## Genes in action: Polytene chromosomes from Drosophila larval salivary gland

**(http://people.hsc.edu/faculty-staff/edevlin/edsweb01/courses/Development/labmanual/new_page_7.htm http://web.as.uky.edu/biology/faculty/kellum/bio315/Lab%204A,%204B%20Exercise-Spring%202015.pdf)**

The regular chromosome complement of *Drosophila melanogaster* consists of three pairs of autosomes and a pair of sex chromosomes. The stained appearance of mitotic chromosomes in most cells is as follows: Two of the autosome pairs are large and V-shaped (chromosome 2 and 3) whereas the third pair are small dots (chromosome 4). The X chromosome is long and rod-shaped, and a male Y chromosome is J-shaped.

In the salivary glands, however, the chromosomes do not appear this way at all. The glands have evolved cells of tremendous size that are controlled by proportionately enlarged nuclei. In the process of enlargement, the chromosomes in each nucleus duplicate themselves many times over (endoduplication/endoreduplication- without the cell dividing), but in such a way that each duplicate is lined up parallel to all other duplicates of the same chromosome, (polyteny). A single salivary chromosome may thus consist of about a thousand copies (1024) lined up together. The "banding" of the salivary chromosome actually represents small density differences along the chromosome that are amplified by the fact that the duplicated chromosomes lie in register.

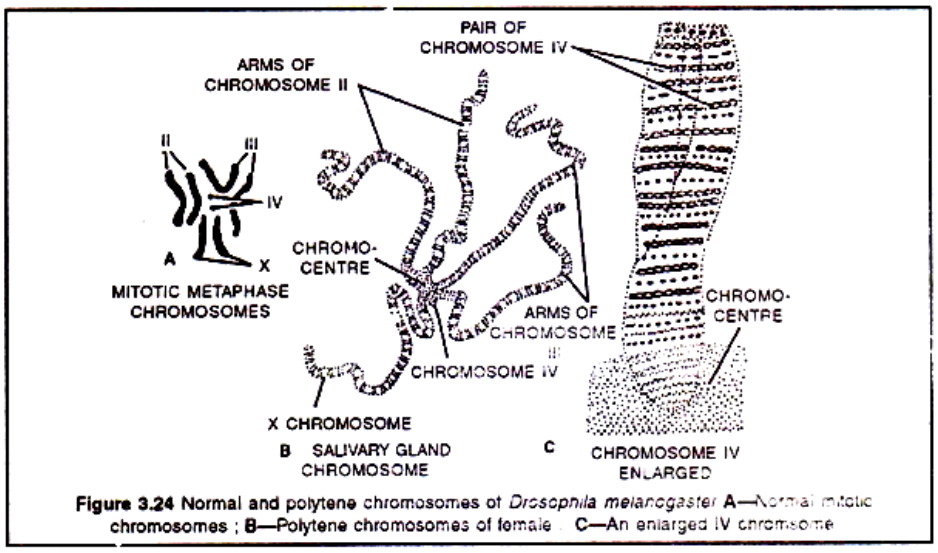
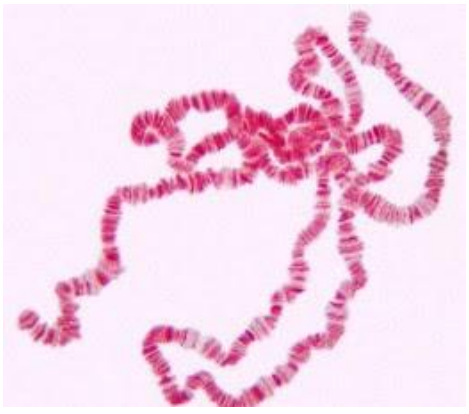



**Figure 3.24** Normal and polytene chromosomes of *Drosophila melanogaster* **A**—Normal mitotic chromosomes ; **B**—Polytene chromosomes of female . **C**—An enlarged IV chromsome

