DEVELOPMENTAL BIOLOGY 108, 341-354 (1985)

The *Jonah* Genes: A New Multigene Family in *Drosophila melanogaster*

JOHN R. CARLSON[1] AND DAVID S. HOGNESS[2]

Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305

Received August 24, 1984; accepted in revised form October 22, 1984

The *Jonah* gene family consists of approximately 20 genes, distributed in small clusters at eight or more widely dispersed chromosomal sites. Gene clusters differ in the number of genes per cluster, ranging up to four, and in the arrangement of genes within clusters, which include direct repeats, inverted repeats, and combinations of direct and inverted repeats. In the third-instar larva the *Jonah* genes are abundantly expressed as transcripts of a single size class, located exclusively in the midgut. The *Jonah* genes were initially defined by homology to a reference cDNA clone. However, the structural heterogeneity within this gene family is so large that *Jonah* genes that fail to hybridize to the reference cDNA, yet hybridize to other members of the family that do hybridize to that cDNA have been subsequently identified. These findings raise questions regarding the sharpness of gene family boundaries.

INTRODUCTION

Multigene families in *Drosophila melanogaster* exhibit patterns of chromosomal organization that are marked by two extremes. At one extreme, exemplified by the actin genes, dispersion is maximized so that each member of the family occupies a separate chromosomal site (Tobin *et al.*, 1980; Fyrberg *et al.*, 1980). By contrast, dispersion is minimized and clustering maximized at the other extreme, exemplified by the histone genes, where all members are localized at a single site (Lifton *et al.*, 1978). The *Jonah* gene family described in this and the accompanying article (Carlson and Hogness, 1985) exhibits a pattern of the third kind, in which the members are both dispersed and clustered.

Our attention was first drawn to this new multigene family by studies in this laboratory on the *Sgs* genes that encode the glue polypeptides synthesized and excreted by the larval salivary gland (Wolfner, 1980; Muskavitch and Hogness, 1980, 1982; Meyerowitz and Hogness, 1982; Guild, 1984). Transcription of the *Sgs* genes is initiated during the third instar and is turned off at the end of this last larval stage in response to a dramatic increase in the concentration of the steroid molting hormone, ecdysterone, which signals the beginning of metamorphosis to the adult fly. The procedure for isolating the *Sgs* genes therefore involved the construction of a library of cDNA clones from the poly(A)$^+$ RNA of third-instar salivary glands obtained prior to this rise in ecdysterone concentration, and the screening of that library for clones representing those RNAs whose abundance rapidly decreases at the end of larval life, when the ecdysterone concentration is high. While most of the cDNA clones that passed this screen did indeed derive from the *Sgs* genes, one group of homologous clones did not. Our analysis of the *Jonah* family has its origins in the cDNAs of this group.

The first indication that these *Jonah* cDNAs differed from others that passed the screen was the observation that they contained repeated sequences found at several dispersed chromosomal sites (Wolfner, 1980). More startling, however, was our subsequent finding that the *Jonah* RNA, detected by its hybridization to the cloned *Jonah* cDNA, was not detectable in hand-isolated salivary glands. Rather, *Jonah* RNA is specific to the midgut, where it is sufficiently abundant that small amounts of midgut in the mass-isolated salivary glands used for the isolation of the cDNA clones gave rise to significant amounts of *Jonah* RNA. It was this surprising discovery that prompted the name *Jonah*, after the Biblical prophet who unexpectedly turned up in the gut.

In this first article, we analyze the chromosomal arrangement and structural variation of the *Jonah* genes by isolating cloned genomic DNA segments carrying the genes from different chromosomal sites and determining the position, orientation and, at low resolution, structure of the genes within these segments. The results reveal a complex family of imperfectly repeated genes that are dispersed among at least eight chromosomal sites, most of which contain small clusters of up to four genes. In the accompanying article (Carlson and Hogness, 1985), we examine the develop-

[1] Present address: Department of Biology, Yale University, New Haven, Conn. 06520.

[2] To whom reprint requests should be addressed.

0012-1606/85 $3.00

mental regulation of *Jonah* gene transcription and consider the function of the resulting transcripts.

MATERIALS AND METHODS

Materials

Restriction endonucleases were purchased from New England BioLabs, Bethesda Research Labs, or Boehringer-Mannheim, except as follows: *Bam*HI was a gift of C. Mann, one preparation of *Sal*I was a gift of J. Lis and another a gift of K. Burtis, and one preparation of *Eco*RI was a gift of D. Finnegan and another was purified according to a scheme based on the method of Modrich and Zabel (1976). *Escherichia coli* DNA polymerase I was a gift of S. Scherer and J. Widom, DNase I was purchased from Worthington, T4 DNA ligase was a gift of S. Scherer, reverse transcriptase was obtained from J. Beard and further purified by R. Padgett, RNase A was obtained from J. Ridge, and one preparation of SI nuclease was obtained from J. Widom and another purchased from Boehringer-Mannheim. ^{32}P-labeled nucleoside triphosphates were purchased from Amersham, and ^{125}I- and ^{3}H-labeled nucleoside triphosphates were purchased from New England Nuclear. Single-stranded ϕX174 DNA was a gift of J. Kobori.

Unless otherwise indicated, *D. melanogaster* used in these experiments was of the strain Oregon-R and derived from a stock obtained from M. Green (University of California, Davis) and maintained in continuous culture as described in Elgin and Miller (1978). For *in situ* hybridizations, larvae of the genotype *gt-1* w^a/*gt x-11 y sc* were used as sources of salivary polytene chromosomes; these larvae were generated from the strains *gt-1* w^a and *gt x-11 y sc*/FM6, obtained from M. Pardue (Massachusetts Institute of Technology, Cambridge).

Recombinant DNA

Unless otherwise noted, the cDNA clones used in these experiments were isolated as described by Wolfner (1980), and were constructed by insertion of cDNA segments into the *Bam*HI site of the plasmid pSC105 (Cohen *et al.*, 1973) by $(dA)_n \cdot (dT)_n$ joints, or, in the case of adm135A8, into the *Pst*I site of pBR322 (Bolivar *et al.*, 1977) by $(dG)_n \cdot (dC)_n$ tailing. The primary genomic clones derived from the λbDm library of Meyerowitz and Hogness (1982). Genomic subclones were obtained from the λbDm clones by a procedure similar to that of Davis *et al.* (1980, p. 138); λdDm refers to subclones employing the λ590 vector (Murray *et al.*, 1977) and aDm refers to subclones constructed by insertion into pBR322 at the *Eco*RI site (aDm2350 and aDm2352), the *Hin*dIII site (aDm2351, aDm2353, aDm2354, aDm2355, aDm2356, aDm2357, aDm2359, and aDm2360), at the *Bam*HI and *Eco*RI sites (aDm2358), or at the *Pst*I site, following $(dG)_n \cdot (dC)_n$ tailing of *Kpn*I ends (aDm2361). The following subclones derived from the indicated parent clones: aDm2352 from λbDm2307; aDm2350 from λbDm2300; aDm2359 and aDm2360 from λbDm2309; λdDm2340, aDm2353, and aDm2357 from λbDm2306; aDm2355 from λbDm2311; aDm2356 from λbDm2302. Transfections were performed using the *E. coli* strain K802 and the method of Davis *et al.* (1980); transformations were performed using HB101 by a method essentially that of Davis *et al.* (1980).

Phage pools were screened essentially according to the method of Benton and Davis (1977). Hybridizing phage were retested, plaque purified, grown, and then purified on two successive CsCl block density gradients as described in Davis *et al.* (1980).

Isolation of DNA

Phage removed from CsCl gradients with a syringe were mixed with an equal volume of 1 *M* Tris (pH 8.0), 0.25 vol of 0.5 *M* EDTA (pH 8.0), and 2.25 vol formamide (Matheson, Coleman, Bell). This mixture was allowed to sit at room temperature or at 37°C for ~15 min, and then phage DNA was precipitated with ethanol. In some cases phage DNA was purified by a rapid procedure similar to that of Davis *et al.*(1980).

Drosophila DNA was purified by a method similar to that of Schachat and Hogness (1973), except that occasionally a rapid lysis purification procedure was used: single flies were individually ground in a conical grinder in 200 μl of 0.2 *M* sucrose, 0.05 *M* EDTA, 0.1 *M* Tris (pH 9.0), 0.5% SDS, 0.1% diethylpyrocarbonate, and the grindate was heated to 65–70°C for 30–60 min. Then 50 μl of 3 *M* KAc was added and the mixtures were placed on ice for 90 min, spun in a microfuge, and the supernatants extracted with phenol and precipitated with ethanol. The pellets were rinsed, dried at 37°C, and resuspended in 10 m*M* Tris (pH 7.4), 1 m*M* EDTA containing 10 μg/ml RNase A at 37°C.

Plasmid DNA was isolated either as described in Muskavitch (1980; pkdm26H2, pkdm8G8) or by the procedure described in Wolfner (1980; adm135A8).

