Nucleic Acids Research, 1998, Vol. 26, No. 19 **4471–4475**

Drosophila RpS3a, a novel *Minute* gene situated between the segment polarity genes *cubitus interruptus* and *dTCF*

Moniek van Beest, Mark Mortin[1] and Hans Clevers*

Department of Immunology, University Hospital, Utrecht, The Netherlands and [1]Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

Received May 26, 1998; Revised and Accepted August 12, 1998

DDBJ/EMBL/GenBank accession no. AF034971

ABSTRACT

Genetic analysis of the small chromosome 4 of *Drosophila* has been hampered by the virtual lack of recombination. The segment polarity gene *cubitus interruptus* (*ci*) maps to the most intensively studied locus on this chromosome. Up to four complementation groups have been found to be associated with *ci*. We and others have recently characterized a second segment polarity gene, *dTCF* or *pan*, 12 kb upstream of *ci*, in a head-to-head configuration. During the course of these studies we identified a transcription unit in the intergenic region. We report here the cloning of cDNAs from this transcription unit, which encode the *Drosophila* homologue of the human ribosomal protein S3a (RpS3a). The *RpS3a* gene is expressed ubiquitously and throughout development. A *Minute* allele, *M(4)101*, linked tightly to *ci*, was found to harbour an integration of a Doc retroposon in the promotor region of *RpS3a*. Thus, like other *Minute* loci, *M(4)101* encodes a component of the protein synthesis machinery. These data further unravel the complex genetics surrounding the *ci* and *dTCF* loci.

INTRODUCTION

Chromosome 4 of *Drosophila* represents only 3.5% of the total genomic content of the fly. It has an unusual property, in that recombination rarely takes place between homologs. This is most likely the result of the small size of the chromosome, as well as its relatively high content of heterochromatin. The lack of recombination in combination with the presumed paucity of genetic loci has made chromosome 4 difficult to study by genetic means. Nevertheless, a few chromosome 4 loci have been studied in some detail (1,2). One of the regions of chromosome 4 that has received attention contains the *cubitus interruptus* (*ci*) locus, encoding a segment polarity gene. The only known *Minute* locus on chromosome 4 [*M(4)101*] has been tightly linked to the *ci* locus (1).

The multiplicity of recessive and dominant phenotypes associated with *ci* was, until recently, thought to be linked to two different loci (3,4). Locke and Tartoff (5), however, proposed that all these mutants arise from one single complex locus on chromosome 4. These same authors placed all *ci* mutations in three different complementation groups, based on complementation assays using the prototypic ci^D mutant fly. Recently, we and others have shown that the ci^D fly bears a compound mutation affecting both the *ci* gene and a novel segment polarity gene termed *dTCF* or *pangolin*. One of these proposed *ci* complementation groups, *l(4)13*, actually represents *dTCF* (6,7). The *dTCF* gene is positioned upstream of *ci* on the proximal portion of chromosome 4 and is transcribed in the opposite direction (6). A more detailed analysis of this region showed that the transcription start sites of these genes are separated by ~12 kb (8). Orenic *et al.* (4) report the presence of an unidentified transcription unit in the intergenic region between *ci* and *dTCF*, further complicating the genetics of this area.

In the course of analyzing the *dTCF* gene, we isolated cDNA clones representing this transcription unit from an embryonic mixed stage cDNA library. Detailed analysis of this gene revealed it to be the *Drosophila* homolog of the human ribosomal protein S3a. Since Minute phenotypes often result from mutation in structural ribosomal genes, we analysed the status of *Drosophila RpS3a* in $M(4)101^{57g}$ flies and found an insertion of a Doc retroposon in the promotor, likely inactivating the gene. Identification of this novel gene may shed light on the complex genetics of proximal chromosome 4.

MATERIALS AND METHODS

Fly stocks

The fly stocks $M(4)101^{57g}/ci^D$, ci^D/Ey^D and $df(4)M62f/Ey^D$ were kindly provided by R. Holmgren and the Bloomington Stock Center (Bloomington, IN).