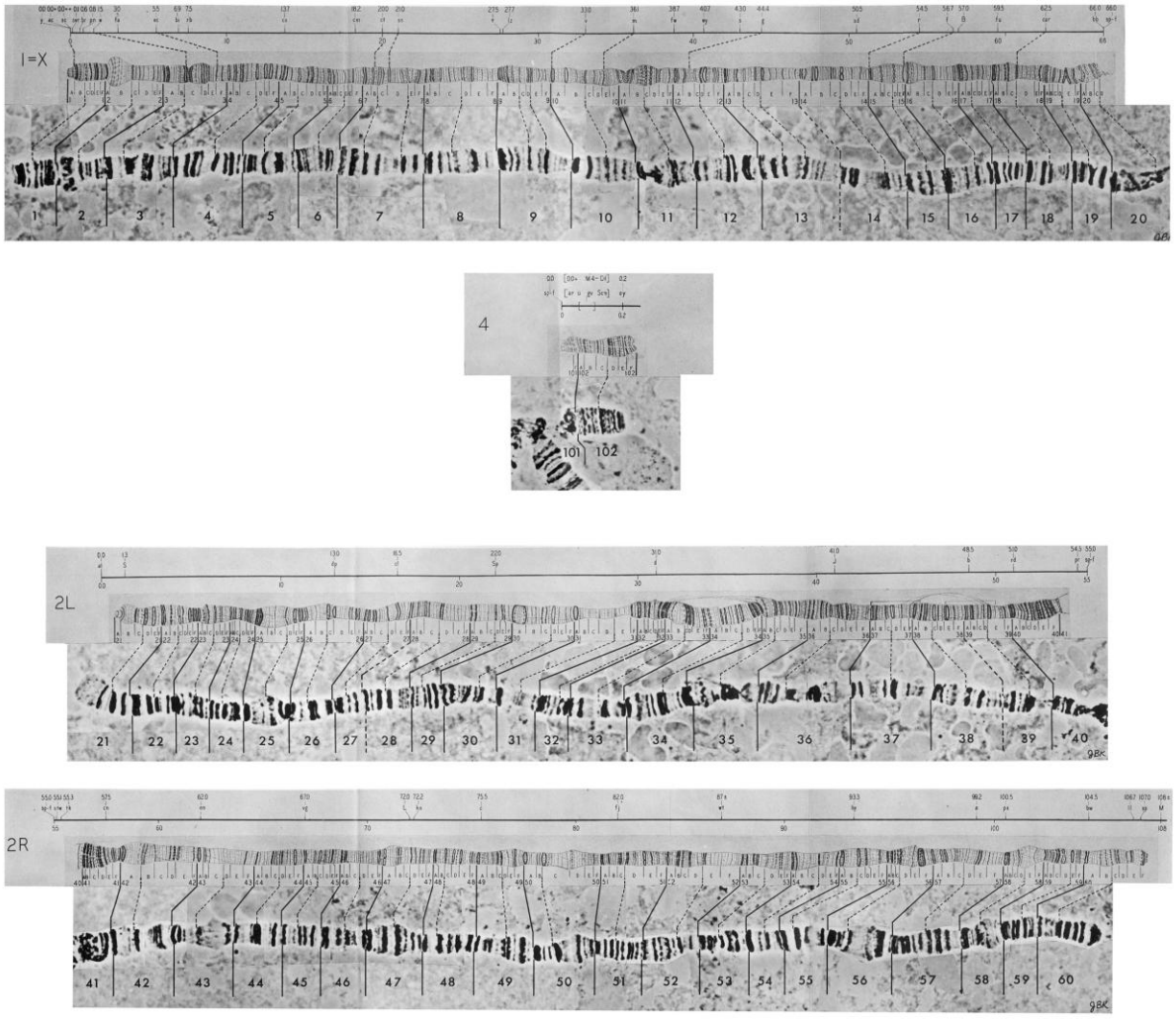


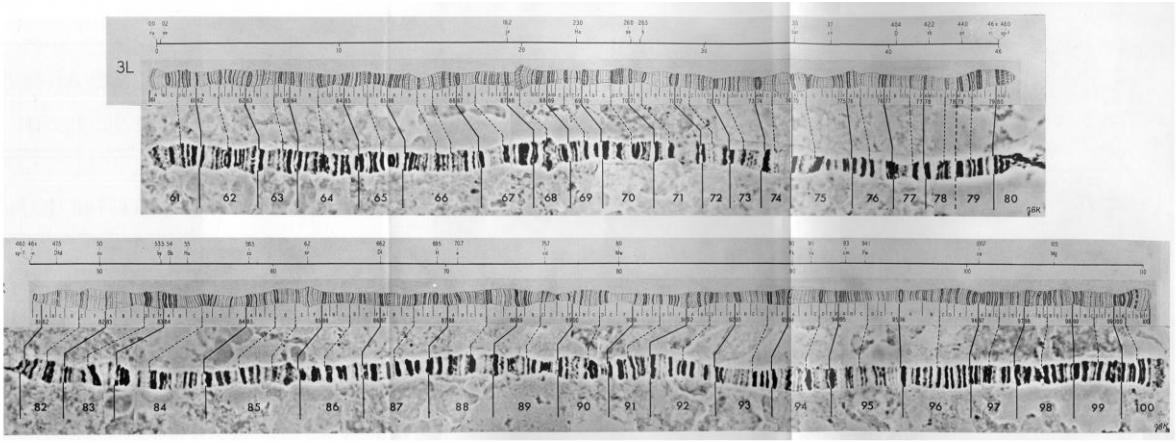
**Labelling the Polytene Chromosomes:**

The reference system proposed by Bridges divides the limbs of salivary gland chromosomes into 102 sections called "divisions" designated by number from 1 to 102. Each of the five main limbs (X, 2L, 2R, 3L, and 3R) contains 20 divisions; the short chromosome 4 contains only two divisions. The divisions are started with a prominent band and divided further into 6 subdivisions, each designated with capital letters from A to F. Each subdivision starts with a sharp band. Thus each individual band of salivary gland chromosomes can be identified by giving the division number, subdivision, and the number of the band starting from the beginning of the subdivision. Bridges presents the following minimum numbers of bands for the salivary gland chromosomes of *Drosophila melanogaster*: 537 bands for the X chromosome, 1032 bands for the second chromosome, 1047 bands for the third chromosome, and 34 bands for the fourth chromosome, totalling a minimum of 2650 bands for the whole genome. In this initial count doublets were listed as single bands; more recent interpretations give the total number of bands as 3286 (Sorsa, 1988).