DNA Digestion, Electrophoresis,and Labeling

DNA was digested under conditions specified by the suppliers. Electrophoresis was performed on horizontal slab gels (McDonell *et al.*, 1977), using either Tris-borate-EDTA buffer (Peacock and Dingman, 1968), or Tris-acetate-EDTA buffer (Davis *et al.*, 1980). DNA

was radioactively labeled by nick translation (Rigby *et al.*, 1977) except that [^{32}P]cDNA and the λ strand-specific probe were prepared with reverse transcriptase, using the method of Myers *et al.* (1977) as modified by Lis *et al.* (1978).

A labeled fragment of pkdm8G8 used in screening the genomic library was isolated, following nick translation of pkdm8G8, by digestion with *Sal*I and electrophoresis of the products on a 0.7% agarose gel. The gel was stained with ethidium bromide, and the 1.6-kb band was excised, heated until fluid, and mixed with the other ingredients of the hybridization buffer described below.

DNA Transfer and Hybridization

DNA blots (Southern, 1975) were baked 1½–2 hr, prehybridized in 50% formamide, 5× SSCP (1× SSCP = 120 m*M* NaCl, 15 m*M* Na_3citrate, 20 m*M* Na phosphate (pH 6.8)) and 1× Denhardt's solution (0.02% BSA, 0.02% Ficoll, 0.02% polyvinylpyrrolidone) (Denhardt, 1966) for at least 4 hr, and then hybridized at 42°C for 12–48 hr. Except where specified otherwise, the hybridization buffer consisted of 50% formamide, 5× SSCP, 1 mg/ml carrier DNA, 0.5 mg/ml poly(A) or poly(U) [poly(U) was used in [^{32}P]cDNA hybridizations] and 1× Denhardt's solution. The probe was heated at 80°C for 2–5 min in the hybridization buffer before use. Following hybridization, the filters were washed at 42–45°C for 3 hr with four changes in 50% formamide, 5× SSC (1× SSC = 150 m*M* NaCl, 15 m*M* Na_3citrate). Filters were exposed to Kodak XR-5 X-ray film at −70°C with a Dupont Cronex Lightning-Plus intensifying screen.

In situ hybridizations were carried out essentially as described in Wolfner (1980). The hybridization buffer contained 0.33 *M* NaCl, 10 m*M* Na phosphate (pH 6.8), 10 m*M* $MgCl_2$, 1× Denhardt's solution, 1 mg/ml carrier DNA, and, with iodinated probes only, 1 m*M* iododeoxycytidine. Hybridizations were for approximately 15–18 hr at 65°C, and 2 × 10^5 cpm/slide were used. Slides were washed 1× 30 min in 2× SSC with 1× Denhardt's solution, 4× 30 min in 2× SSC at 60–63°C, and finally, 3× 2× SSC at room temperature.

Electron Microscopy of DNA

The heteroduplex procedure of Davis *et al.* (1980) was followed; where indicated, the isodenaturing 80/50% formamide conditions of Davis and Hyman (1971) were used. Measurements were made with a Hewlett-Packard digitizer; φX174 was generally used as a single-stranded size standard (Sanger *et al.*, 1977) and the λ arms of the λbDm clones were used as an internal double-strand size standard (see Davis *et al.*, 1980).

Isolation of Inverted Repeats

Inverted repeats were isolated by a procedure based on the method of Ohtsubo and Ohtsubo (1976). Phage DNA was denatured in 100 m*M* NaOH at room temperature for 15 min in a microfuge tube. The mixture was then chilled and 200 m*M* Tris (pH 7.0) was added. The mixture, which in the experiment depicted in Fig. 5 contained phage DNA at a concentration of approximately 130 μg/ml, was placed at 68°C for 25 sec, transferred immediately to ice, and a 12-fold excess of cold S1 buffer [30 m*M* NaAc (pH 4.4), 250 m*M* NaCl, 1 m*M* $ZnCl_2$] was added. In the case of the experiment shown in Fig. 5, the mixture was divided into 100-μl aliquots containing ~1 μg of DNA, and various amounts of S1 (10–1250 units) were added. The mixtures were incubated at 37°C for 30 min, chilled to 0°C, 300 m*M* Tris (pH 8.95) and 5 μg tRNA added, and the nucleic acid precipitated with ethanol, rinsed, resuspended in 10 μg/ml RNase A, and electrophoresed on a 1.4% agarose gel.

RESULTS

The Jonah RNA in Third-Instar Larvae is Confined to the Midgut and Restricted to a Single Size Class

The reference cDNA clone, pkdm26H2, used to define the *Jonah* family and its transcripts consists of a duplex cDNA segment, dm26H2, inserted into the *Bam*HI site of the "pk" vector plasmid (the pSC105 of Cohen *et al.*, 1973) by $(dA)_n \cdot (dT)_n$ joints. Its cDNA segment derives from the poly(A)$^+$ RNA found in preparations of mass-isolated salivary glands obtained from Oregon-R third-instar larvae in the intermolt stage. pkdm26H2 is a member of the set of clones representing the class of RNAs whose abundance drops precipitously at the end of third instar, and it belongs to group VI of the six nonoverlapping homologous sequence groups into which this set was divided (see Introduction; Muskavitch, 1980; Wolfner, 1980). We selected this clone over other members of group VI because the 0.9 kb length of the dm26H2 segment (plus joints) most closely approaches the length of the *Jonah* RNA.

Figure 1 shows that glyoxal-denatured *Jonah* RNA detected by hybridization with pkdm26H2 migrates in electrophoretic gels as a single band; its mobility is that expected of molecules 0.91 ± 0.01 kb in length, as determined in a separate experiment by comparison to glyoxalated restriction fragments of pBR322 (data not shown). In the experiment shown in Fig. 1, individual organs were dissected from a single third-instar larva and the RNA from different organs was extracted, fractionated by gel electrophoresis, transferred to DBM

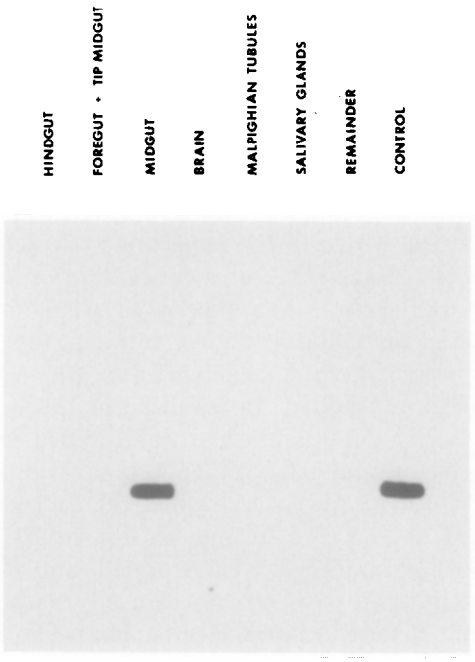

FIG. 1. Tissue specificity of RNA homologous to pkdm26H2. A single, late third-instar larva was dissected in Robb's insect medium in a glass well and isolated organs were placed in microfuge tubes containing medium. The hindgut was severed from the rest of the gut at a position slightly posterior to the region at which the Malpighian tubules join; a small amount of tissue at the extreme posterior end may have been lost during dissection. The foregut included, in addition to the esophagus and the proventriculus, the gastric caecae and most of the stomach, which are properly considered as the anterior tip of the midgut (Bodenstein, 1965). The track labeled "Midgut," then, consists of the remainder of the midgut, extending to a position slightly anterior to the region at which the Malpighian tubules join the gut; a small amount of contaminating tracheal tissue is also included. Following the isolation of the salivary glands, virtually all remaining material, including the carcass and the liquid in the well, were removed and included in the sample labeled "Remainder." As a control, another late third-instar larva was surgically opened in Robb's medium, but no dissection was attempted; rather, the opened larva and surrounding medium were transferred immediately to a microfuge tube. RNA extraction, electrophoresis, transfer to DBM paper, and hybridization were as described in Carlson and Hogness (1985).

paper, and hybridized with ^{32}P-labeled pkdm26H2 DNA. The surprising result is that *Jonah* RNA was detected only in the midgut and control tracks, the latter consisting of the RNA from an undissected third-instar larva. No *Jonah* RNA was detected in the salivary glands even when, in another experiment, several hand-dissected glands were pooled and their RNA was examined in this manner. Given these results, it is evident that the *Jonah* RNA which is present in the mass-isolated salivary gland preparations, and from which the group VI cDNA clones derive, arises from small amounts of midgut in the preparations.