Cloning of RpS3a

A genomic clone 5′ of *dTCF* was obtained as described elsewhere (8). A random primed mixed staged embryo cDNA library was kindly provided by B. Hovemann (described in 9). This library was screened at low stringency with a probe derived from the genomic phage clone as indicated in Figure 1 according to standard procedures. Positive clones were subcloned into pBluescript SK and sequenced. Genomic structure was determined by direct sequencing of the original genomic *dTCF* clones. The full-length sequence is available under GenBank accession no. AF034971.

*To whom correspondence should be addressed. Tel: +31 30 250 7674; Fax: +31 30 251 7107; Email: h.clevers@lab.azu.nl

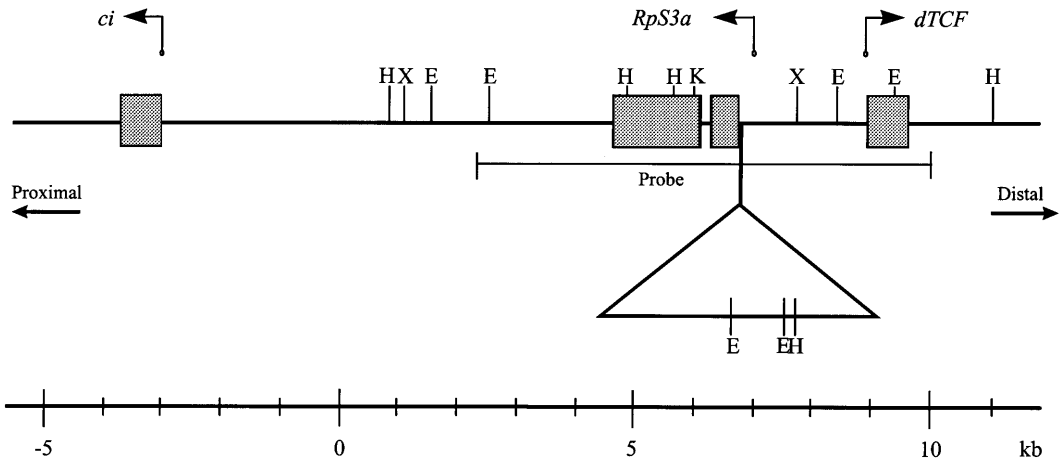

Figure 1. Representation of the genomic area between *ci* and *dTCF*. The top part of the figure represents the proximal part of chromosome 4 between *ci* and *dTCF*. The first exons of *ci* and *dTCF* are depicted, as well as two *RpS3a* exons. The triangle reflects the size and location of the inserted Doc retroposon. The probe used for screening is indicated. The kilobase (kb) numbering is taken from figure 2 of Locke and Tartoff (5). E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; X, *Xho*I.

Northern blot analysis

RNA was isolated according to Chomczynski *et al.* (10), adjusted for whole organisms. Briefly, adult flies, larvae or timed embryos were homogenized in solution D (6.3 M guanidine thiocyanate, 0.04 M sodium citrate, 0.8% Sarcosyl) using a Dounce homogenizer. After this, RNA extraction was performed according to the standard protocol (10). RNA was subjected to electrophoresis and was blotted onto nitrocellulose and probed with the *RpS3a* cDNA clone.

In situ hybridization

Whole mount digoxigenin *in situ* hybridization was performed as described elsewhere (6), using *RpS3a* cDNA as probe.

Southern blot analysis

Genomic DNA was isolated from adult flies according to protocol 47 of Ashburner (11). An aliquot of 1 µg DNA was digested overnight with either *Hin*dIII or *Eco*RI. The digested DNA was subjected to electrophoresis and blotted onto nitrocellulose and the Southern blot subsequently probed with either the cDNA clone or with the genomic clone.

