

A Physical Map of the Polytenized Region (101EF–102F) of Chromosome 4 in *Drosophila melanogaster*

John Locke, Lynn Podemski, Nicole Aippersbach, Hilary Kemp and Ross Hodgetts

Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

Manuscript received December 1, 1999
Accepted for publication March 22, 2000

ABSTRACT

Chromosome 4, the smallest autosome (~5 Mb in length) in *Drosophila melanogaster* contains two major regions. The centromeric domain (~4 Mb) is heterochromatic and consists primarily of short, satellite repeats. The remaining ~1.2 Mb, which constitutes the banded region (101E–102F) on salivary gland polytene chromosomes and contains the identified genes, is the region mapped in this study. Chromosome walking was hindered by the abundance of moderately repeated sequences dispersed along the chromosome, so we used many entry points to recover overlapping cosmid and BAC clones. *In situ* hybridization of probes from the two ends of the map to polytene chromosomes confirmed that the cloned region had spanned the 101E–102F interval. Our BAC clones comprised three contigs; one gap was positioned distally in 102EF and the other was located proximally at 102B. Twenty-three genes, representing about half of our revised estimate of the total number of genes on chromosome 4, were positioned on the BAC contigs. A minimal tiling set of the clones we have mapped will facilitate both the assembly of the DNA sequence of the chromosome and a functional analysis of its genes.

THE fruit fly *Drosophila melanogaster* has two sex chromosomes and three autosomes. The smallest chromosome, 4, is ~5 Mb length (Locke and McDerimid 1993) and consists of two major regions. The centromeric region is heterochromatic and the three bands, h59, h60, and h61, identified by Hoechst-33258 staining and N banding, contain primarily short satellite repeats (Lohe *et al.* 1993). This region forms part of the highly condensed, underreplicated chromocenter seen in polytene chromosome spreads. The other region, which contains the remaining ~1.2 Mb, constitutes cytogenetic bands 101E–102F on polytene salivary gland chromosomes. Of the ~75 genes Hochman (1976) predicted on chromosome 4, few have been mapped but all are likely in the 101E–102F region. Most *Drosophila* species have a similar, presumably homologous, small chromosome that appears as a “dot” in metaphase spreads. It is referred to as the F element in Muller’s description of chromosome segments in *Drosophila* species (Muller 1940).

In addition to its small size, chromosome 4 is atypical in several other respects, all of which relate to the long-standing presumption that this chromosome is, in some ways, heterochromatic in nature. First, the banded region of chromosome 4 often shows a diffuse and poorly defined appearance that is similar to regions of heterochromatin. A second similarity to heterochromatin was revealed when it was shown that the chromosomal pro-

tein HP1, thought to be an important constituent of heterochromatin, bound to several sites along the chromosome (Eissenberg *et al.* 1992). Third, repetitive DNA sequences normally confined to heterochromatin are distributed throughout the banded region of 4 (Miklos *et al.* 1988). Recently, a new very short repeat sequence, DINE-1, has been found distributed along chromosome 4 (Locke *et al.* 1999a,b). Fourth, similarities to heterochromatin are indicated by the behavior of *P*-element transgenes inserted into this region since they frequently show variegated expression of a *white*⁺ marker gene (Wallrath and Elgin 1995; Wallrath *et al.* 1996). This is a typical characteristic of insertions near heterochromatic boundaries. Finally, chromosome 4 normally fails to undergo crossing over during female meiosis (Bridges 1935).

The small size of chromosome 4 makes cytogenetic localization of its genes difficult, while the lack of crossing over precludes the construction of a typical genetic map. Consequently, many of the genes on 4 are uncharacterized at present. Hochman (1976), who undertook a systematic mutational screen to catalog the loci on 4, recovered lethal alleles in 36 essential loci and estimated that an additional 30 loci were missed in his screens. The mutants were roughly mapped to one of three regions using the two existing deletions. When we initiated this study, relatively few of the genes had been accurately mapped. More recently, however, several cloned chromosome 4 genes have been localized using *in situ* hybridization. We undertook the construction of a physical map of the chromosome to facilitate an ordering of the genes.

Corresponding author: John Locke, Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada. E-mail: john.locke@ualberta.ca

Using known, unique chromosome 4 sequences as probes, we isolated cosmid clones and initiated several chromosome walks with the goal of forming one cosmid contig spanning the gene-containing region defined cytologically as 101EF–102F. Despite screening one existing cosmid library and building three others, our cosmid map contained several gaps. However, by screening three new bacterial artificial chromosome (BAC) libraries, all but two of these gaps were filled. We now have clones that cover 1.1 Mb of chromosome 4. The physical map and its construction are presented here.

MATERIALS AND METHODS

Maintenance of stocks: *Drosophila* were grown at room temperature on a standard yeast-sucrose medium described in Nash and Bell (1968). In the later stages of the project, a standard cornmeal and molasses medium was used.

***D. melanogaster* cosmid libraries:** *Drosophila* genomic DNA was prepared as described in Locke and Tartof (1994) from adults of the Oregon-R strain. A *Sau3AI* cosmid library was constructed by ligating partially digested, size selected, and dephosphorylated genomic DNA fragments into *Bam*HI linearized pDO193 (Ahmed 1994). The cloning site of pDO193 was modified slightly to introduce flanking promoter sites for T7 and T3 RNA polymerase. The ligated products were packaged *in vitro* using Gigapack XL extracts (Stratagene, La Jolla, CA). Insert-containing colonies were picked into 384-well microtiter plates, stamped onto Genescreen Plus membranes, and processed using standard procedures (NEN Life Science Products, Inc.). A similar procedure was used to construct a *Kpn*I cosmid library and a *Hind*III fosmid library, using the vector pBeloBAC11 (Wang *et al.* 1997). A *Drosophila* (iso-1) genomic cosmid library, made with *Sau3AI* (Tamkun *et al.* 1992), was also screened.