Jonah cDNA Sequences are Repeated and Dispersed within the D. melanogaster Genome

When total embryonic DNA from our Oregon-R stock was digested to completion by each of four restriction endonucleases and the digests were fractionated by gel electrophoresis, blot transferred to nitrocellulose paper, and hybridized with a ^{32}P-labeled *Jonah* cDNA probe, the autoradiogram shown in Fig. 2 was obtained. Approximately 11–14 bands are visible in each track; these bands vary widely in intensity, and on longer exposure several more can be seen in each track. The same result was obtained whether the hybridization probe was the reference cDNA clone, pkdm26H2, or another member of group VI, the pkdm8G8 clone used in Fig. 2. The dm8G8 and dm26H2 cDNA segments share a common restriction map and are homologous in cross-hybridization experiments (Carlson and Hogness, 1984); pkdm8G8 sequences have therefore sometimes been used in place of pkdm26H2 in the probing of the *Jonah* genes.

Figure 3 shows that the *Jonah*-containing restriction

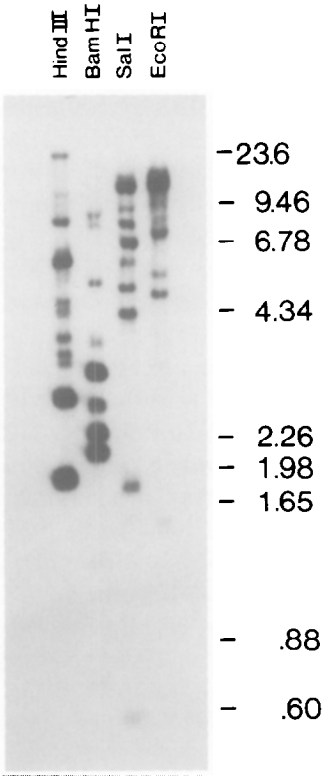

FIG. 2. Genomic DNA fragments containing *Jonah*-homologous sequences. Approximately 3 μg of embryonic DNA from our Oregon-R stock was digested with the indicated restriction endonucleases, electrophoretically fractionated on a 0.5% agarose gel in Tris-borate-EDTA buffer, transferred to nitrocellulose, and hybridized with a ^{32}P-labeled *Jonah* cDNA probe, as described under Materials and Methods.

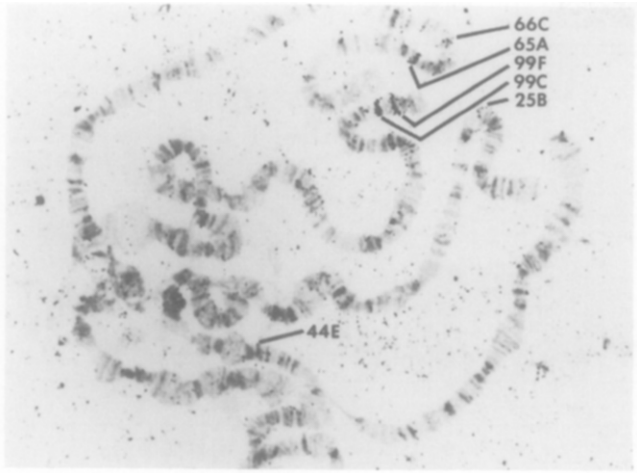

FIG. 3. *In situ* hybridization of *Jonah* cDNA to salivary gland polytene chromosomes. The cDNA clone pkdm26H2 was labeled with ^{3}H by nick translation and hybridized to salivary gland *gt 1/gt x-11* chromosomes as described under Materials and Methods. Exposure: 16 weeks.

fragments identified in Fig. 2 are dispersed among several chromosomal sites. Thus, *in situ* hybridization of ^{3}H-labeled pkdm26H2 to salivary gland polytene chromosomes reveals two major sites at 99C and 99F in 3R (i.e., the right arm of chromosome 3), of which 99C generally exhibits the stronger label. Four minor sites are observed in Fig. 3 at 25B in 2L, 44E in 2R, and 65A and 66C in 3L. Longer exposures reveal another site at 67B in 3L (Wolfner, 1980).

The observation that the number of bands per track (Fig. 2) is consistently and considerably greater than the number of chromosomal sites suggests that the *Jonah* genes are clustered at one or more of the sites. The variation in the intensity of cDNA hybridization both among the chromosomal sites and among the bands could result from variation in the number and topography of clustered genes at different sites. Alternatively, or additionally, structural heterogeneity among the *Jonah* genes might cause this differential hybridization. As will be apparent from the following analysis of the genomic DNA clones, the *Jonah* genes are heterogeneous both in structure and in the gene number and topography of clustering.

Isolation and Classification of Genomic Clones

Primary genomic clones containing *Jonah* RNA-homologous sequences were obtained by plaque-hybridization screening (Benton and Davis, 1977) of a library of λbDm recombinant phage (Meyerowitz and Hogness, 1982) with ^{32}P-labeled sequences from the pkdm8G8 *Jonah* cDNA clone (Materials and Methods). The λbDm recombinants consist of randomly sheared segments of embryonic DNA (Dm segments) from our stock of the Oregon-R wild-type strain of *D. melanogaster* that were inserted by $(dA)_n \cdot (dT)_n$ joints into the "λb" vector (the λSep6 phage of Meyerowitz and Hogness, 1982). Almost all of the primary genomic clones used in our analysis of the *Jonah* genes were obtained in a single screen of 5×10^4 plaques that yielded 25 λbDm clones containing *Jonah* RNA-homologous sequences.

Cleavage of the DNA from the 25 λbDm phage with *Xho*I and *Bam*HI and analysis of the resulting restriction fragments by gel electrophoresis, blot transfer, and hybridization with the *Jonah* cDNA probes divided the phage into the seven classes shown in Fig. 4, where the sequence overlap between Dm segments within a class was confirmed by mapping additional restriction sites. *In situ* hybridization of one or more representatives of each class to polytene chromosomes revealed a single major site of hybridization which was used to designate the class (Fig. 4).

The seven classes derive from six of the seven chromosomal sites previously identified by *in situ* hybridization with the reference cDNA probe; no genomic clones from the weak 67B site were identified and two of the seven classes (99Cα and 99Cβ) derive from the 99C site. Secondary sites of hybridization were identified with most of the λbDm phage and these were generally subsets of the major hybridization sites. The λbDm2309 phage of the 99Cα class is an exception, as it hybridized to approximately 20 additional sites. Furthermore, when ^{32}P-labeled λbDm2309 DNA was hybridized to an electrophoretically fractionated *Hin*dIII digest of total embryonic DNA, it labeled many bands in addition to those labeled by the *Jonah* cDNA probe (Fig. 2). Further analysis (summarized in Fig. 7) of these additional repetitive sequences indicates that they derive from a dispersed repeated element that occupies ≥7 kb at the left end of the 99Cα locus, and which we suspect is one of the numerous transposable elements that populate the *D. melanogaster* genome (reviewed by Spradling and Rubin, 1981).

The Topography of Clustering

The positions and orientations of 17 *Jonah* genes are indicated by the filled arrows above the restriction maps for each of the 7 chromosomal arrangements defined in Fig. 4. Only one locus (44E) contains a single gene; the others contain clusters of multiple genes.

Following the convention used for the *D. melanogaster* actin genes (Zulauf *et al.*, 1981; Fyrberg *et al.*, 1983), we denote an individual gene according to its chromosomal position, modifying that convention to accom-