*Drosophila melanogaster* salivary gland chromosomes as drawn by C B Bridges:

(http://www.sdbonline.org/sites/fly/aimorph/puffing.htm)



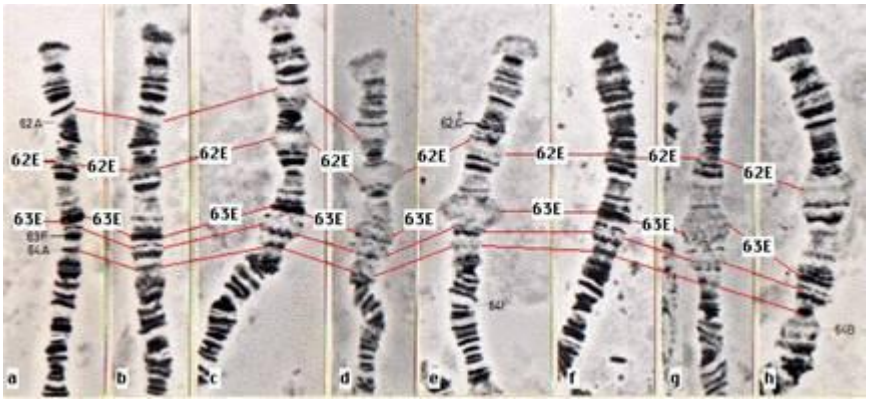



**Studying gene activity using polytene chromosomes:**

When a particular gene is being transcribed, its corresponding band assumes a puffed appearance in which diffuse material extends away from the axis of the chromosome. It is of interest that the same bands are always puffed at a given developmental stage of the larvae and that different bands are puffed at different developmental stages. Their chromosomes can be viewed at the link below:

http://www.ncbi.nlm.nih.gov/projects/mapview/map_search.cgi?taxid=7227&query=

These eight photomicrographs (courtesy of Dr. Michael Ashburner, University of Cambridge) show the changes in the puffing pattern of equivalent segments of chromosome 3 in **Drosophila melanogaster** over the course of some 20 hours of normal development. Note that during this period, when the larvae were preparing to pupate, certain puffs formed, regressed, and formed again. However, the order in which they did often differed. For example, in the larva, band 62E becomes active before 63E (c, d, and e), but when pupation begins, the reverse is true (g, h).