***D. melanogaster* BAC libraries:** Three BAC genomic libraries were screened during the course of this work. The RPCI-98 BAC library (*Eco*RI partial digest) was constructed by Aaron Mammoser and Kazutoyo Osoegawa at Roswell Park Cancer Institute, in collaboration with the *Drosophila* Genome Project at Lawrence Berkeley National Laboratory. The DrosBACN library (*Nde*I partial digest) and the DrosBACH library (*Hind*III partial digest) were both constructed by Alain Billaud at the Centre d'Étude du Polymorphisme Humaine (CEPH) and were distributed by the UK Human Genome Mapping Project (HGMP) Resource Centre. The BAC libraries were obtained as high-density gridded filters.

***Escherichia coli* host strains:** The *Sau3AI* library was plated in *E. coli* host strain DH5 α (GIBCO BRL Life Technologies, Gaithersburg, MD). The *Kpn*I and *Hind*III cosmid and the fosmid library were plated in *E. coli* host strain DH10B (GIBCO BRL Life Technologies). The *Sau3AI* and *Kpn*I cosmid libraries were plated additionally in *E. coli* host strain PMC107 (Raman *et al.* 1997). All of the BAC clones were grown in host strain DH10B.

Radiolabeling of probes: DNA probes (cDNAs or genomic fragments) were labeled by random priming with [α -³²P]dCTP as described by Feinberg and Vogelstein (1984). Fragments used as probes for screening library filters were shown to be free of repetitive DNA by hybridizing to blots of genomic DNA digested with a restriction endonuclease. RNA transcripts labeled with [α -³²P]UTP were made using T7 and T3 RNA polymerases following the manufacturer's recommendations (Boehringer Mannheim, Indianapolis). Transcripts likewise

were shown to be unique before being used to screen the library filters.

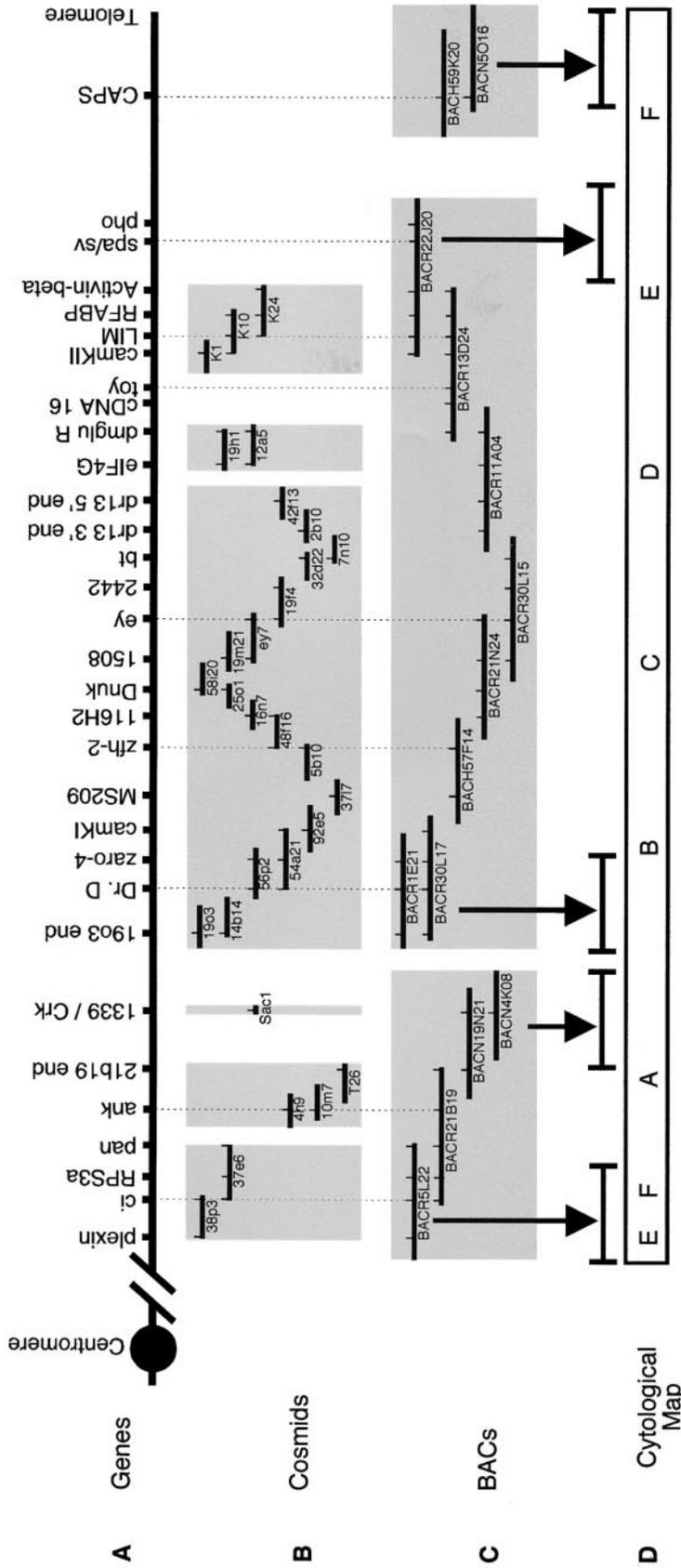
Screening of library filters: DNA-DNA hybridization reactions were performed in 50 mM Na₃PO₄ (pH 7.2), 6 \times SSC, 1% SDS at 65°. RNA-DNA hybridization reactions were performed in 50% formamide, 6 \times SSC, 1% SDS at 42°. High stringency washes consisted of a 30-min wash in 2 \times SSC, 0.1% SDS at 65°, a 30-min wash in 0.2 \times SSC, 0.1% SDS at 65°, and a 30-min wash in 0.1 \times SSC, 0.1% SDS at 65°.

