolford JL Jr: **Ribosomal protein synthesis is not regulated at the translational level in Saccharomyces cerevisiae: balanced accumulation of ribosomal proteins L16 and rp59 is mediated by turnover of excess protein.** *Genes Dev* 1988, **2**:664-676.
 116. Deutschbauer AM, Jaramillo DF, Proctor M, Kumm J, Hillenmeyer ME, Davis RW, Nislow C, Giaever G: **Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast.** *Genetics* 2005, **169**:1915-1925.
 117. Schultz J: **The Minute reaction in the development of Drosophila melanogaster.** *Genetics* 1929, **14**:366-419.
 118. Lyamouri M, Enerly E, Kress H, Lambertsson A: **Conservation of gene order, structure and sequence between three closely linked genes in Drosophila melanogaster and Drosophila virilis.** *Gene* 2002, **282**:199-206.
 119. Yuan G, Klambt C, Bachelier JP, Brosius J, Huttenhofer A: **RNomics in Drosophila melanogaster: identification of 66 candidates for novel non-messenger RNAs.** *Nucleic Acids Res* 2003, **31**:2495-2507.
 120. Accardo MC, Giordano E, Riccardo S, Digilio FA, Iazzetti G, Calogero RA, Furia M: **A computational search for box C/D snoRNA genes in the Drosophila melanogaster genome.** *Bioinformatics* 2004, **20**:3293-3301.
 121. Huang ZP, Zhou H, Liang D, Qu LH: **Different expression strategy: multiple intronic gene clusters of box H/ACA snoRNA in Drosophila melanogaster.** *J Mol Biol* 2004, **341**:669-683.
 122. Huang ZP, Zhou H, He HL, Chen CL, Liang D, Qu LH: **Genome-wide analyses of two families of snoRNA genes from Drosophila melanogaster, demonstrating the extensive utilization of introns for coding of snoRNAs.** *Rna* 2005, **11**:1303-1316.
 123. Kiss T: **Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions.** *Cell* 2002, **109**:145-148.
 124. Tycowski KT, Steitz JA: **Non-coding snoRNA host genes in Drosophila: expression strategies for modification guide snoRNAs.** *Eur J Cell Biol* 2001, **80**:119-125.
 125. Rudra D, Warner JR: **What better measure than ribosome synthesis? Genes Dev 2004, **18**:2431-2436.**
 126. Weijers D, Franke-van Dijk M, Vencken RJ, Quint A, Hooykaas P, Offringa R: **An Arabidopsis Minute-like phenotype caused by a semi-dominant mutation in a RIBOSOMAL PROTEIN S5 gene.** *Development* 2001, **128**:4289-4299.
 127. Amsterdam A, Sadler KC, Lai K, Farrington S, Bronson RT, Lees JA, Hopkins N: **Many ribosomal protein genes are cancer genes in zebrafish.** *PLoS Biol* 2004, **2**:E139.
 128. Gazda HT, Zhong R, Long L, Niewiadomska E, Lipton JM, Ploszynska A, Zaucha JM, Vlachos A, Atsidaftos E, Viskochil DH, et al.: **RNA and protein evidence for haplo-insufficiency in Diamond-Blackfan anaemia patients with RPS19 mutations.** *Br J Haematol* 2004, **127**:105-113.
 129. Oliver ER, Saunders TL, Tarle SA, Glaser T: **Ribosomal protein L24 defect in belly spot and tail (Bst), a mouse Minute.** *Development* 2004, **131**:3907-3920.
 130. Leger-Silvestre I, Caffrey JM, Dawaliby R, Alvarez-Arias DA, Gas N, Bertolone SJ, Gleizes PE, Ellis SR: **Specific role for yeast homologs of the Diamond Blackfan anemia-associated Rps19 protein in ribosome synthesis.** *J Biol Chem* 2005, **280**:38177-38185.
 131. Panic L, Tamarut S, Sticker-Jantschkeff M, Barkic M, Solter D, Uzelac M, Grabusic K, Volarevic S: **Ribosomal protein S6 gene haploinsufficiency is associated with activation of a p53-dependent checkpoint during gastrulation.** *Mol Cell Biol* 2006, **26**:8880-8891.
 132. Panic L, Montagne J, Cokaric M, Volarevic S: **S6-haploinsufficiency activates the p53 tumor suppressor.** *Cell Cycle* 2007, **6**:20-24.