During the normal development of the larval salivary gland of *Drosophila*, considerable changes occur in the patterns of puffing activity. These can be seen as changes in the puffs of the gland's polytene chromosomes, and occur as a consequence of changes in the concentration of the insect's growth and molting hormone, ecdysone. In addition to the changes in gene activity in normal development, there are changes in the activities of a set of

genes that occur as a direct consequence of subjecting animals to a wide variety of experimental insults, for example, a brief heat shock. The discovery of the induction of a unique set of puffs by heat shock has led the way to an analysis of gene function and structure in *Drosophila* that is, so far, unique.

If *Drosophila* are subjected to a brief heat shock (40 minutes at 37°C, the normal culture temperature being 25°C) puffs are induced at specific sites. There are nine inducible puffs in *Drosophila melanogaster* at bands 33B, 64C, 64F, 67B, 70A, 87A, 87C, 93D, and 95D. Additionally, all other puffs active at the time the temperature shock begins, regress. A second dramatic heat shock response is evident in the pattern of protein synthesis. The synthesis of a small number of heat shock polypeptides (HSPs) is induced and the synthesis of most other proteins ceases; the protein synthesizing machinery is therefore somehow modified so that only HSPs are translated.

The cells must have a coordinated response to protect themselves in the event of excess heat; the same responses are also seen as the result of a number of other unrelated environmental insults such as Azide treatment, hydrogen peroxide, valinomycin treatment, or oxygen deprivation. This coordinated response, in which whole batteries of genes are shut down and others are turned on, is thought to be analogous to developmental switching mechanisms such as going from blastula to gastrula or determining dorsal/ventral axes.

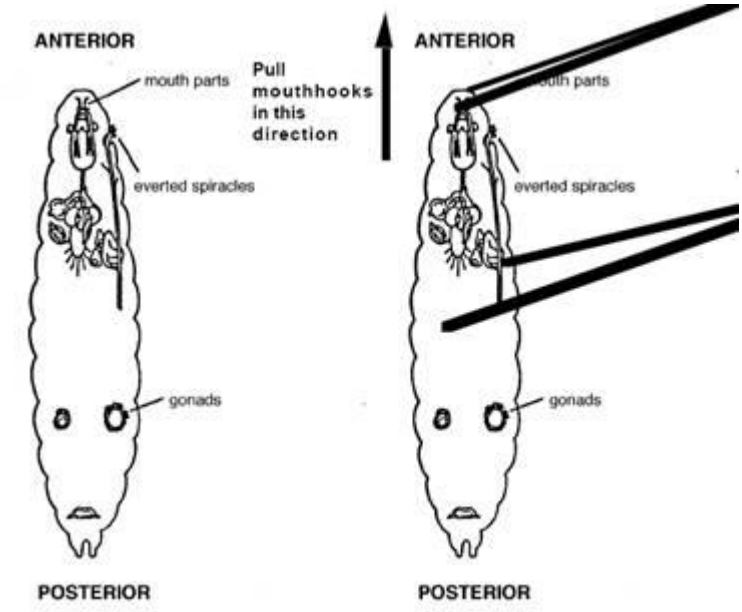
This coordinated response, exemplified by the heat shock phenomenon, can be thought of as a general mechanism used by developing organisms. In the heat shock response, a whole battery of genes is turned off and another battery of genes turned on, in a matter of minutes, in response to some signal or stimulus. Embryonic development probably proceeds in such a fashion. A tissue might, for instance, make a decision to be anterior or posterior, dorsal or ventral, thoracic segment or abdominal segment, by switching on a series of related genes. Thus, the heat shock response is useful as a model system to study the mechanisms of coordinate gene regulation.