Assembly of clones into contigs: Clones that were deemed positive for a particular probe were propagated in *E. coli*. Cosmid DNA was extracted by alkaline lysis. BAC clone DNA was isolated with a modified alkaline lysis protocol from the UK HGMP Resource Centre using ProCipitate reagent (Ligochem, Inc.). The clones were then assembled into contigs using a combination of cross-screening (Locke *et al.* 1996), restriction fragment analysis, and hybridization reactions using unique end fragments. *In situ* hybridization to polytene squashes was then used to confirm the contig location on chromosome 4.

***In situ* hybridization to salivary chromosomes:** Polytene chromosome squashes were prepared from larval salivary glands using the translocation *T(1:4)w^{m5}* (Lindsley and Zimm 1992), which reciprocally exchanges the banded region of chromosome 4 (101EF–102F) with the tip of the X chromosome (1A–3C). This translocation moves chromosome 4 away from the chromocenter thereby aiding *in situ* cytology. DNA probes were made by labeling whole BAC clones with either digoxigenin-11-dUTP or biotin-16-dUTP, and these were hybridized to polytene squashes using standard procedures (Boehringer Mannheim). Digoxigenin-labeled probes were detected with anti-digoxigenin rhodamine, Fab fragments. Biotin-labeled probes were detected by binding with streptavidin fluorescein, followed by a signal amplification step using biotinylated anti-streptavidin and a final binding step with streptavidin fluorescein. DAPI (4',6-diamidino-2-phenylindole) was used as a counterstain. Rhodamine and fluorescein signals were observed using a fluorescence microscope with the appropriate filters.

RESULTS

Construction of cosmid contigs: Cosmid libraries were initially screened with unique probes mapped to chromosome 4: *cubitus interruptus* (*ci*), *zfh-2*, and *eyeless* (*ey*). Clones, recovered using these initial probes, furnished fragments that were used to identify adjacent cosmid clones in the libraries. As additional unique chromosome 4 probes were obtained, each was used to initiate a new chromosome walk or was mapped to clones in an existing contig. Chromosome walks were extended as long as unique terminal fragments could be identified or until a gap was reached. Using this methodology, we were able to obtain clones that covered about 800 kb of the estimated 1200 kb of the cytological region 101EF–102F (Figure 1). They comprised six contigs of which five were three cosmids or less in length. The sixth, which was estimated to be \sim 400 kb long, covered the region 102B to 102D as determined by *in situ* hybridization of the end clones to polytene chromosomes (data not shown). Unique probes from many of the genes shown at the top of Figure 1 were mapped to various cosmid clones, although a number (*e.g.*, 1339,



101 102

Figure 1.—Diagram of *D. melanogaster* chromosome 4 showing the extent of the mapped region. Shown is an ordered array of the 32 genes and probes (A) used to assemble a contig of 28 cosmid clones (B) and 14 BAC clones (C). The location of these clones is shown relative to the cytological map (D). The downward arrows from the five BAC clones in C show their location on the salivary gland polytene chromosome cytological map as determined by *in situ* hybridization (see Figure 2). Note: the relative order of cDNA16 and toy, as well as the order of Rf β BP and dActivin- β , could not be established from our data. The gray boxes indicate the extent of contigs, while the dotted vertical lines serve as a guide to show the positions of a subset of the probes.

spa and *pho*) were not represented in the cosmid libraries.

Construction of BAC contigs: When the three BAC clone libraries became available (see materials and methods), we began by screening them with selected unique probes from the existing cosmid clones. We did this first to confirm the cosmid contigs and second to obtain BACs that spanned the gaps between the cosmid contigs. Using the BAC clones, we were able to coalesce the six cosmid contigs into two BAC contigs (Figure 1). Furthermore, the BAC libraries contained clones that hybridized to calcium-activated protein secretion (CAPS) (Walent *et al.* 1992), a probe that had not been used to screen for cosmid clones. This produced an additional BAC contig located distally on the chromosome that consisted of the three overlapping clones (BACH59K20, BACN40B03, and BACN5O16), two of which are shown in Figure 1.

The three BAC contigs (Figure 1) extend over the cytological region 101EF to 102F, the full length of the visible polytene chromosome. We used the terminal BAC clones (BACR5L22 and BACN5O16) that defined the proximal and distal limits of the three contigs as probes to polytene chromosomes to map their cytological location. Clone BACR5L22 hybridizes to the most proximal region on chromosome 4, 101EF, while the clone BACN5O16 hybridizes to the distal-most band at 102F and should include the telomere region (Figure 2A) since BACN5O16 also hybridizes to the telomeres of other chromosomes.

The distal gap is substantially larger than the proximal gap. When its flanking clones, BACR22J20 and BACH59K20, are used as probes to polytene chromosomes, there is a visible gap that spans several polytene bands (Figure 2B). On the basis of the width of the adjacent BAC clone hybridization signals we estimate this gap to be about one or two BACs in length (about 50–200 kb). We have a probe derived from a chromosome 4 *P*-element insertion, 118-E15 (Wallrath and Elgin 1995), that we mapped to within this gapped region by *in situ* hybridization. This probe also cross-hybridizes to a single site on the X chromosome. We have used this fragment to screen the BAC libraries and recovered >20 clones, all of which overlap and correspond to the X-chromosome site. This provides a direct comparison of the likelihood for recovering chromosome 4 *vs.* X-chromosome BAC clones and represents what may be a typical bias against chromosome 4 sequences in this, and possibly other, clone libraries.