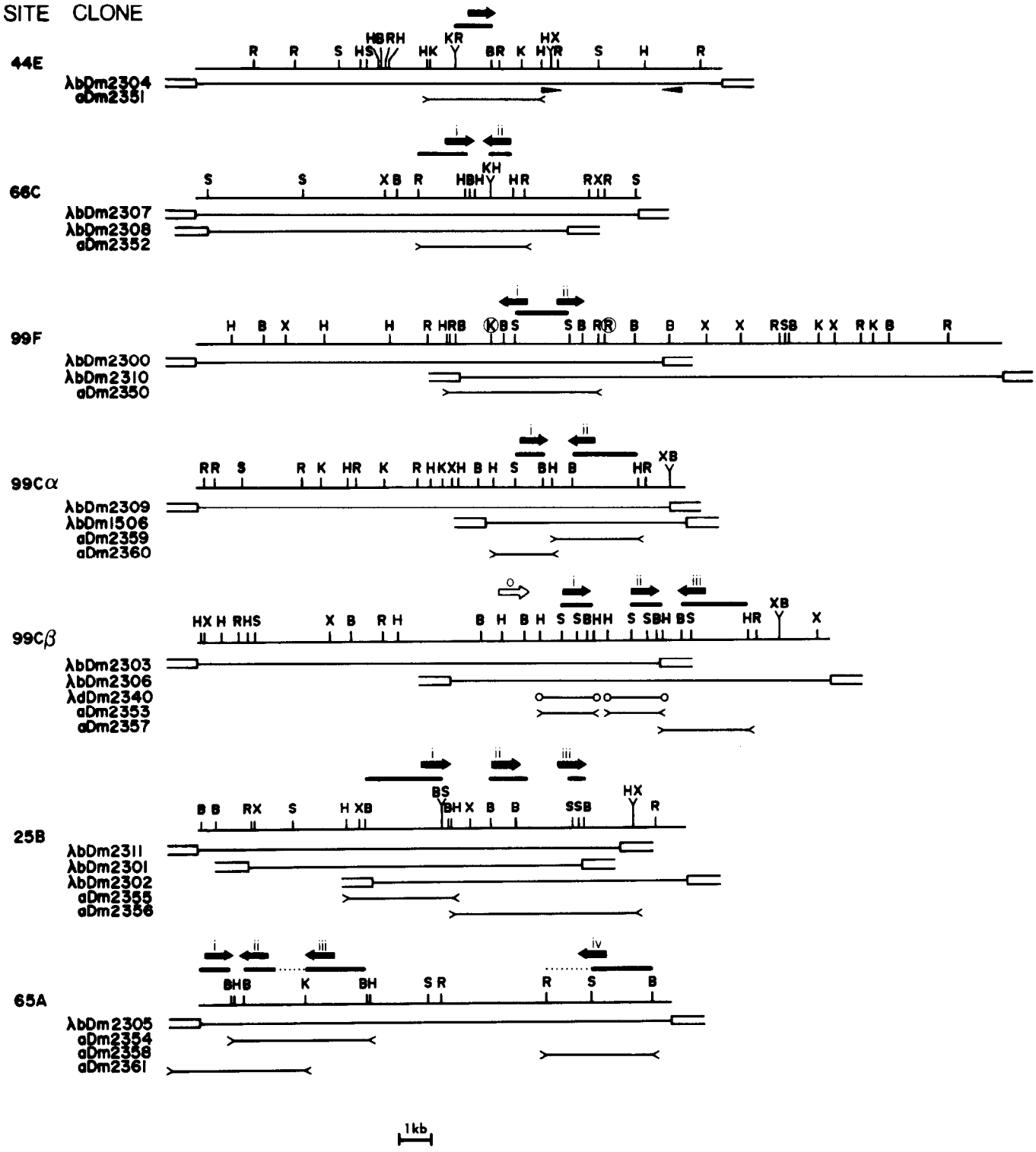

FIG. 4. *Jonah* gene arrangements at different chromosomal sites. The restriction map above the overlapping Dm segments from the indicated chromosomal sites applies to each segment with the exception of the two circled restriction sites on the 99F map, which are present in Dm2310 but not Dm2300. The arms of the λ vector DNA are shown by open bars, with the left λ arm on the left for all clones except λbDm1506, λbDm2303, and λbDm2301, where it is on the right. The λbDm phage hybrids are primary clones obtained by screening the genomic library with *Jonah* cDNA probes, while the λdDm phage and aDm plasmid hybrids were obtained as subclones of the primary clones (see text). The λdDm2340 and aDm2353 subclones in the 99Cβ group are shown in two positions because this region appears to be tandemly repeated (see text). Each arrow above a restriction map represents the position of a *Jonah* gene, while the associated Roman numerals distinguish among the genes in a cluster. See the text for methods of mapping these positions and for the definition of the black bars below the arrows. The oppositely oriented arrowheads shown below the Dm2304 segment from 44E represent an inverted repeat that does not contain *Jonah* sequences. The dotted lines indicate regions of Dm2305 in which additional hybridization was observed with adm135A8. B, *Bam*HI; H, *Hin*dIII; K, *Kpn*I; R, *Eco*RI; S, *Sal*I; X, *Xho*I.

modate clustering by adding a lowercase Roman numeral to distinguish among genes in a cluster. Thus, the single gene at 44E is *Jon44E*, while the two genes in the 66C cluster are *Jon66Ci* and *Jon66Cii*, the three genes in the 99Cβ cluster are *Jon99Cβi*, *Jon99Cβii*, and *Jon99Cβiii*, etc.—the Roman numerals being assigned in the order shown in Fig. 4. We have not attempted to define a locus in the polytene chromosomes beyond the lettered regions (44E, 66C, etc.) because the resolutions of *in situ* hybridization by autoradiography seldom allows assignment of a DNA sequence to an individual band within a lettered region (Spierer *et al.*, 1983). Furthermore, different clusters that map to the same lettered region, such as the 99Cα and 99Cβ clusters, may not occupy different positions within the region, but may instead represent polymorphic variants, as appears to be the case in this instance (see below). We have therefore adopted the pragmatic convention of distinguishing among clusters from the same region by assigning a Greek letter to each.

The following parts of this section describe the four lines of evidence that we have used to delineate the topography of the *Jonah* genes shown in Fig. 4.

(i) Restriction mapping and cDNA hybridization. The DNAs of representative λbDm clones defining each locus were digested with six restriction endonucleases, singly and in all possible pairs, to yield fragments whose lengths were used to deduce the restriction maps that encompass each locus (Fig. 4). Blot transfer to nitrocellulose of electrophoretically fractionated fragments and subsequent hybridization with ^{32}P-labeled sequences from the pkdm26H2 and pkdm8G8 cDNA clones provided a preliminary map of the *Jonah* RNA-homologous regions at each locus.

Restriction fragments containing these RNA-homologous regions were then subcloned by insertion into either the "λd" phage vector (the λ590 of Murray *et al.*, 1977) or the "a" plasmid vector (the pBR322 of Bolivar *et al.*, 1977) to yield the λdDm and aDm subclones indicated in Fig. 4. Digestion of these subcloned DNAs with seven additional restriction endonucleases provided supplemental restriction maps of which five are shown in the accompanying paper (Carlson and Hogness, 1985). Localization of the *Jonah* RNA-homologous regions was, in some cases, further refined by probing fragments defined by these maps with cDNA sequences, yielding the positions for these regions indicated by the thick lines shown below the arrows in Fig. 4.

Comparison of the restriction maps for genomic and cDNA segments was used for more precise localization of several genes. Given the structural heterogeneity within the *Jonah* family, this method is limited by the variety of cDNA clones that have been isolated. For example, both the reference cDNA segment, dm26H2, and the almost identical dm8G8 segment exhibit restriction maps that are exactly overlapped by a pattern of restriction sites found only in the three RNA-homologous regions of the 99Cβ cluster (Carlson and Hogness, 1985), and which can therefore be used for determining the position of only the *Jon99Cβi*, *Jon99Cβii*, and *Jon99Cβiii* genes. Several other cDNA clones were defined in studies on the expression of the *Jonah* family reported in the accompanying article (Carlson and Hogness, 1985). Their restriction maps correspond to, and were therefore used to map the positions of, the *Jon25Bi*, *Jon65Ai*, *Jon65Aiv*, *Jon99Cαi*, and *Jon99Cαii* genes.

Seven of the above eight genes (i.e., all but *Jon65Aiv*) contain a *Bam*HI site ~0.1 kb from their 3′ ends. (Determination of gene orientation is described below; the positions of the 3′ ends were approximately determined from the 3′ ends of the corresponding cloned cDNA segments.) Similarly, seven of the nine remaining RNA-homologous regions exhibit a *Bam*HI site at or near the 3′ ends of these regions (Fig. 4). Consequently, the 3′ ends of these seven genes were assumed to lie 0.1 kb beyond these *Bam*HI sites, yielding gene positions consistent with those determined from electron microscopic examination of heteroduplexes that are described below. The positions of the last two genes (*Jon66Cii* and *Jon65Aiii*) within their RNA-homologous regions were determined from such heteroduplex maps. The lengths of all 17 genes shown in Fig. 4 were assumed to be 0.9 kb, a value taken from the length of the *Jonah* RNA and the lengths of homology regions observed in heteroduplexes and determined from the analysis of isolated inverted repeats described below.

(ii) Isolation and analysis of inverted repeats. Electron microscopy of renatured single strands of λbDm phage DNAs revealed a single inverted repeat in six of the seven arrangements; only the 25B locus (represented by λbDm2311) contains none. Five of these six inverted repeats contain *Jonah* RNA-homologous sequences and, indeed, appear to consist of oppositely oriented pairs of *Jonah* genes in each of five clusters; 44E, the only locus represented by a single gene, contains the only inverted repeat lacking *Jonah* sequences (Fig. 4). These conclusions derive both from electron microscopic mapping of the inverted repeats and from their isolation and analysis.

Each of seven λbDm phage DNAs representing the seven arrangements was denatured and then allowed to renature during a period determined to be too short for appreciable bimolecular renaturation but long enough to allow unimolecular annealing of the inverted repeats (Materials and Methods). The single strand regions in these strands were then digested with S1

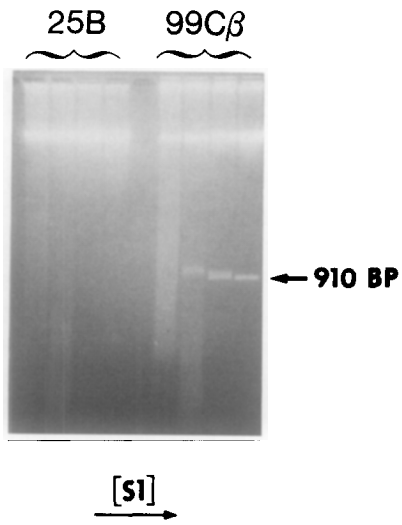

FIG. 5. Isolation of inverted repeats by brief renaturation and S1 digestion. Two genomic *Jonah* clones from 25B (λbDm2311) and 99Cβ (λbDm2306) were denatured, briefly renatured, treated with increasing concentrations of S1, and electrophoretically fractionated, as described under Materials and Methods.

nuclease, leaving intact the duplex regions consisting of the inverted repeats, which were isolated by agarose gel electrophoresis and detected with ethidium bromide.