**Procedure:**

**1)** Dissecting scope (one per Lab Group)
**2)** Drosphila larvae (third instar)
**3)** 10% Acetic acid
**4)** 45% Acetic Acid
**5)** Aceto Orcein stain
**6)** Tissue paper
**7)** Microscope slides and Coverslips
**8)** A pair of dissection needles
**9)** Nail polish

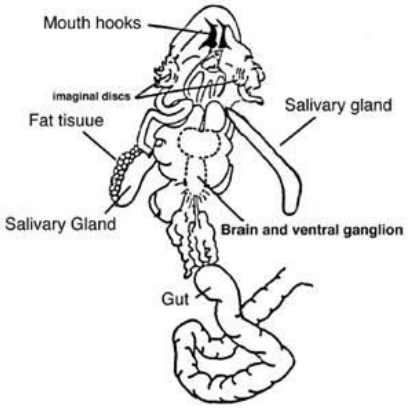
**<u>Part 1: Preparing Salivary Gland Chromosomes from Un-treated larvae.</u>**

**1.** Place a drop of 10% Acetic acid in the middle of a microscope slide. Use moist brush to transfer one third instar larva to the liquid drop.
**2.** Use one dissection needle to hold the anterior end of the larva in place at the spot just above the mouth hooks in, while using another pin to hold the larva at about one third body length from the anterior end.




**3.** Pull the pins apart gently in one smooth motion. The larva should tear just below the mouth hooks.

4. The salivary glands and ventral ganglion (brain) will usually remain attached to the head region, separate from the rest of the body, after a clean dissection. You may have to use your pins to search for the salivary glands among the dissected material. If this fails, simply start over with a new larva. These dissections are difficult to master, but reasonable results can be obtained with a little practice.




**5.** Clear the debris and retain only the salivary gland on the slide. Ensure that the gland does not dry out at any point. Remove excess acetic acid by using a tissue paper. Tilt the slide and gently touch the acetic acid with the tip of tissue paper. Do not let the tissue paper touch the gland else they will adhere to the tissue paper.

**6.** Then place a drop of **Aceto Orcein** stain on the glands. Let the glands stain for about 15-20 minutes. Make sure that the stain does not dry out. If necessary, place the slides in a moist chamber (a Perti plate with wet tissue paper).

7. Remove excess stain using tissue paper.

8. Place a drop of 45% Acetic acid on the glands.

**Watch the demonstration before attempting steps 9 to 11.**

**9.** Place a clean coverslip (dust blown from its surface) onto the surface of the glands. Wrap the coverslip and slide with a tissue paper and firmly press down with your thumb. Make sure that the coverslip does not move.

10. Using the back of a pencil, gently tap the coverslip (over the tissue paper) in a circular motion, starting at the centre and moving outwards.
11. Seal the edges of the coverslip with clear nail polish.
**12.** Place the slide on your microscope. Using the lowest power objective and standard light source, scan around the slide until you find the squashed material. It is often useful to first locate the squashed material on the slide by eye. Then place the slide on the microscope stage with this material centered in the light beam. This will make it easier to find your chromosomes under the microscope. First locate the focal plane of your chromosomes by adjusting the focus knob until the material becomes visible. You can then begin to scan around the slide to find the chromosomes.
**13.** If the chromosomes look sufficiently well spread, you can now continue with your observations.
**14.** If the nuclei were found to be still intact, you can attempt to rupture them if there is evidence of liquid between the slide and coverslip as evidenced by movement of background material. To rupture the nuclei, use the back of a pencil as in step 10. It may be necessary to start over with some newly dissected salivary glands. Preparing polytene chromosome squashes is part science/part art. It takes a lot of practice and a little luck to get the quality of squash.

**Part 2: Preparing Salivary Gland Chromosomes from heat shocked larvae.**

Repeat steps 1-14 using larvae that have been subjected to heat shock.
To heat shock, larvae have been placed in an incubator at 37C temperature for 30 minutes.
Compare the chromosomes from the two types of larvae.