Genomic size-selected plasmid library of *M(4)101^{57g}/ci^{D}*

Southern blot data revealed a mutant *Hin*dIII band on the *M(4)101^{57g}* chromosome of 4.2 kb, hybridizing to the 5′-portion of the *RpS3a* cDNA. Genomic DNA from the *M(4)101^{57g}/ci^{D}* flies was digested with *Hin*dIII overnight. After electrophoresis, DNA of ~4.2 kb was cloned into pBluescript SK. Amongst 2×10^5 clones, ~40 colonies hybridized to the *RpS3a* cDNA. Several of these were isolated and sequenced.

Rapid amplification of cDNA ends (RACE)

RNA was isolated from three clones and cDNA was generated using AMV reverse transcriptase and *RpS3a*-specific primer (VEA, gtctttgcttcgacaatagc; Isogen, The Netherlands). After separation of the reaction products from excess primers by use of microspin columns (Pharmacia Biotech, The Netherlands), cDNAs were extended using dATP and TdT (Pharmacia Biotech). Typically, cDNA samples were incubated for 10 min at 37°C using 1 mM dATP and 1 U TdT in the TdT buffer provided by the manufacturer. Subsequently, 1/50 of the tailed products was subjected to PCR using oligo(dT) and *RpS3a*-specific primer (MVK, cacttttaaccatcgacc; Isogen). The following PCR program was used: two cycles of 30 s at 94°C, 30 s at 42°C, 30 s at 72°C; 25 cycles of 30 s at 94°C, 30 s at 55°C, 30 s at 72°C; a final extension of 7 min at 72°C. PCR products were cloned in vector pGEM T (Promega, Madison, WI) and sequenced.

RESULTS

During analysis of the 12 kb region between *ci* and *dTCF* (6), a mixed stage embryonic cDNA library was screened with a genomic probe located 5′ of *dTCF* (Fig. 1). Two independent cDNA clones were isolated. Sequencing of these clones revealed that they were derived from the same gene. The largest clone of 802 bp was used for further study. The smaller clone contained 699 bp and was completely internal to the larger clone. Database searches revealed that the encoded protein was the orthologue of a ribosomal protein identified in multiple diverse species (Table 1). Comparison with prototypic KRP-A protein from *Aplysia* (11) indicated that the reading frame encoded by our largest cDNA clone contained a translation start site, but lacked approximately five codons at the 3′-end. In order to determine the missing sequence, the genomic clone was directly sequenced. The full open reading frame, combining cDNA and genomic sequences, is depicted in Figure 2. Further sequencing revealed the genomic structure of the *Drosophila RpS3a* gene. It was found that the coding region of the *RpS3a* gene consists of two exons separated by an intron of 101 bp (Figs 1 and 2). The *RpS3a* gene was transcribed in the same direction as the *ci* gene and mapped ~2.5 kb upstream of *dTCF* (Fig. 1).

Table 1. Table of ribosomal protein S3a orthologs

Gene product	Species	Identity (%)	Reference
RpS3a	*Drosophila melanogaster*	–	This report
KRP-A	*Aplysia californica*	69	12
KRP-Y1	*Saccharomyces cerevisiae*	64	12
KRP-Y2/MFT	*Saccharomyces cerevisiae*	63	12,13
hRpS3a	*Homo sapiens*	70	14,15
fte-1	*Rattus norvergicus*	69	16,17
TU-11	*Mus musculus*	70	17,18
Cyc07	*Caranthes roseus*	57	19
C3	*Anopheles gambiae*	74	20

The identity column reflects the percentage identity of each clone with respect to the *Drosophila* amino acid sequence.