Figure 5 shows the results of two such experiments: one with the λbDm2311 DNA from the 25B locus and the other with λbDm2306 from 99Cβ. Both DNAs contain multiple RNA-homologous regions but only the 99Cβ DNA contains an inverted repeat observable in the electron microscope. Similarly, only the 99Cβ DNA yields a duplex fragment which is registered in Fig. 5 as a band that becomes progressively sharper with increasing S1 concentration and consists of duplexes 0.91 kb long. The agreement between this length and that of the *Jonah* RNA suggests that the inverted repeat consists of oppositely oriented *Jonah* genes in the 99Cβ cluster. This suggestion is supported by the observation that the 0.91-kb fragment hybridized with the reference cDNA probe after blot transfer to nitrocellulose paper. It is substantiated by electron micrographic mapping of this inverted repeat in heteroduplexes (e.g., see Fig. 6C), which demonstrated that members of the repeat occupy positions corresponding to those of the *Jonah* genes mapped above; more specifically, one member is always *Jon99Cβiii*, while the other member is either *Jon99Cβii* or *Jon99Cβi* in a ratio of approximately 25:1, respectively (see below). The length of the inverted repeat observed in the microscope (1.05 ± 0.20 kb; $N = 15$) is in reasonable agreement with that determined by S1 analysis, given the high standard deviation resulting from poor definition of the boundary between the stem and small loop of the hairpins formed by the adjacent *Jon99Cβii* and *iii* genes, which comprise most of the hairpins in the measured sample.

The λbDm DNAs from four of the five other arrangements also generated S1-resistant fragments similar in length to the *Jonah* RNA and containing sequences homologous to it; thus, DNA from 65A (λbDm 2305) and 66C (λbDm 2307) generated fragments of 0.82 and 0.96 kb, respectively, while both 99Cα (λbDm 2309) and 99F (λbDm 2300) produced 1.0-kb fragments. When combined with the data for the 99Cβ repeat (Fig. 5), these results yield a mean length and standard deviation of 0.94 ± 0.08 kb for the S1-resistant, RNA-homologous fragments from all five arrangements. We hesitate to interpret differences among these values as reflecting structural variation among the *Jonah* genes comprising the inverted repeats since these small differences could be due to experimental error resulting from the sensitivity of the electrophoretic mobilities to the level of S1 treatment (Fig. 5). In the case of the 65A and 99F DNAs, however, we observed a smaller secondary band of 0.48 and 0.88 kb, respectively, which may reflect regions of mismatch between the two genes comprising each of these inverted repeats.

Microscopic mapping of the positions and lengths of these four inverted repeats is consistent with the S1 analysis in indicating that they derive from pairs of oppositely oriented *Jonah* genes. These pairs consist of the only two genes in the 66C, 99Cα, and 99F clusters, and of the adjacent *Jon65Ai* and *ii* genes in 65A, the other two genes of this cluster not having been observed in the eight hairpins of the measured sample. The stems comprising these and the 99Cβ inverted repeats appear to be continuous duplexes and give no indication of looped regions of nonhomology, which would be detectable in these experiments were they ≥0.1 kb long.

The remaining locus, 44E, exhibits an inverted repeat in the electron microscope (Fig. 6B) that is 0.66 ± 0.12 kb ($N = 7$) long and does not contain *Jonah* sequences, as determined both by microscopic mapping (Fig. 4) and by the absence of S1-resistant fragments that hybridize with the reference cDNA probe.

(iii) Orientation of transcription. The preceding results indicate the relative orientations of the genes in a given cluster, but do not determine the orientation of transcription. This orientation was delineated for the three genes of the 25B cluster and for the first two genes of the 99Cβ cluster by directly determining the 5′ to 3′ orientation of the DNA strand that contains the *Jonah* RNA sequence, using the procedure described by Muskavitch and Hogness (1980). Briefly, the strands of λbDm2302 (containing the three commonly oriented *Jon25B* genes) and λdDm2340 (containing *Jon99Cβi* or *Jon99Cβii*; Fig. 4) were separated by gel electrophoresis,

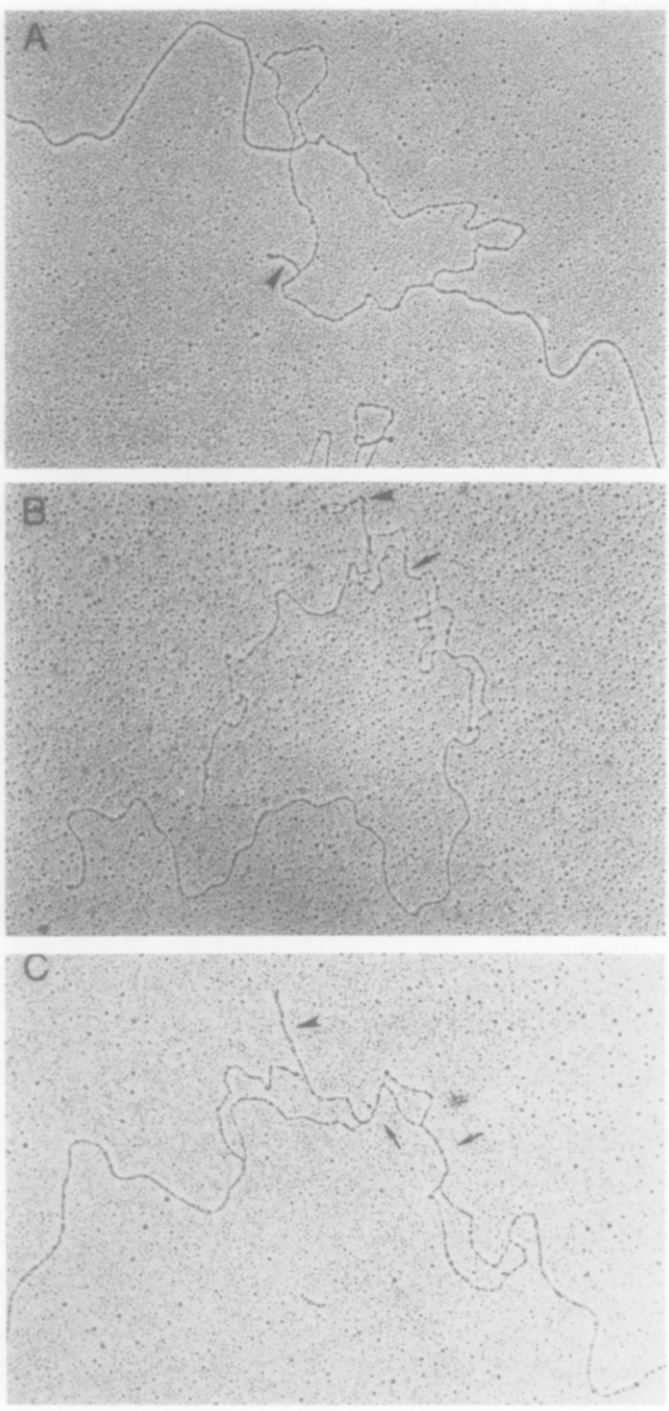

FIG. 6. Heteroduplexes between a 25B clone and test clones from 44E, 66C or 99Cβ. Heteroduplexes were formed and analyzed according to the procedure of Davis *et al.* (1980). The clone from 25B was λbDm2302 in all three cases. (A) The test clone is λbDm2307 from 66C; the arrowhead indicates the inverted repeat in Dm2307 consisting of its two *Jonah* genes. (B) The test clone is λbDm2304 from 44E. The arrowhead indicates the non-*Jonah* inverted repeat, while the arrow indicates duplex formation between *Jonah* genes in the two Dm segments. The other, shorter region where the strands are aligned has arisen through random crossing of the strands as opposed to the annealing of two strands. (C) The test clone is λbDm2306 from 99Cβ. The arrowhead indicates an inverted repeat formed by oppositely oriented *Jonah* genes in Dm2306; no loop is visible in this micrograph. The two arrows indicate duplex regions formed between *Jonah* genes from the two Dm segments.

blot transferred to nitrocellulose paper, and hybridized with two ^{32}P-labeled probes: one to identify the strand that contains the *Jonah* RNA sequence (single-stranded cDNA copies of the poly(A)$^+$ RNA from third-instar larvae), and one to identify the 5′ to 3′ orientation of the strands relative to the Fig. 4 restriction maps (a strand-specific λ probe). The results of these hybridizations demonstrate that *Jon99Cβi* and *ii*, and all of the *Jon25B* genes are transcribed from left to right, as is indicated in Fig. 4 by the directions of the filled arrows. Since the microscopic mapping of the inverted repeat in 99Cβ showed that *Jon99Cβiii* is oppositely oriented to the above two 99Cβ genes, it must be transcribed from right to left.