 133. Huang SL, Baker BS: **The mutability of the Minute loci of Drosophila melanogaster with ethyl methanesulfonate.** *Mutat Res* 1976, **34**:407-414.
 134. Spencer FA, Hoffmann FM, Gelbart WM: **Decapentaplegic: a gene complex affecting morphogenesis in Drosophila melanogaster.** *Cell* 1982, **28**:451-461.
 135. Karch F, Weiffenbach B, Peifer M, Bender W, Duncan I, Celniker S, Crosby M, Lewis EB: **The abdominal region of the bithorax complex.** *Cell* 1985, **43**:81-96.
 136. Dorer DR, Rudnick JA, Moriyama EN, Christensen AC: **A family of genes clustered at the Triplo-lethal locus of Drosophila melanogaster has an unusual evolutionary history and significant synteny with Anopheles gambiae.** *Genetics* 2003, **165**:613-621.
 137. Prado A, Canal I, Ferrus A: **The haplolethal region at the 16F gene cluster of Drosophila melanogaster: structure and function.** *Genetics* 1999, **151**:163-175.
 138. Lefevre G Jr: **The eccentricity of vermilion deficiencies in Drosophila melanogaster.** *Genetics* 1969, **63**:589-600.
 139. **National Center for Biotechnology Information: Entrez Protein** [<http://www.ncbi.nlm.nih.gov/sites/entrez?db=Protein>]
 140. **FlyBase: BLAST** [<http://flybase.org/blast/>]
 141. **ExpASY Proteomics Server: Compute pI/Mw tool** [http://www.expasy.org/tools/pi_tool.html]
 142. **Pôle Bioinformatique Lyonnais, Network Protein Sequence Analysis: CLUSTALW** [http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_clustalw.html]
 143. Combet C, Blanchet C, Geourjon C, Deleage G: **NPS@: network protein sequence analysis.** *Trends Biochem Sci* 2000, **25**:147-150.
 144. Yang Z: **PAML: a program package for phylogenetic analysis by maximum likelihood.** *Comput Appl Biosci* 1997, **13**:555-556.
 145. Kent WJ: **BLAT - the BLAST-like alignment tool.** *Genome Res* 2002, **12**:656-664.
 146. **UCSC Genome Bioinformatics** [<http://genome.ucsc.edu/>]
 147. **EBI Sequence Analysis Tools: ClustalW** [<http://www.ebi.ac.uk/Tools/clustalw/index.html>]
 148. **National Center for Biotechnology Information: HomoloGene** [<http://www.ncbi.nlm.nih.gov/sites/entrez?db=homologene>]
 149. Powell JR: *Progress and Prospects in Evolutionary Biology: The Drosophila Model* New York: Oxford University Press; 1997.
 150. **FlyBase: Map Conversion Table** [http://flybase.org/static_pages/docs/cytotable3.html]
 151. Hoskins RA, Smith CD, Carlson JW, Carvalho AB, Halpern A, Kaminker JS, Kennedy C, Mungall CJ, Sullivan BA, Sutton GG, et al.: **Heterochromatic sequences in a Drosophila whole-genome shotgun assembly.** *Genome Biol* 2002, **3**:RESEARCH0085.
 152. **FlyBase: BAC In Situ Images** [http://cane.bio.indiana.edu:7062/images/lk/bac_insitu_pic/]
 153. Schalet AP: **The distribution of and complementation relationships between spontaneous X-linked recessive lethal mutations recovered from crossing long-term laboratory stocks of Drosophila melanogaster.** *Mutat Res* 1986, **163**:115-144.
 154. Golic KG, Golic MM: **Engineering the Drosophila genome: chromosome rearrangements by design.** *Genetics* 1996, **144**:1693-1711.
 155. Qian S, Hongo S, Jacobs-Lorena M: **Antisense ribosomal protein gene expression specifically disrupts oogenesis in Drosophila melanogaster.** *Proc Natl Acad Sci USA* 1988, **85**:9601-9605.
 156. Reed B: **The genetic analysis of endoreduplication in Drosophila melanogaster.** In *PhD thesis* University of Cambridge; 1992.
 157. Kongsuwan K, Dellavalle RP, Merriam J: **Deficiency analysis of the tip of chromosome 3R in Drosophila melanogaster.** *Genetics* 1986, **112**:539-550.