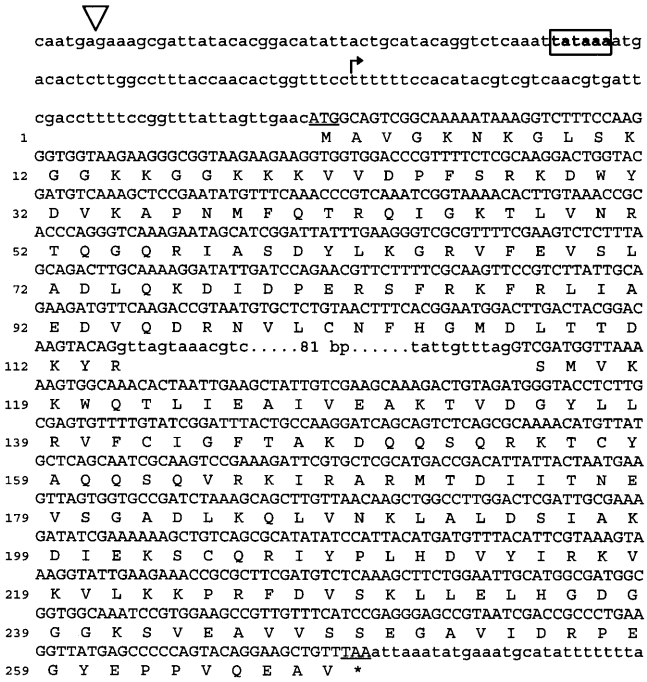

Figure 2. Nucleotide and amino acid sequence of *RpS3a*. The coding sequence is in upper case. The trancription start site is indicated by the hooked arrow. The translation start and stop codons are underlined. The box indicates the TATA box. The location of the Doc retroposon insertion is indicated by the triangle. The full-length sequence of *RpS3a* is also available under GenBank accession no. AF034971.

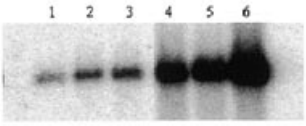

Figure 3. Expression of *Drosophila RpS3a* during development. A northern blot of different stages of development was composed. Lanes 1–3, wild-type embryos of 0–2, 2–4 and 4–8 h; lane 4, mixed stage larvae; lanes 5 and 6, male and female adults.

Expression of *RpS3a* during development of the fly was examined using northern blot analysis. RNA was isolated at different stages of development: embryos of 0–2, 2–4 and 4–8 h, mixed stage larvae and male and female adult flies. The blotted RNA was probed with the complete cDNA clone. Analysis of the hybridizing bands revealed that at every stage an mRNA species of ~1.7 kb could be detected (Fig. 3). The intensity of each band, after correction for the amount of RNA loaded, was comparable. The observed expression in the 0–2 h embryos indicated maternal contribution of the *RpS3a* mRNA. Whole mount *in situ* hybridization revealed high ubiquitous expression of maternally derived *RpS3a* mRNA in early cleavage stage embryos (Fig. 4A). Expression also remained ubiquitous during germband extension (Fig. 4B). In order to distinguish between ubiquitous expression and background staining, the offspring of *df(4)M62f/Ey*^{D} (see below) were tested. Of this population, 25% should be homozygous for *df(4)62f*, which deletes a large part of chromosome 4 containing both *ci* and *dTCF*. Indeed, 25% of the population showed no staining after germband extension (Fig. 4C), confirming the ubiquitous expression pattern of *RpS3a*.

We next sought to determine the *in vivo* function of *RpS3a*. We performed a Southern blot analysis on total genomic *Drosophila* DNA using several restriction enzymes. Probing with the cDNA clone revealed simple hybridization patterns with bands of the expected sizes, indicating that the *Drosophila* genome contained only one *RpS3a* gene (Fig. 5, lanes 4). In the same experiment *df(4)M62f/Ey*^{D} flies were analyzed. The *Ey*^{D} locus maps to the distal region of chromosome 4 and is not expected to show abberations in the pertinent region. It was predicted that *df(4)M62f* deficiency would remove the *RpS3a* gene. Indeed, *df(4)M62f/Ey*^{D} flies showed a 50% reduction in the signal as analyzed by phosphorimaging (not shown), in concordance with the proposed physical map.