The orientations of *Jon44E*, *Jon65Aiii*, and *Jon65Aiv* were determined by examining heteroduplexes between a DNA containing one or more of these genes (λbDm2304 or λbDm2305) and one of the above two phage DNAs. The restriction site homologies shown in Fig. 7 for the right-hand parts of 99Cα and 99Cβ indicate that *Jon99Cαi* and *ii* are transcribed in the same directions as *Jon99Cβii* and *iii*, respectively. The six remaining genes form the inverted repeats observed in 65A, 66C, and 99F, and their orientations of transcription were assigned by assuming that the single *Bam*HI site present in five of them marks their 3′ ends. Although this assumption is supported by the orientations determined for the other genes, these last assignments are clearly less sure.

(iv) Heteroduplex analyses of sequence homologies among the loci. Heteroduplexes formed from a dozen pairs of primary genomic λbDm clones have been analyzed in the electron microscope. With two exceptions, each consisted of a reference strand from the 25B locus (λbDm2301 or 2302) against which a strand from another locus was tested for sequence homologies. [The exceptional heteroduplexes were formed between the above two clones from 25B, and between clones from 99Cβ (λbDm2306) and 66C (λbDm2308)]. The test strands were from the 44E (λbDm2304), 65A (λbDm2305), 66C (λbDm2307 and 2308), 99Cβ (λbDm2303 and 2306), and 99F (λbDm2300) loci. These heteroduplexes yielded structures consistent with the maps given in Fig. 4. Three examples are considered here. Figure 6A illustrates the simple structures obtained when the test strand contains a single oppositely oriented pair of *Jonah* genes, as is the case for the 66C DNA shown here. No regions of interstrand homology are seen because the two genes in the test strand pair

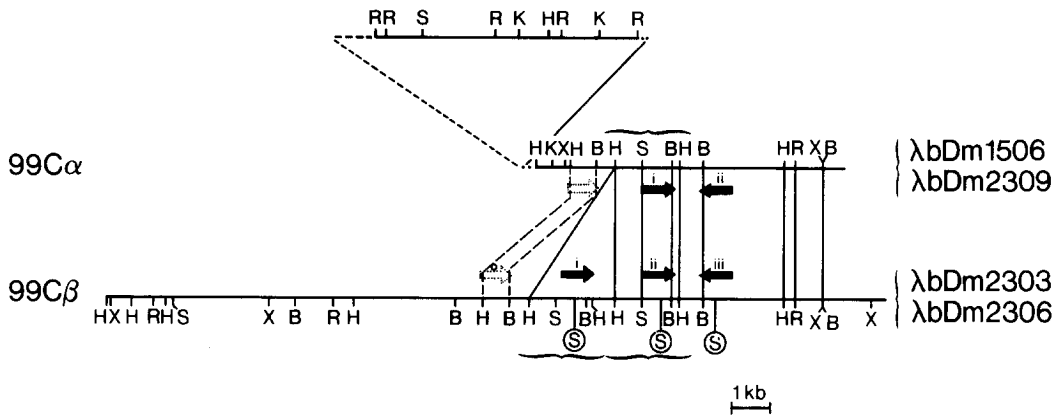

FIG. 7. Two arrangements at 99C. The restriction maps and arrows indicating transcription units are drawn as in Fig. 4; an open dashed arrow is also drawn at 99Cα at a position analogous to that of *Jon99Cβo*. Those restriction sites found in the transcription units of the 99Cβ arrangement but not in those of the 99Cα arrangement are circled. The brackets show the positions of the 1.8-kb *Hin*dIII fragments, which apparently constitute part of a direct repeat. The dotted lines connect restriction sites to the left of the 1.8-kb *Hin*dIII fragment which are shared by the 99Cα and 99Cβ arrangements. The raised line represents sequences contained within the dispersed repeated element (see text). The right endpoint of the element has not been determined precisely; the dotted extension at the right end of the repetitive element and the left end of the 99Cα line below it mark the region in which the right end of the element lies. The dashed extension at the left end of the repetitive element is drawn to indicate uncertainty as to the position of its left end; genomic blotting experiments indicate only that this terminus lies to the left of the *Eco*RI site nearest the *Sal*I site in the repetitive element. B, *Bam*HI; H, *Hin*dIII; K, *Kpn*I; R, *Eco*RI; S, *Sal*I; X, *Xho*I.

preferentially with each other to form the hairpin seen in this figure and analyzed above.

The test strand of the heteroduplex in Fig. 6B is from 44E and contains a single *Jonah* gene whose position corresponds to that of the single region of interstrand homology seen in these heteroduplexes. This region occupies different positions in the reference strand that correspond to the three *Jon25B* genes which it contains. (Note that the test and reference strands are distinguished by the *Jonah*-independent inverted repeat in the 44E test strand.) The length of the interstrand region is 0.8 ± 0.29 kb (N = 6), and exhibits a high standard deviation, presumably, because of differences among the three *Jon25B* genes.

A more complex set of structures was generated between a 99Cβ test strand (λbDm2306) and the λbDm2302 reference strand. As noted earlier, this set can be divided into two classes according to the kind of hairpin formed by the three genes in the test strand. Thus, among 49 interpretable heteroduplexes out of 53 examined, 47 exhibited hairpins in which the stem position corresponds to the adjacent *Jon99Cβii* and *iii* genes, while in the remaining 2 it corresponds to *Jon99Cβi* and *iii*, with *Jon99Cβii* in the loop. The majority class divides into two subclasses: one consists of 19 heteroduplexes containing one region of interstrand homology, and the second consists of 28 heteroduplexes containing *two* such regions, an example of which is shown in Fig. 6C.

The second subclass was entirely unexpected because two of the three 99Cβ genes in the test strand are paired in the hairpin, leaving only one to pair with the genes in the reference strand. Measurements of micrographs such as that in Fig. 6C show that the 0.85 ± 0.13-kb (N = 10) region of interstrand homology proximal to the hairpin does indeed map to the position of the *Jon99Cβi* gene in the test strand. However, the second, distal region of homology, which also exhibits a length within the range expected of a *Jonah* gene (0.77 ± 0.19 kb; N = 10), maps 0.92 ± 0.14 kb (N = 11) to the left of *Jon99Cβi*, at a position where we failed to detect *Jonah* RNA-homologous sequences by hybridization with the cDNA reference probe, pkdm26H2 (Fig. 4). The existence of this second region of homology indicates either a *Jonah*-independent element that is present in both the 99Cβ and 25B clusters, or a new *Jonah* gene in the 99Cβ cluster whose structure is sufficiently different from the other members of this cluster that it escaped detection by filter hybridization with the pkdm26H2 probe, yet anneals with one or more of the *Jon25B* genes under the conditions of heteroduplex formation.

The position of the two regions of homology in the 25B reference strand, λbDm2302, strongly favors the proposition that the 99Cβ cluster contains a fourth *Jonah* gene. Two map distributions of approximately equal frequencies were obtained for these positions. In one, the positions of the two regions correspond to the *Jon25Bi* and *ii* genes, while in the second, they correspond to *Jon25Bii* and *iii*. Thus, the distances from the right end of the reference Dm2302 segment to the right or 3′ ends of the *Jon25Bi*, *ii*, and *iii* genes are 7.5, 5.4, and 3.2 kb, respectively, as determined by restriction and cDNA hybridization mapping (Fig.

4), while the distances to the right ends of the two homology regions in the first distribution are 7.6 ±0.36 kb and 5.5 ± 0.33 kb, and in the second distribution are 5.6 ±0.62 kb and 3.1 ±0.33 kb, as determined by measuring 10 heteroduplexes. Moreover, the region in Dm2302 to which the *Jon99Cβi* gene anneals in the first distribution (5.5 kb on the above scale) is coincident with that to which the novel 99Cβ element anneals in the second distribution (5.6 kb)—a conclusion that is not dependent upon a knowledge of the positions of the *Jon25B* genes.

In view of these results, we have included the novel element within the *Jonah* family, denoting it as *Jon99Cβo* and indicating its position and orientation by an open arrow in Fig. 4. This inclusion implies an expansion of the family definition so as to embrace chromosomal DNA regions of appropriate length that do not hybridize with the reference cDNA under standard conditions, but can anneal with other members of the family that do—an expansion that we shall consider further in the last section of these Results and in the Discussion.