The proximal gap, flanked by clones BACN4K08 and BACR1E21, is very small. Using these BACs as probes to polytene chromosomes reveals no visible gap between the two regions of hybridization (Figure 2C). Probes near the ends of both BAC clones identify common-sized restriction fragments on Southern transfers of genomic DNA. From their size we estimate that this gap is less than 1 kb. Although the very ends of both BACs

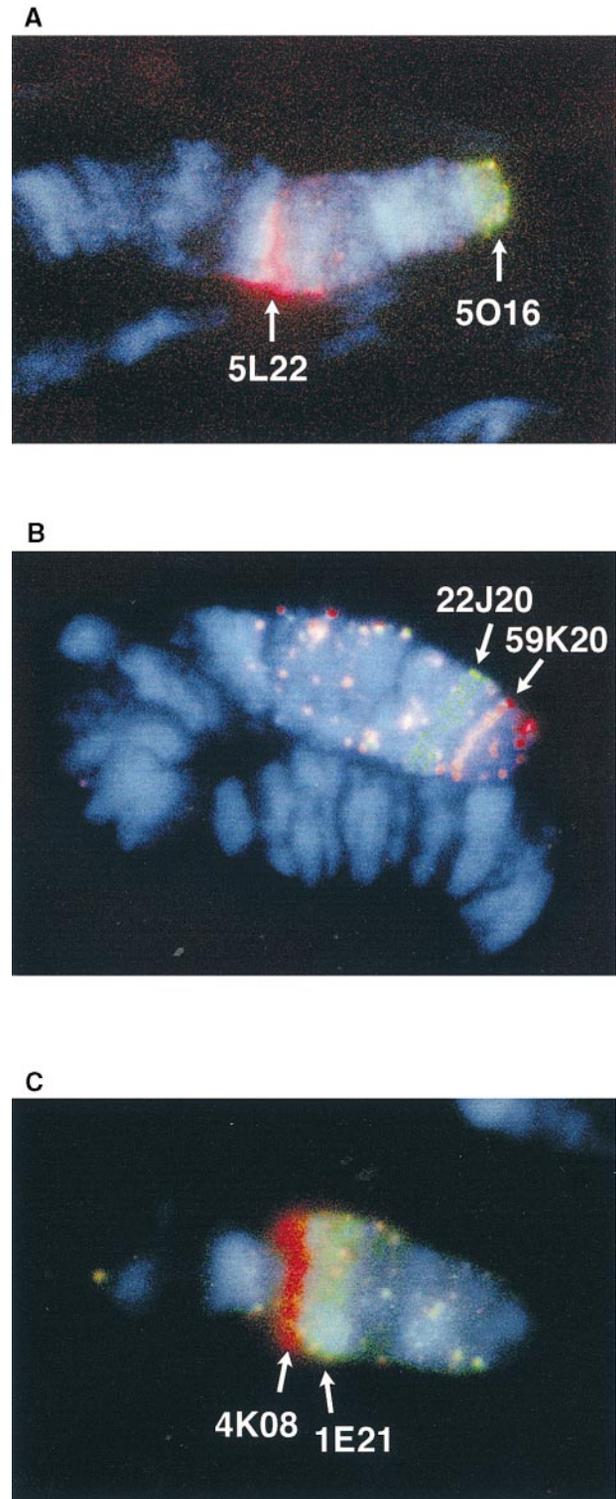


Figure 2.—(A) *In situ* hybridization of BAC clones to correlate the physical and cytogenetic maps. BAC clone DNA was labeled with digoxigenin or biotin and detected with rhodamine- or fluorescein-tagged antibodies. Salivary gland chromosomes from the translocation-bearing strain $T(1;4)w^{m5}$ were probed with BACR5L22 and BACN5O16, which are the terminal BAC clones shown on the map in Figure 1. (B) Hybridization of the BAC clones BACR22J20 and BACH59K20, which flank the distal gap, to chromosome 4 in $T(1;4)w^{m5}$. (C) Hybridization of the BAC clones BACN4K08 and BACR1E21, which flank the proximal gap, to chromosome 4 in $T(1;4)w^{m5}$.

BACN4K08 and BACR1E21 are repetitive, single-copy sequences somewhat removed from the gap in both clones were identified and used to probe the three BAC libraries. However, these probes failed to identify any new BACs that would have closed the gap. The repetitive sequences at the BAC clone ends also precluded us from using PCR methods to span this gap.

Localization of previously identified loci: During the construction of this physical map, we obtained a set of 23 cloned sequences from genes or anonymous cDNAs that had been mapped to chromosome 4 by various methods. These were positioned on the BAC and cosmid clones and their unambiguous order along the minimal tiling set of clones is shown at the top of Figure 1. Where known, the distal-proximal orientation of genes is shown. The location shown for the highly repetitive clone described as "Dr. D" (Miklos *et al.* 1988), which contains sequences of both DINE-1 and the Hoppel transposon (Locke *et al.* 1999a,b), was determined using a single-copy subfragment. The unique probes also hybridized to BAC clones not shown in Figure 1 that are listed in the appendix. The number of clones that hybridized to a particular probe varied from 1 to 25, but on average 5 clones were identified by each unique probe from the three libraries.

Distribution of *Sau3AI* restriction fragments in regions represented by or lacking cosmid clones: Some regions, for example, those represented by the clone 1339, were absent from our cosmid libraries but present in BAC clones, suggesting a bias may have existed in the distribution of cosmid clones. A bias of *Sau3AI* sites (used to make the largest cosmid library) could have distorted the production and/or recovery of the smaller cosmid clones. Since we now have BAC clones from regions lacking cosmid clones, we can ask whether the frequency of *Sau3AI* sites (as revealed by the size of *Sau3AI* fragments) differs in these regions compared to those represented in the cosmid libraries. We took three BAC clones from a cosmid-poor region (BACN28B18, BACN4K08, BACR1E21) and three from a cosmid-rich region (BACR44L03, BACR21N24, BACR30L15) and digested their DNA with *Sau3AI* (Figure 3A). The restriction fragments of both groups are similar in size suggesting no difference in the distribution of *Sau3AI* sites.