Unexpectedly, an ~4.8 kb insert previously proposed to be present on the *ci*^{D} chromosome (5) and located very close to the *dTCF* promotor turned out to be absent from *ci*^{D}*/Ey*^{D} flies but was detected in *M(4)101*^{57g}*/ci*^{D} flies (Fig. 5). This insert, which we tentatively mapped to the M(4)101 Minute chromosome, is indicated in Figure 1. Since mutations in many ribosomal genes cause Minute phenotypes (21,22), we hypothesized that the insertion in M(4)101^{57g} would affect the *RpS3a* gene, which likely encodes a component of the *Drosophila* ribosomal protein complex.

Detailed analysis of the insert by Southern blotting indicated the presence of internal *Hin*dIII and *Eco*RI sites. The insert could be mapped between the *Kpn*I site in exon 2 of the *RpS3a* gene and a *Xho*I site ~1 kb upstream of the *RpS3a* gene. Upon probing with cDNA, a mutant *Hin*dIII fragment of 4.2 kb was observed. This fragment was predicted to cover a region from the *Hin*dIII site in exon 2 to a *Hin*dIII site located in the insert (Fig. 1). In order to determine the exact localization and nature of the insert, we constructed a size-selected plasmid library from *M(4)101*^{57g}*/ci*^{D} genomic DNA restricted with *Hin*dIII. A primary library of 2×10^5 recombinant clones was obtained and screened with the cDNA. Four independent plasmids were sequenced and found to contain identical inserts. Sequence analysis of these clones and comparison with the wild-type genomic sequences revealed the presence of an insert 139 bp upstream of the translation start site of *RpS3a* (Fig. 2).

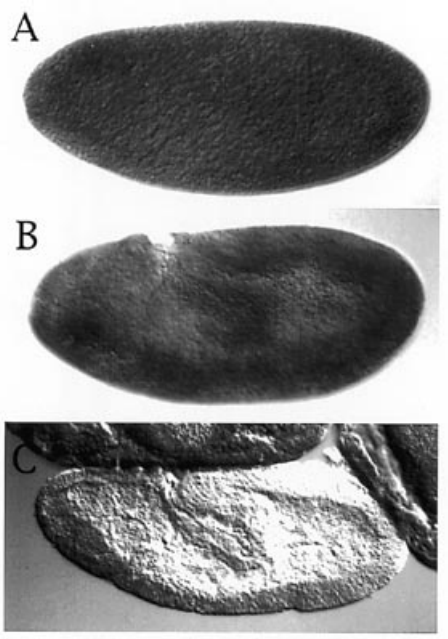

Figure 4. Expression pattern of *RpS3a* in embryos. *In situ* hybridization for *RpS3a* was performed on whole mount embryos. (**A**) A 0–2 h embryo showing the maternal contribution of *RpS3a* mRNA. (**B**) Germband extended wild-type embryo. (**C**) Germband extended *df(4)M62f* homozygous embryo.