Polymorphism Accounts for the Two Clusters at 99C

A comparison of the restriction maps for the 99Cα and 99Cβ DNAs given in Fig. 7 suggests that they represent allelic variants of a single polymorphic locus rather than different loci in the same chromosome. In examining these two maps for equivalent restriction site topographies, we first note that the left half of 99Cα consists of DNA from a dispersed repeated element inserted into this region but not into 99Cβ; hence, we expect no equivalence here and restrict our examination to the right half, or more specifically to the 99Cα DNA including and extending to the right of the *Hin*dIII site that is adjacent to the central *Xho*I site of *99Cα*. The equivalent region in 99Cβ includes and extends to the right of the *Hin*dIII site adjacent to the 5′ end of the unusual *Jon99Cβo* gene, the two *Hin*dIII sites being connected by a dotted line in Fig. 7.

The principal difference between these two regions is the apparent tandem duplication in 99Cβ of a unit that appears singly in 99Cα. The duplication was first recognized by the observation that *Jon99Cβi* and *ii* are contained in two apparently identical 1.8-kb *Hin*dIII fragments of the same orientation that are separated by 0.5 kb, suggesting that this pair of genes was generated by tandem duplication of a 2.3-kb unit from which the 1.8-kb fragments derive. The *Jon99Cαi* gene is also contained in a 1.8-kb *Hin*dIII fragment with a restriction map equivalent to that for the duplicated 1.8-kb fragments of 99Cβ, with the exception of the circled *Sal*I site uniformly found in *Jon99Cβi*, *ii*, and *iii*, but not in *Jon99Cαi* or *ii* (Fig. 7). Indeed, if one ignores these *Sal*I sites, then conceptual deletion of either one of the tandemly duplicated units makes the restriction map for the right half of 99Cβ equivalent to that for the right half of 99Cα. This equivalence not only indicates 99Cα and 99Cβ have a common origin, but suggests that 99Cα also contains a *Jonah* gene equivalent to *Jon99Cβo* located near the insertion site of its dispersed repeated element.

To test whether 99Cα and 99Cβ are allelic variants located on different chromosomes in a polymorphic Oregon-R population, or derive from different closely linked loci in the same chromosome, we isolated DNA from single flies in our Oregon-R stock, digested this DNA with *Eco*RI, separated the fragments by gel electrophoresis and, after blot transfer to nitrocellulose, probed for fragments that distinguish and are characteristic of the 99Cα and 99Cβ DNAs. The results indicate that some of the tested flies contain the 99Cα, but not the 99Cβ DNA, while others exhibit the reciprocal distribution, indicating that each arrangement is found on a different chromosome rather than both on the same chromosome; we therefore presume that they are allelic. In the same vein, we suspect, but have not similarly demonstrated, that the two clones from 99F (λbDm2300 and λbDm2310; Fig. 4) derive from allelic variants because the restriction maps for the 6.5-kb region of their overlap are identical except for the two circled sites that are present in λbDm2310 but not λbDm2300.

A cDNA Clone from Jon65Aiv Reveals Jonah Sequences at a New Chromosomal Locus

In the accompanying paper (Carlson and Hogness, 1985), we describe *Jonah* cDNA clones which exhibit restriction maps and primary sites of *in situ* hybridization to polytene chromosomes different from those of the reference cDNA clone. Two of the new clones, adm135A8 and adm2372, contain dm segments with identical restriction maps that indicate they derive from *Jon65Aiv*. Consistent with this assignment, 65A is the primary site of *in situ* hybridization for both clones, and within the 65A DNA shown in Fig. 4, the 3.3-kb *Eco*RI-*Bam*HI fragment encompassing *Jon65Aiv* hybridizes most strongly with this class of cDNA clones.

The adm135A8 clone was isolated by Wolfner (1980), who showed that in addition to the primary 65A site, it hybridized weakly to 25B, 44E, 66C, 67B, 99C, and 99F, as expected for a *Jonah* cDNA. However, it also hybridized to 74E with an intermediate intensity—a quite unexpected result since no *Jonah* clone had previously been observed to label this novel site. In particular, the reference cDNA clone, pkdm26H2, and the equivalent pkdm8G8 clone, failed to label 74E in

an extensive set of experiments that included autoradiographic exposures much longer than that required to detect 74E labeling by adm135A8.

The simplest explanation of these results is that the 74E site contains one or more *Jonah* genes so divergent in sequence from the genes represented by the reference cDNA (i.e., *Jon99Cβi, ii,* or *iii;* Carlson and Hogness, 1985) that the two fail to anneal under the conditions of *in situ* hybridization, and that on this scale of sequence divergence, the *Jon65Aiv* gene occupies an intermediate position, exhibiting a weak but sufficient homology to the 74E and 99Cβ genes that its cDNA hybridizes to both. The following experiments support this explanation as opposed to the alternative hypothesis that the adm135A8 and adm2372 cDNAs derive not from *Jon65Aiv* but from a closely linked gene of a quite different family that cohabits the seven *Jonah* chromosomal loci and is the lone inhabitant of the novel 74E site.

Two types of hybridization experiment indicate that the adm135A8 cDNA from 65A does, in fact, exhibit weak homology with cDNAs from the *Jon99Cβi, ii,* or *iii* genes. Heteroduplexes were formed between adm135A8 and a cDNA clone from the *Jon99Cβi, ii,* or *iii* genes (adm2373; Carlson and Hogness, 1985) that was chosen instead of the reference cDNA, pkdm26H2, because its vector sequences are the same as those in adm135A8. Formed under nonstringent conditions, the heteroduplexes were spread under conditions of different stringencies. When spread under the standard conditions of 40/10% formamide (Davis *et al.*, 1980), no unpaired single-stranded regions were observed, but when spread under the more stringent 80/50% formamide conditions (Davis and Hyman, 1971), a denaturation bubble appeared in the approximate region of the cDNA segments. Similarly, in filter hybridizations using ^{32}P-labeled pkdm26H2 as a probe under a range of conditions of different stringencies, weak homology was detected to the excised dm135A8 segment and to a restriction fragment contained within this segment. On the basis of these and other experiments, we estimate that the two cDNAs segments exhibit a sequence mismatch of ~15%. Finally, we note that our isolation of the adm135A8-equivalent clone, adm2372, was effected by screening a cDNA library for clones that will hybridize with the 1.8-kb *Hin*dIII fragment that contains the *Jon99Cβi* or *ii* genes (Carlson and Hogness, 1985).

These results indicate that sequences in the gene from which the adm135A8 and adm2372 cDNAs derive exhibit a sufficient homology to those in the reference cDNA to justify the inclusion of this gene in the *Jonah* family as *Jon65Aiv.* The question of expanding the definition of the *Jonah* family to include genomic elements such as that at 74E and *Jon99Cβo* is considered in the Discussion.

DISCUSSION

The *Jonah* genes constitute a complex family of approximately 20 imperfectly repeated genes distributed in small clusters at eight or more widely distributed chromosomal sites. This pattern of organization combines the characteristics of two types of gene family organization: the clustered pattern manifested by the 100 copies of the *Drosophila* histone genes, tandemly repeated at a single site (Lifton *et al.*, 1978), and the 6 *Drosophila* actin genes, singly dispersed at six chromosomal sites (Fyrberg *et al.*, 1980; Tobin *et al.*, 1980). If a particular structural feature endows a gene with the ability to proliferate at a given site, and if another structural feature enables a gene to become dispersed at distant sites, then perhaps an ancestral *Jonah* gene acquired both features, and the capacity to proliferate *in situ* was maintained following dispersion events.

We have designated those genomic sequences homologous to the reference *Jonah* cDNA probe as "genes"; however, the relationship between the number of RNA-homologous regions and the number of *Jonah* genes deserves comment. Although the location of two RNA-homologous regions in an inverted repeat indicates the presence of two distinct *Jonah* genes, the location of two RNA-homologous regions in a direct repeat could in principle represent either two distinct *Jonah* genes or two exons of a single gene. The observation that multiple structures arise between certain pairs of Dm segments in heteroduplex experiments and the existence of a common *Bam*HI site near the 3′ ends of 14 *Jonah* RNA-homologous regions suggest that the number of *Jonah* RNA-homologous regions shown in Fig. 4 is a good approximation to the number of *Jonah* genes.

Structural variation among *Jonah* sites is extensive. Gene clusters are heterogeneous in gene number and in topography of clustering, and structural heterogeneity among individual *Jonah* genes is revealed both by hybridization experiments and by restriction site heterogeneities, shown in Fig. 4 and in greater detail by restriction mapping of the kind shown in Fig. 4 of the accompanying paper (Carlson and Hogness, 1985). This mapping has shown that some *Jonah* genes share common restriction site patterns; in all such cases, genes bearing common patterns are clustered at the same site. Examples include the *Bam*HI-*Sal*I pattern of *Jon99Cβi, ii,* and *iii,* the distinct *Bam*HI-*Bgl*I-*Sal*I pattern of *Jon99Fi* and *ii,* and the *Bam*HI-*Bgl*I pattern of *Jon99Cαi* and *ii.* Perhaps local duplication events

have occurred more recently than distant dispersion events, or dispersion may be accompanied by changes in gene structure.