To show that we could detect a restriction enzyme cleavage bias through restriction fragment size differences, we used two other restriction enzymes, *MseI* and *HhaI*. Whereas *Sau3AI* cleaves at the sequence GATC, in which AT and GC content are equal, *MseI* cleaves at an AT-rich site (TTAA), while *HhaI* cleaves at a GC-rich site (GCGC). A comparison of the size of restriction fragments produced by these enzymes in the same BAC clones shows a difference in cleavage frequency. *MseI* produces mostly small fragments of <1 kb (Figure 3C), whereas *HhaI* produces fragments that are mostly greater than 1 kb (Figure 3B). The easily observable

difference in fragment lengths indicates we could have detected differences in *Sau3AI* restriction if they had existed. Thus we can rule out a bias in *Sau3AI* cleavage sites as the cause of the incomplete coverage of the chromosome by the cosmid clones. Furthermore, these data support the idea that chromosome 4, at least as it is represented by these BAC clones, is AT rich.

DISCUSSION

The presence of many gaps in our initial cosmid contig is not unexpected because several factors are known to bias the recovery of genomic regions using cosmid vectors. These include a bias in the distribution of cleavage sites for the restriction enzyme used to construct the libraries, the inability of a fragment to be propagated in *E. coli*, and the chance probability of a clone not being included. Attempts to walk across several of the gaps were confounded by the presence of repeated sequences at the termini of the flanking clones. Locke *et al.* (1999a,b) estimated that approximately half the restriction fragments from chromosome 4 cosmid clones contain repeated sequences. To minimize this problem we screened BAC libraries, which contained inserts about three times as long as those in the cosmid clones, and recovered clones that reduced the number of gaps to two. It appears that sequences from these two gaps are not present in any of the three BAC libraries, although these libraries do contain a fair representation of other chromosome 4 sequences. The appendix shows that cloned regions vary in coverage from 1 to 25 BACs, with an average of ~5 clones per unique probe region.

Absence of clones in libraries built using partial digestion with *Sau3AI* might be due to a bias in the distribution of its cleavage site, GATC. This possibility was discarded for our partial *Sau3AI* cosmid library because BAC clones from regions within the gapped regions of the cosmid contigs had *Sau3AI* fragment sizes equivalent to those of BAC clones included in cosmid clones. Thus these sequences must be absent from these cosmid libraries for other reasons.

Previous *in situ* hybridization data have shown that chromosome 4 has a sequence distribution that differs from the other chromosomes (Pardue *et al.* 1987; Lowenhaupt *et al.* 1989). Probing with the simple sequence (dC-dA)_n·(dG-dT)_n showed hybridization to the euchromatic arms of all the chromosomes except chromosome 4. The (dC-dA)_n·(dG-dT)_n probe distribution parallels that of the presence of crossing over (all euchromatic arms except chromosome 4). A large-scale genomic DNA sequence available for chromosome 4 was obtained from two cosmids located at some distance from each other (Ahmed and Podemski 1997; Locke *et al.* 1999a,b). Both sequences are very AT rich and together show only 34.8% G + C. This is substantially less than the genome-wide average of 43% (Laird and McCarthy 1969) or the value of 40.8% for an ~2.9-