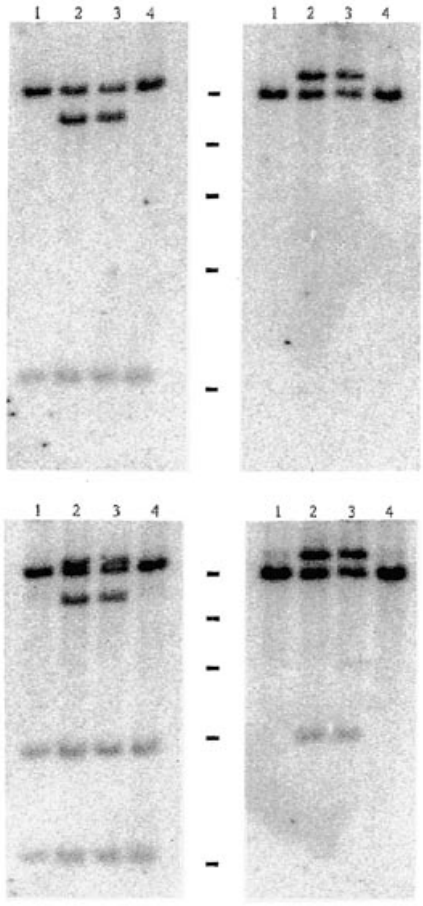

Figure 5. Southern blots of *M(4)101*^{57g} and *ci*^{D} genomic DNA. Genomic DNA was isolated and digested with *Hin*dIII (left blot of each panel) or *Eco*RI (right blot). The top panel was hybridized with *RpS3a* cDNA. The lower panel was hybridized with a 5 kb *Eco*RI fragment of the genomic *dTCF* clone (Fig. 1). Lanes 1, *ci*^{D}/*Ey*^{D} DNA; lanes 2 and 3, *M(4)101*^{57g}/*ci*^{D}; lanes 4, wild-type DNA. The horizontal marks reflect the size of the DNA, with 1 kb for the lowest followed by 2–5 kb above. (**Top**) Lanes 2 and 3 show an extra 4.2 kb band for *Hin*dIII and an extra 5.3 kb band for *Eco*RI. (**Bottom**) Lanes 2 and 3 show extra 4.2 and 5.2 kb bands for *Hin*dIII and extra 2.0 and 5.3 kb bands for *Eco*RI.

Using 5′ RACE, we determined the transcription start site of *RpS3a* (Fig. 2). A putative TATA box is located ~30 bp upstream of this start site. The Doc insert was located 43 bp upstream of the TATA box, presumably within the *RpS3a* promotor. Database searching revealed that the inserted sequence was identical to the 4.8 kb Doc retroposon, at least over the 350 bp that we analyzed. Furthermore, the *Eco*RI and *Hin*dIII sites predicted from Southern blotting coincided with sites in this retroposon (23). The Doc retroposon was originally identified as responsible for a spontaneous mutation in the *white-one* allele (24).

DISCUSSION

In this report, we describe a molecular and genetic characterization of a *Drosophila* gene encoding a homolog of the mammalian ribosomal protein S3a. We originally cloned cDNAs encoding a transcription unit located directly between the two segment polarity genes, *ci* and *dTCF*. Having identified the nature of the encoded protein, we determined the status of the *RpS3a* gene in *M(4)101*^{57g}, a chromosome harbouring a mutation mapping near *ci* and causing a Minute phenotype. We could show the insertion of a Doc retroposon in the predicted promotor region of *RpS3a*, which very likely causes inactivation of the gene. These data add yet another structural ribosomal gene to the list of genes mutated in flies with dominant Minute phenotypes. Furthermore, our report unravels part of the complex and incompletely understood genetics of the pertinent region on proximal chromosome 4.

The Minute phenotype results from a mutation in any of >50 loci scattered throughout the *Drosophila* genome. The mutations are usually dominant and cause a phenotype consisting of short, thin bristles, slow development, reduced viability, rough eyes, small body size and etched tergites. The mutations causing Minute phenotypes are thought to occur in elements of the ribosomal machinery (21,22,25,26). Most cloned ribosomal proteins have been mapped to chromosomal regions near *Minute* loci (27,28). For six ribosomal protein genes this correlation has been confirmed: *M(3)99D* is *RP49* (21); *M(2)60E* is *RPL19* (29); *M(3)95A* is *RPS3* (27); *M(2)32A* is *RPS13* (22); *M(2)32D* is *RPL9* (25); *M(1)15D* is *RPS5* (30). Conversely, mutations in ribosomal proteins do not always cause Minute phenotypes, as evidenced by recessive lethal mutations in the *RpS14* genes (31) which do not display visible phenotypes in heterozygotes. The current study adds another example of a Minute phenotype caused by a mutation in a gene coding for a ribosomal protein.