Not all genes at a site are homogeneous in structure; we have identified a *Jonah* gene in the 99Cβ cluster, *Jon99Cβo*, which does not hybridize with the reference cDNA under standard conditions, but can hybridize to other genes that do. We have also presented evidence that the cDNA clone adm135A8 represents a *Jonah* gene, designated *Jon65Aiv*, and has diverged sufficiently in sequence that it does not show homology to the reference cDNA under standard conditions; we have observed that adm135A8 shows homology to one or more genes at another site, 74E, which had not been detected using other *Jonah* probes. The identification of poorly conserved members of a *Drosophila* gene family has also been reported in the case of the 70-kDa heat-shock genes (Ingolia and Craig, 1982).

Is there a sharp boundary to the *Jonah* gene family, and if so, how many members does it encompass? Would the gene at 74E, in turn, show homology to a *Jonah* gene at another novel site? One wonders to what lengths such genomic leap-frogging might be extended. Such a process need not be unidimensional: other isolated *Jonah* genes might be points of departure for other such paths. Expanding gene families might in some cases arise as a consequence of exon shuffling (Gilbert, 1978). A gene consisting of two domains, *a* and *b*, which were, at least originally, exons, would yield a reference cDNA which would define a family including all genes with exons *a* or *b*. One of these genes might consist of *a* and *c*, and the family would thereby expand to include all genes containing *c*, etc. In any case, the scenario observed for the *Jonah* genes invites questions as to how far the structural relationships among genes are followed by functional relationships, and in a certain sense it is this question which underlies the experiments described in the accompanying paper (Carlson and Hogness, 1985).

We thank Stewart Scherer for helpful discussions throughout the course of this work. We also thank Mariana Wolfner for performing and interpreting several of the *in situ* hybridizations and also for her valuable contributions to the work on adm135A8. This work was supported by grants from the National Institutes of Health and the National Science Foundation. J.C. was supported by a National Science Foundation Predoctoral Fellowship and an N.I.H. Traineeship.

REFERENCES

BENTON, W. D., and DAVIS, R. W. (1977). Screening λgt recombinant clones by hybridization to single plaques *in situ*. *Science (Washington, D. C.)* **196,** 180–182.

BOLIVAR, F., RODRIGUEZ, R. L., GREENE, P. J., BETLACH, M. C., HEYNEKER, H. L., and BOYER, H. W. (1977). Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* **2,** 95–113.

CARLSON, J. R., and HOGNESS, D. S. (1985). Developmental and functional analysis of *Jonah* gene expression. *Dev. Biol.* **108,** 355–368.

COHEN, S. N., CHANG, A. C. Y., BOYER, H. W., and HELLING, R. B. (1973). Construction of biologically active bacterial plasmids *in vitro*. *Proc. Natl. Acad. Sci. USA* **70,** 3240–3244.

DAVIS, R. W., BOTSTEIN, D., and ROTH, J. R. (1980). "Advanced Bacterial Genetics." Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

DAVIS, R. W., and HYMAN, W. (1971). A study in evolution: The DNA base sequence homology between coliphages T7 and T3. *J. Mol. Biol.* **62,** 287–301.

DENHARDT, D. T. (1966). A membrane-filter technique for the detection of complementary DNA. *Biochem. Biophys. Res. Commun.* **23,** 641–646.

ELGIN, S., and MILLER, D. (1978). Mass rearing of flies and mass production and harvesting of embryos. *In* "The Genetics and Biology of *Drosophila*" (M. Ashburner and T. R. F. Wright, eds.), Vol. 2A, pp. 112–121. Academic Press, New York.

FYRBERG, E. A., KINDLE, K. L., and DAVIDSON, N. (1980). The Actin genes of *Drosophila*: A dispersed multigene family. *Cell* **19,** 365–378.

FYRBERG, E. A., MAHAFFEY, J. W., BOND, B. J., and DAVIDSON, N. (1983). Transcripts of the six *Drosophila* actin genes accumulate in a stage- and tissue-specific manner. *Cell* **33,** 115–123.

GILBERT, W. (1978). Why genes in pieces? *Nature (London)* **271,** 501.

GUILD, G. M. (1984). Molecular analysis of a developmentally regulated gene which is expressed in the larval salivary gland of *Drosophila*. *Dev. Biol.* **102,** 462–470.

INGOLIA, T. D., and CRAIG, E. A. (1982). *Drosophila* gene related to the major heat-shock-induced gene is transcribed at normal temperatures and not induced by heat shock. *Proc. Natl. Acad. Sci. USA* **79,** 525–529.

LIFTON, R. P., GOLDBERG, M. L., KARP, R. W., and HOGNESS, D. S. (1978). The organization of the histone genes in *Drosophila melanogaster:* Functional and evolutionary implications. *Cold Spring Harbor Symp. Quant. Biol.* **42,** 1047–1051.

LIS, J. T., PRESTIDGE, L., and HOGNESS, D. S. (1978). A novel arrangement of tandemly repeated genes at a major heat shock site in *D. melanogaster*. *Cell* **14,** 901–919.

MCDONELL, M. W., SIMON, M. N., and STUDIER, F. W. (1977). Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. *J. Mol. Biol.* **110,** 119–146.

MEYEROWITZ, E. M., and HOGNESS, D. S. (1982). Molecular organization of a *Drosophila* puff site that responds to ecdysone. *Cell* **28,** 165–176.

MODRICH, P., and ZABEL, D. (1976). *Eco*RI endonuclease physical and catalytic properties of the homogeneous enzyme. *J. Biol. Chem.* **251,** 5866–5874.

MURRAY, N. E., BRAMMAR, W. J., and MURRAY, K. (1977). Lambdoid phages that simplify the recovery of *in vitro* recombinants. *Mol. Gen. Genet.* **150,** 53–61.

MUSKAVITCH, M. A. T. (1980). "Molecular Analysis of a Gene in a Developmentally Regulated Puff of *Drosophila melanogaster*." Ph.D. thesis, Stanford University, Stanford, California.

MUSKAVITCH, M. A. T., and HOGNESS, D. S. (1980). Molecular analysis of a gene in a developmentally regulated puff of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **77,** 7362–7366.

MUSKAVITCH, M. A. T., and HOGNESS, D. S. (1982). An expandable gene that encodes a *Drosophila* glue protein is not expressed in variants lacking remote upstream sequences. *Cell* **29,** 1041–1051.

MYERS, J. C., SPIEGELMAN, S., and KACIAN, D. L. (1977). Synthesis of

full-length DNA copies of avian myeloblastosis virus RNA in high yields. *Proc. Natl. Acad. Sci. USA* **74**, 2840–2843.

OHTSUBO, H., and OHTSUBO, E. (1976). Isolation of inverted repeat sequences, including IS1, IS2 and IS3, in *Escherichia coli* plasmids. *Proc. Natl. Acad. Sci. USA* **73**, 2316–2320.

PEACOCK, A. C., and DINGMAN, C. W. (1968). Molecular weight estimation and separation of ribonucleic acid by electrophoresis in agarose-acrylamide composite gels. *Biochemistry* **7**, 668–674.

RIGBY, P. W. J., DIECKMANN, M., RHODES, C., and BERG, P. (1977). Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J. Mol. Biol.* **113**, 237–251.

SANGER, F., AIR, G. M., BARRELL, B. G., BROWN, N. L., CORELSON, A. R., FIDDES, J. C., HUTCHISON, C. A., III, SLOCOMBE, P. M., and SMITH, M. (1977). Nucleotide sequence of bacteriophage ϕX174DNA. *Nature (London)* **265**, 687–695.

SCHACHAT, F., and HOGNESS, D. S. (1973). Repetitive sequences in isolated Thomas circles from *Drosophila melanogaster*. *Cold Spring Harbor Symp. Quant. Biol.* **38**, 371–381.

SOUTHERN, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503.

SPIERER, P., SPIERER, A., BENDER, W., and HOGNESS, D. S. (1983). Molecular mapping of genetic and chromomeric units in *Drosophila melanogaster*. *J. Mol. Biol.* **168**, 35–50.

SPRADLING, A. C., and RUBIN, G. M. (1981). *Drosophila* genome organization: Conserved and dynamic aspects. *Annu. Rev. Genet.* **15**, 219–264.

TOBIN, S. L., ZULAUF, E., SANCHEZ, F., CRAIG, E. A., and MCCARTHY, B. J. (1980). Multiple actin-related sequences in the *Drosophila melanogaster* genome. *Cell* **19**, 121–131.

WOLFNER, M. F. (1980). "Ecdysone-Responsive Genes of the Salivary Gland of *Drosophila melanogaster*." Ph.D. thesis, Stanford University, Stanford, California.

ZULAUF, E., SANCHEZ, F., TOBIN, S. L., RODEST, U., and MCCARTHY, B. J. (1981). Developmental expression of a *Drosophila* actin gene encoding actin I. *Nature (London)* **292**, 556–558.