**Questions:**

1. Draw a sketch of your best polytene chromosome spread. Are you able to see all five major chromosome arms (X, 2L, 2R, 3L, 3R)? Are you able to see the small 4th chromosome? Can you locate the chromocenter in your spread? Are you able track along a single chromosome arm from its centromere to telomere?
2. Compare the polytene chromosomes from heat shocked and un-treated larvae. Can you make out the puffs? Can you locate all puffs?
3. Comment on the differences between the chromatin structure at the puffed and non-puffed regions.
4. Design an experiment to check if there is a temporal order in the activation of Heat Shock Proteins.

# SALIVARY CHROMOSOME MAPS

## With a Key to the Banding of the Chromosomes of *Drosophila Melanogaster*

CALVIN B. BRIDGES

*Carnegie Institution of Washington*
*Resident at California Institute of Technology, Pasadena, California*

SINCE Heitz and Bauer have shown that the enormously enlarged chromosomes present in the nuclei of dipteran salivary gland cells are to be regarded as normal chromosomes rich in constant detail, and since Painter and his colleagues have demonstrated that the series of structures observable along the length of such chromosomes can be correlated to the series of genic loci on the linkage maps, it has become imperative for every Drosophila worker to make use of this new method of analysis. While the field of employment of salivary analysis is very wide, it is especially in those cases in which aberrations of the chromosomes may be involved, viz., deficiencies, duplications, translocations, inversions, etc., that the greatest saving in time and clarity of result may be expected. For such analysis two types of chromosome maps are prerequisites: first, linkage maps which give the sequence and locations of the genes for all mutant characters which may be involved, and, second, accurate detailed charts of all the normal salivary chromosomes against which to check the precise points of breakage or area of disturbance. A necessary adjunct to the detailed maps of the chromosome banding is an objective system of referring to a particular band or section of a chromosome. The general adoption of a system of "Salivary chromosome coordinates" by workers in this field is thus a matter of prime importance. The plan outlined below has proved in practice to have great advantages in being both accurate and elastic.

For many years I have maintained current linkage maps summarizing the genetic evidence on the locations of genes. These have been published from time to time, and revised copies have also been sent to many individuals upon request. The latest revision will appear in a forthcoming number of this JOURNAL and will be distributed widely through the "Drosophila Information Service."

### The Wealth of Detail Observable in Salivary Chromosomes

Painter has just published in this JOURNAL (December, 1934, pp. 465-476) his survey of the normal salivary chromosomes, with maps which show the salient features. But, as he states, the amount of detail to be seen in the chromosomes is far in excess of that shown. During the past year I have attempted to make, for the structures of the normal salivary chromosomes, maps which would show this finer detail. Such detailed maps were found indispensable in precise study of chromosome aberration. They are herewith presented at a magnification sufficient to carry most of the detail, though to insure that the faintest lines observable be not lost entirely in reproduction, it has been necessary to sacrifice much of the range of intensities in the original drawings which were shown at the annual exhibit of the Carnegie Institution of Washington, December 14-17, 1934. Hence in using these maps it must be remembered that the fainter lines of the chromosomes and the fainter areas appear disproportionately too conspicuous and dark upon the maps reproduced herewith.

### Refinements Which Aid Observation

For observing the finest detail several refinements of technique are required.