Based on the high level of sequence conservation, all genes displayed in Table 1 are likely to code for the ribosomal protein S3a. However, with the exception of human S3a (15), the genes were not cloned in a search for ribosomal components. The *Aplysia* ortholog was cloned in a search for genes specifically expressed in the large neurons of this organism (12). In the same study, the yeast orthologs were cloned by homology (12). The

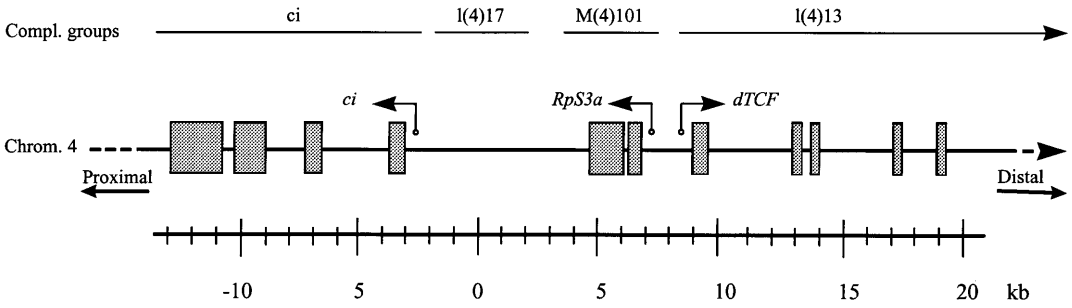

Figure 6. Physical map of *ci*, *RpS3a* and *dTCF*. The proximal region of chromosome 4 is depicted bearing the genes *ci*, *RpS3a* and *dTCF*. Four potential complementation groups, *ci*, *l(4)17*, *M(4)101* and *l(4)13*, can be discriminated. The kilobase (kb) numbering is taken from figure 2 of Locke and Tartoff (5).

mouse ortholog, termed TU-11, was originally identified as a TNF-inducible protein (18), whilst the rat ortholog fte-1 was proposed to be an effector of the v-*fos* oncogene (16). The different settings in which the same gene has apparently been identified may reflect particular limiting requirements for structural ribosomal proteins in individual physiological/developmental processes in the cell.

The starting point of this study was the complex and conflicting genetic and molecular data on the *ci* locus. The confusion is largely caused by application of the compound *ci/dTCF* mutant ci^D as the prototypic complementation partner in genetic crosses. Originally, two complementation groups, *cell* and ci^D, were believed to map to this region (3), but were later found to both affect the *ci* gene (5). The latter authors assign three complementation groups to the *ci* locus and map the *M(4)101* locus directly downstream of the *ci* gene. We and others have since shown that one of the three proposed *ci* complementation groups, *l(4)13*, represents mutations in the independent segment polarity gene *dTCF* (6,7). Moreover, data from Holmgren and colleagues (32) provide evidence that the *l(4)17* complementation group has molecular abberations in upstream regulatory sequences of *ci*. The third complementation group in *ci* affects composition of the *ci* mRNA directly (5,32). Finally, this study now shows the *M(4)101* locus to be upstream of *ci*.

Based on published data from our laboratory and Basler and colleagues (6,7) and on the current study we can now draw a new physical map of this region on proximal chromosome 4, which harbours the three genes *ci*, *RpS3a* and *dTCF* (Fig. 6). Elucidation of the physical map should serve as a framework for the further characterization of other mutations in this region.

ACKNOWLEDGEMENTS

We thank Drs F. Staal and N. Barker for critically reading the manuscript and the members of the Clevers laboratory for helpful discussions. This work was supported by the Netherlands Organization for Scientific Research–Physics Research Netherlands (NWO-SON).