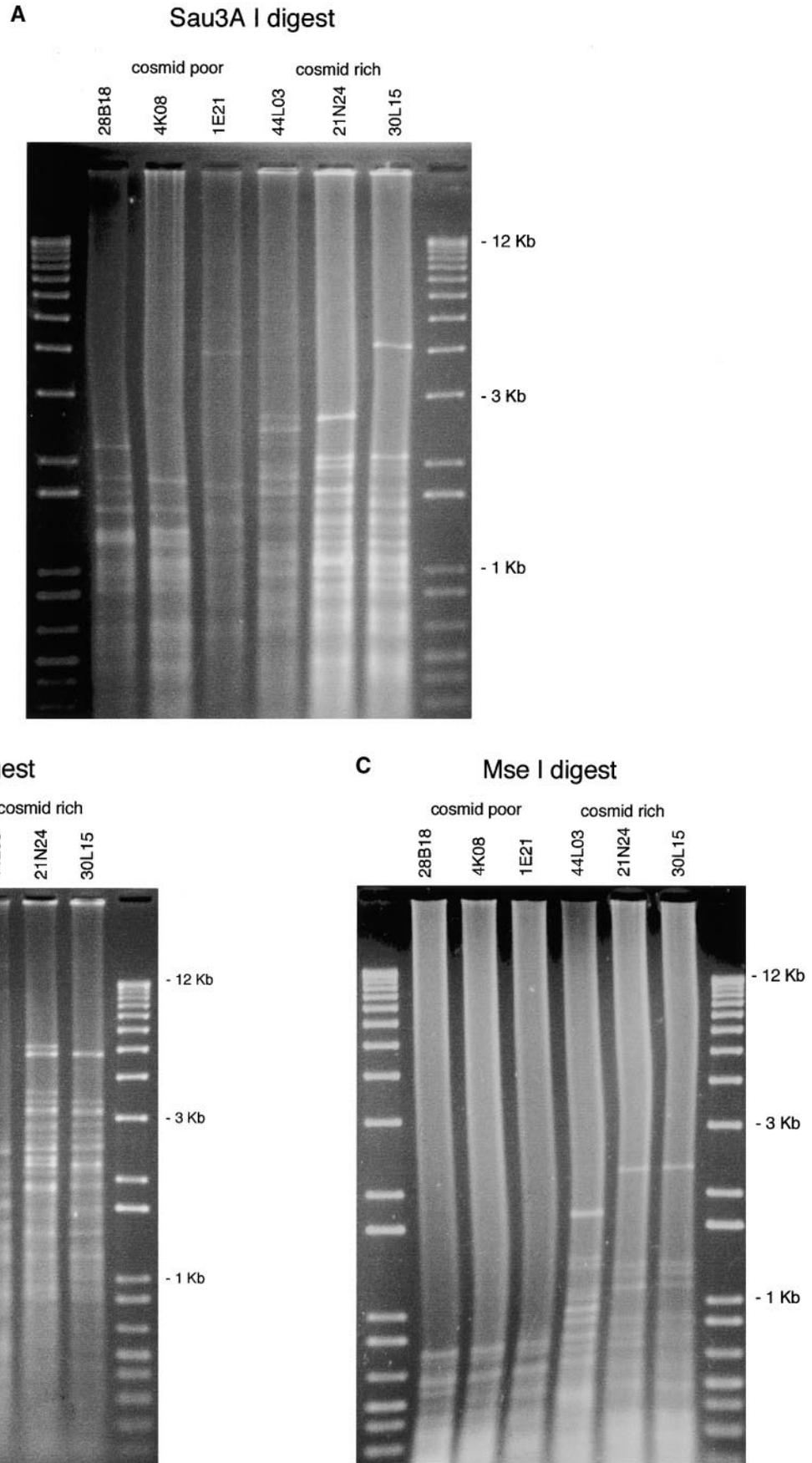


Figure 3.—Gel electrophoresis of restriction digests of BAC clones, showing the size of restriction fragments from cosmid-poor (BACN28B18, BACN4K08, BACR1E21) and cosmid-rich (BACR44L03, BACR21N24, BACR30L15) regions of the contig. The DNA was digested with *Sau3A*I (A), *Hha*I (B), or *Mse*I (C) and fractionated on a 1.2% agarose gel.

Mbp section of euchromatin from chromosome 2L (Ashburner *et al.* 1999). The large reduction in GC content on chromosome 4 is further reflected in the prevalence of the restriction site for *MseI* (TTAA) compared to that for *HhaI* (GCGC) in the BAC clones we tested. The sequence peculiarities in chromosome 4 DNA [*e.g.*, AT richness and lack of (dC-dA)_n·(dG-dT)_n] may be relevant to the lack of crossing over on this chromosome.

We have localized 23 genes and anonymous cDNAs to the clone set shown in Figure 1. Included among the genes are four (*ci*, *ey*, *spa/sv*, and *bt*) of the seven visible mutations known to lie on the chromosome. Hochman (1976) described eight visible mutations on chromosome 4; however, it has recently been shown that *spa* and *sv* are alleles at a single complex locus (Fu *et al.* 1998). We have placed this locus within the region of overlap between BACR17F10 and BACR22J20 (Figure 1). Hochman (1976) used an approximation to the Poisson distribution for those loci that mutated either once or twice to predict that ~72 loci were present on the chromosome. However, a reexamination of his data, to include all genes with fewer than six alleles (25 in total), suggests the total number of loci on the chromosome is ~48. Another estimate of gene number, one protein coding gene per 13 kb (Ashburner *et al.* 1999), suggests chromosome 4 should contain 92 genes. But the analysis of two chromosome 4 cosmids containing 63 kb (Locke *et al.* 1999a,b) indicates that the gene density on chromosome 4 may not be as high as that on 2L and supports the revised lower estimates. These estimates may soon be out of date since the two major autosomes have been mapped (Hoskins *et al.* 2000) and the *Drosophila* genome sequence is being determined. The annotation of chromosome 4 sequences will provide a better estimate of gene number.

The repetitive sequences within the ~2.9 Mb region of 2L consist of six LINE elements, 11 retrotransposons, and one short (~171 bp) repeat (Ashburner *et al.* 1999). While our contig for chromosome 4 has not been probed for all the different *Drosophila* transposable elements, we do know that, on average, chromosome 4 clones have a repeated sequence every other restriction fragment (Locke *et al.* 1999a,b). While not truly quantitative, this observation clearly differentiates chromosome 4 from chromosome 2L, on which 1.8% of the sequence is repetitive (Ashburner *et al.* 1999), and from the *Drosophila* euchromatin in general, in which the repetitive DNA content is 9%. If the sequence from the two chromosome 4 cosmids provides an accurate view, then we know that the nature and interspersal pattern of repetitive DNA on chromosome 4 differs from that on the other chromosomes. Analysis of >63 kb showed that four transposable elements (including three Hoppel elements) and 18 of the short repeats we have called DINE-1 were present (Locke *et al.* 1999a,b). The euchromatic distribution of both Hoppel and DINE-1 is

almost exclusively confined to chromosome 4 and explains why extending chromosome walks on this chromosome is more difficult than elsewhere in the genome (Bender *et al.* 1983a,b; Garber *et al.* 1983; Scott *et al.* 1983).

Our physical map will assist in the assembly of the DNA sequence for the entire chromosome. There are perhaps as many as 25 genes represented by lethal alleles and three more by visible mutations that have yet to be coupled to specific sequenced genes. We (Locke *et al.* 1999a,b) and others (Miklos *et al.* 1988) have invoked the relative richness of repeats on chromosome 4 to explain its partly heterochromatic properties. The BAC clones will provide a useful resource for studying the genes on chromosome 4 and the functional significance of the interspersal pattern of its repetitive elements on the expression of these genes.

We thank B. Hepperle, G. Rairdan, D. Adams, and E. Woloshyn for excellent technical assistance. We also thank J. Bell, D. Pilgrim, K. Roy, D. Nash, H. McDermid, and A. Ahmed for their earlier support and interest. In addition, we thank the many researchers who provided chromosome 4 sequence probes used to identify the cosmid and BAC clones in this work: K. Arora, B. Berwin, P. Callaerts, A. R. Campos, H. Clevers, J. L. Couderc, R. Dubreuil, S. C. R. Elgin, M. Forte, E. Frei, E. A. Fyrberg, C. S. Goodman, Y. Grau, T. E. Haery, M.-L. A. Joiner, J. A. Kassis, S. Kunes, R. K. Kutty, P. Lasko, R. W. Lewis, J. Mehrens, M. Noll, J. N. Noordermeer, C. J. O'Kane, S. N. Robinow, S. Russel, A. Scully, I. Sidén-Kiamos, B. Stronach, C. Sung, S. Sweeney, L. L. Wallrath, A. Whitworth, M. Winberg, A. Worthington, the *Drosophila* Genome Project at Lawrence Berkeley National Laboratory, and the UK Human Genome Mapping Project Resource Centre. L.P. conducted or supervised the experimental work with the help of N.A. and H.K. J.L. and R.H. jointly conceived the project, designed the experimental approach and participated in the data analysis. J.L. wrote the article with substantial contributions from R.H. and L.P. This work was funded by grants from the Canadian Genome Analysis and Technology Program, the Medical Research Council of Canada, and the Natural Sciences and Engineering Research Council of Canada.