REFERENCES

1 Hochman,B. (1971) *Genetics*, **67**, 235–252.
2 Hochman,B. (1974) *Cold Spring Harbor Symp. Quant. Biol.*, **38**, 581–589.
3 Orenic,T., Chidsey,J. and Holmgren,R. (1987) *Dev. Biol.*, **124**, 50–56.
4 Orenic,T., Slusarski,D., Kroll,K. and Holmgren,R. (1990) *Genes Dev.*, **4**, 1053–1067.
5 Locke,J. and Tartoff,K. (1994) *Mol. Gen. Genet.*, **243**, 234–243.
6 van de Wetering,M., Cavallo,R., Dooijes,D., van Beest,M., van Es,J., Loureiro,J., Ypma,A., Hursh,D., Jones,T., Bejsovec,A., Peifer,P., Mortin,M. and Clevers,H. (1997) *Cell*, **88**, 789–799.
7 Brunner,E., Peter,O., Schweizer,L. and Basler,K. (1997) *Nature*, **385**, 829–833.
8 Dooijes,D., van Beest,M., van de Wetering,M., Boulanger,G., Jones,T., Clevers,H. and Mortin,M. (1998) *Mol. Gen. Genet.*, **258**, 45–52.
9 Clos,J., Westwood,J., Becker,P., Wilson,P., Lambert,K. and Wu,C. (1990) *Cell*, **63**, 1084–1097.
10 Chomczynski,P. and Sacchi,N. (1987) *Anal. Biochem.*, **162**, 156–159.
11 Ashburner,A. (1989) *Drosophila: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
12 Auclair,B., Lang,F., Forest,P. and Desgrosseillers,L. (1994) *Eur. J. Biochem.*, **220**, 997–1003.
13 Garrett,J., Singh,K., Vonder Haar,R. and Emr,S. (1991) *Mol. Gen. Genet.*, **225**, 483–491.
14 Metspalu,A., Rebane,A., Hoth,S., Pooga,M., Stahl,J. and Kruppa,J. (1992) *Gene*, **119**, 313–316.
15 Nolte,D., Taimor,G., Kalff-Suske,M. and Seifart,K. (1996) *Gene*, **169**, 179–185.
16 Kho,C. and Zarbl,H. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 2200–2204.
17 Lecomte,F., Szpirer,J. and Szpirer,C. (1997) *Gene*, **186**, 271–277.
18 Gordon,H., Kucera,G., Salvo,R. and Boss,J. (1992) *J. Immunol.*, **148**, 4021–4027.
19 Ito,M., Kodama,H. and Komamine A. (1991) *Plant J.*, **1**, 141–148.
20 Zurita,M., Reynaud,E. and Kafatos,F. (1997) *Insect Mol. Biol.*, **6**, 55–62.
21 Kongsuwan,K., Yu,Q., Vincent,A., Frisardi,M., Rosbash,M., Lengyel,J. and Merriam,J. (1985) *Nature*, **317**, 555–558.
22 Sæbøe-Larssen,S. and Lambertson,A. (1996) *Genetics*, **143**, 877–885.
23 O'Hare,K., Alley,M., Cullingford,T., Driver,A. and Sanderson,M. (1991) *Mol. Gen. Genet.*, **225**, 17–24.
24 Driver,A., Lacey,S.F., Cullingford,T.E., Mitchelson,A. and O'Hare,K. (1989) *Mol. Gen. Genet.*, **220**, 49–52.
25 Schmidt,A., Hollmann,M. and Schäfer,U. (1996) *Mol. Gen. Genet.*, **251**, 381–387.
26 Sæbøe-Larssen,S., Urbanczyk Mohebi,B. and Lambertsson,A. (1997) *Mol. Gen. Genet.*, **255**, 141–151.
27 Andersson,S., Sæbøe-Larssen,S., Lambertson,A., Merriam,J. and Jacobs-Lorena,M. (1994) *Genetics*, **137**, 513–520.
28 Ntwasa,M., Bachana,S. and Gay,N. (1994) *Biochim. Biophys. Acta*, **1218**, 210–212.
29 Hart,K., Klein,T. and Wilcox,M. (1993) *Mech. Dev.*, **43**, 101–110.
30 McKim,K., Dahmus,J. and Hawley,R. (1996) *Genetics*, **144**, 215–228.
31 Dorer,D., Anane-Firempon,A and Christensen,A. (1991) *Mol. Gen. Genet.*, **230**, 8–11.
32 Slusarski,D., Motzny,C. and Holmgren,R. (1995) *Genetics*, **139**, 229–240.