LITERATURE CITED

- Ahmed, A., 1994 Isolation of nested, bidirectional deletions by the use of exonuclease III on double-origin vectors. *Gene* **151**: 97–101.
- Ahmed, A., and L. Podemski, 1997 Use of ordered deletions in genome sequencing. *Gene* **197**: 367–373.
- Ashburner, M., S. Misra, J. Roote, S. E. Lewis, R. Blazej *et al.*, 1999 An exploration of the sequence of a 2.9-Mb region of the genome of *Drosophila melanogaster*. The *adh* region. *Genetics* **153**: 179–219.
- Bender, W., M. Akam, F. Karch, P. A. Beachy, M. Peifer *et al.*, 1983a Molecular genetics of the bithorax complex in *Drosophila melanogaster*. *Science* **221**: 23–29.
- Bender, W., P. Spierer and D. S. Hogness, 1983b Chromosomal walking and jumping to isolate DNA from the *Ace* and *rosy* loci and the bithorax complex in *Drosophila melanogaster*. *J. Mol. Biol.* **168**: 17–33.
- Bridges, C. B., 1935 The mutants and linkage data of chromosome four of *Drosophila melanogaster*. *Biol. Zh.* **4**: 401–420.
- Eissenberg, J. C., G. D. Morris, G. Reuter and T. Hartnett, 1992 The heterochromatin-associated protein HP-1 is an essential protein in *Drosophila* with dosage-dependent effects on position effect variegation. *Genetics* **131**: 345–352.
- Feinberg, A. P., and B. Vogelstein, 1984 A technique for radiolabelling DNA restriction fragments to high specific activity. *Anal. Biochem.* **137**: 266–267.

- Fu, W., H. Duan, E. Frei and M. Noll, 1998 *shaven* and *sparkling* are mutations in separate enhancers of the *Drosophila* Pax2 homolog. *Development* **125**: 2943–2950.
- Garber, R. L., A. Kuroiwa and W. J. Gehring, 1983 Genomic and complementary DNA clones of the homeotic locus *Antennapedia* in *Drosophila*. *EMBO J.* **2**: 2027–2036.
- Hochman, B., 1976 The fourth chromosome of *Drosophila melanogaster*, pp. 903–928 in *The Genetics and Biology of Drosophila*, edited by M. Ashburner and E. Novitski. Academic Press, London.
- Hoskins, R. A., C. R. Nelson, B. P. Berman, T. R. Laverty, R. A. George *et al.*, 2000 A BAC-based physical map of the major autosomes of *Drosophila melanogaster*. *Science* **287**: 2271–2274.
- Laird, C. D., and B. J. McCarthy, 1969 Molecular characterization of the *Drosophila* genome. *Genetics* **63**: 865–882.
- Lindsley, D. L., and G. G. Zimm, 1992 *The Genome of Drosophila melanogaster*. Academic Press, San Diego.
- Locke, J., and H. McDermid, 1993 Analysis of *Drosophila* chromosome four using pulsed field gel electrophoresis. *Chromosoma* **102**: 718–723.
- Locke, J., and K. D. Tartof, 1994 Molecular analysis of *cubitus interruptus* (*ci*) mutations suggest an explanation for the unusual *ci*-position effects. *Mol. Gen. Genet.* **243**: 234–243.
- Locke, J., G. Rairdan, H. McDermid, D. Nash, D. Pilgrim *et al.*, 1996 Cross-screening: a new method to assemble clones rapidly and unambiguously into contigs. *Genome Res.* **6**: 155–165.
- Locke, J., L. T. Howard, N. Aippersbach, L. Podemski and R. B. Hodgetts, 1999a The characterization of DINE-1, a short, interspersed repetitive element present on chromosome 4 and in the centric heterochromatin of *Drosophila melanogaster*. *Chromosoma* **108**: 356–366.
- Locke, J., L. Podemski, K. Roy, D. Pilgrim and R. Hodgetts, 1999b Analysis of two cosmid clones from chromosome four of *Drosophila melanogaster* predicts two new genes amid an unusual arrangement of repeated sequences. *Genome Res.* **9**: 137–149.
- Lohe, A. R., A. J. Hilliker and P. A. Roberts, 1993 Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster*. *Genetics* **134**: 1149–1174.
- Lowenhaupt, K., A. Rich and M. L. Pardue, 1989 Nonrandom distribution of long mono- and dinucleotide repeats in *Drosophila* chromosomes: correlations with dosage compensation, heterochromatin, and recombination. *Mol. Cell. Biol.* **9**: 1173–1182.
- Miklos, G. L. G., M.-T. Yamamoto, J. Davies and V. Pirrotta, 1988 Microcloning reveals a high frequency of repetitive sequences characteristic of chromosome four and the β -heterochromatin of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **85**: 2051–2055.
- Muller, H. J., 1940 Bearings of the *Drosophila* work on systematics, pp. 185–268 in *The New Systematics*, edited by J. Huxley. Clarendon Press, Oxford.
- Nash, D., and J. Bell, 1968 Larval age and the pattern of DNA synthesis in polytene chromosomes. *Can. J. Genet. Cytol.* **10**: 82–90.
- Pardue, M. L., K. Lowenhaupt, A. Rich and A. Nordheim, 1987 (dC-dA)_n(dG-dT)_n sequences have evolutionarily conserved chromosomal locations in *Drosophila* with implications for roles in chromosome structure and function. *EMBO J.* **6**: 1781–1789.
- Raman, V., D. Woodcock and R. J. Hill, 1997 Typically unstable long tandem repeats in *Escherichia coli* show increased stability in strain PMC107 and are stable when incorporated into the *Drosophila melanogaster* genome. *Anal. Biochem.* **245**: 242–245.
- Scott, M. P., A. J. Weiner, T. I. Hazelrigg, B. A. Polisky, V. Pirrotta *et al.*, 1983 The molecular organization of the *Antennapedia* locus of *Drosophila*. *Cell* **35**: 763–776.
- Tamkun, J. W., R. Deuring, M. P. Scott, M. Kissinger, A. M. Pattatucci *et al.*, 1992 *brahma*: a regulator of *Drosophila* homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell* **68**: 561–572.
- Walent, J. H., B. W. Porter and T. F. Martin, 1992 A novel 145 kd brain cytosolic protein reconstitutes Ca(2+)-regulated secretion in permeable neuroendocrine cells. *Cell* **70**: 765–775.
- Wallrath, L. L., and S. C. R. Elgin, 1995 Position effect variegation in *Drosophila* is associated with an altered chromatin structure. *Genes Dev.* **9**: 1263–1277.
- Wallrath, L. L., V. P. Gunter, L. E. Rosman and S. C. R. Elgin, 1996 DNA representation of variegating heterochromatic P-element inserts in diploid and polytene tissues of *Drosophila melanogaster*. *Chromosoma* **104**: 519–527.
- Wang, K., C. Boysen, H. Shizuya, M. I. Simon and L. Hood, 1997 Complete nucleotide sequence of two generations of a bacterial artificial chromosome cloning vector. *Biotechniques* **23**: 992–994.

Communicating editor: R. S. Hawley

APPENDIX

Probe	Positive BAC clones
plexin	Not tested
<i>ci</i> (<i>I(4)102ABc</i>)	BACR5L22, BACR21B19
RPS3A (<i>M(4)10I</i>)	Not tested
<i>pan</i> (<i>dTCF</i> or <i>I(4)102ABb</i>)	Not tested
ankyrin	BACR21B19
21B19 end fragment	BACN23C02, BACN28I10, BACN28B18, BACN19N21, BACN11A02
cDNA clone 1339/ <i>Crk</i> 19o3/M010.red	BACN42J09, BACN11A02, BACN4K08, BACN19N21, BACN5P15 BACR1E21, BACR2E19, BACR4N17, BACR6C06, BACR8L14, BACR9M13, BACR15G16, BACR22F16, BACR24O10, BACR30L17, BACR30P11, BACR31P13, BACR33K21, BACR33O03, BACR40F19, BACR40P19, BACR42A05, BACR48M02, BACH54L20, BACN29N11, BACN38G09, BACN41M19, BACN45E04, BACN15K03, BACN5A18
Dr. D (unique fragment)	BACR1E21, BACR6C06, BACR9M13, BACR15G16, BACR30L17, BACR30P11, BACR33O03, BACR40F19, BACR40P19, BACR42A05, BACR48M02, BACH54L20
<i>zaro-4</i>	BACR1E21, BACR30L17, BACR30P11, BACR33O03, BACR40F19, BACR40P19, BACH54L20
camKI	BACR30L17, BACR30P11, BACH54L20, BACN33E06, BACN42F03, BACN4K23
Clone MS209	BACH57F14, BACN21F17, BACN23A06, BACN24M16
<i>zfh-2</i>	BACR7F02, BACR7H02, BACR10A06, BACR26O17, BACR28O12, BACR44L03, BACR46D07, BACR47B13, BACH52H02, BACH57F14, BACN32C05, BACN10B22, BACN10B24, BACN10D24, BACN21F17, BACN23A06, BACN24M16
116H2	Not tested
Druk	BACR6K04, BACR21N24, BACR26O17, BACR30L15, BACR36P12, BACR46D07
cDNA clone1508	BACR6K04, BACR16H06, BACR16J06, BACR21N24, BACR30L15, BACR36P12
<i>ey</i> (<i>eyeless</i>)	BACR6K04, BACR16H06, BACR16J06, BACR30L15
cDNA clone 2442	BACR6K04, BACR16H06, BACR16J06, BACR30L15
<i>bt</i> (<i>bent</i>)	BACR30L15
drunc13 (3')	BACR11A04, BACR16N23
drunc13 (5')	BACR11A04, BACR16N23, BACR46C22
eIF4G	BACR11A04, BACR16N23, BACR46C22
DmGluR	BACR11A04, BACR13D24, BAC14M01, BACR16N23, BACR46C22
clone16	BACR13D24, BACR17F10
<i>toy</i> (<i>twin of eyeless</i>)	BACR13D24, BACR17F10
camKII	BACR13D24, BACR14M01, BACR17F10
lim	BACR13D24, BACR14M01, BACR17F10
RfaBP	Not tested
activin- β	BACR13D24, BACR17F10, BACR22J20
<i>spa-1</i> (<i>sparkling</i>)	BACR17F10, BACR22J20, BACR31B07
<i>spa-4</i>	BACR9E01, BACR9F01, BACR17F10, BACR22J20, BACR31B07
<i>spa-3</i>	BACR9E01, BACR9F01, BACR18J01, BACR22J20, BACR35H05
<i>pho</i> (<i>pleohomeotic</i>)	BACR22J20
CAPS	BACH59K20, BACN40B03, BACN5O16

BACR clones are from the RPCI-98 *Drosophila EcoRI* library, BACN clones are from the CEPH *Drosophila NdeI* library (384-well format), and BACH clones are from the CEPH *Drosophila HindIII* library (384-well format).