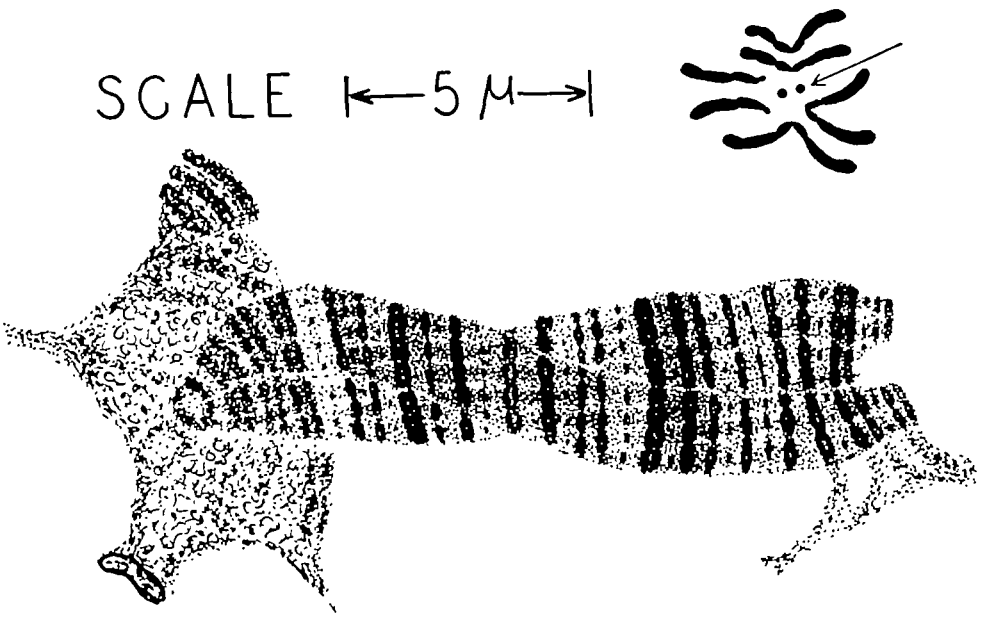
One is relatively light transparent staining of the chromosomes, with avoidance of heavy "contrasty" staining, which may give the heavy lines very dark but the lighter lines not at all. Much iron and heating tend to spoil the finer details. The crispness of detail seen in larvae fully grown in pair cultures at a low temperature is lost in larvae from old cultures, from mass cultures and in larvae which have begun pupation. Especially favorable material has been attached-X (*XXY*) females of the race giant bobbed-11 picking giant larvae and examining the double-thickness chromosomes of certain cells found there.

Another requisite is selection of chromosomes or portions which are straight (i. e., not kinked or coiled) and are stretched somewhat. The lax chromosomes are from 70 to 110 times as long as normal gonial chromosomes, but the somewhat stretched chromosomes which are most favorable for observation are 150-160 times normal length. The maps presented herewith are drawn only from such partially stretched chromosomes, averaging 150 times normal. (For comparison a gonial group at the same magnification is included.) The gross structure of salivary chromosomes is somewhat like that of an accordian, and unless these chromosomes are stretched the doubleness of most bands is not visible and many fine or dotted lines are obscured by their appressed neighbors.

Too much attention cannot be paid to the illumination of the chromosomes. I use a 6-volt ribbon-filament lamp, and control intensity by a 175-ohm 2 amp. variable resistance in the 110-volt current to the transformer. The light from the ribbon is brought to a sharp focus in a small-area image about 30 cm from the filament and 25 cm from the mirror. This image is diaphragmed at that point to a circle the width of the ribbon image (3.5 mm). For turning the red color of the stain black, I set up Wratten filter 58A (deep yellowish green) behind one or more cobalt- blue glasses just before the mirror. A front-silvered or aluminized mirror is preferable. An achromatic 1.3 or 1.4 condenser, oil-immersed, should be focused so that the edge of the filament image and/or the edge of the diaphram are sharp in the field observed. The substage diaphragm should then be gradually closed down, with compensating adjustment of light intensity by the rheostat, until the area outside the image of the filament loses its light-haze and becomes dark, while the details of the chromosome banding increase in sharpness and contrast. Final checking should be made by oblique light (decentering the substage diaphragm) cast along the axis of the chromosome. I prefer 120× apochromatic objective with 10× compensating ocular for observations, but a 90× with 10× or 12.5× eyepieces does nearly as well and does not require such critical attention to illumination, cover glass thickness and other details.

### A System of Cataloguing Salivary Bands

The system of chromosome map nomenclature proposed in this paper divides the five main chromosome limbs (1=X, 2L, 2R, 3L and 3R) each into twenty sections, 100 in all. Sections are numbered 1 to 20 for X, 21 to 40 for 2L, 41 to 60 for 2R, 61 to 80 for 3L and 81 to 100 for 3R. Chromosome 4 has sections 101 and 102. Hence the number of a section is itself a key to the chromosome limb and to the relative position along that limb. Since sharpness and definiteness are essentials, each section begins with a conspicuous and easily recognized band. The division point is always made just to the left of the chosen main band, leaving all minor bands inside the division to the left of it. But since the 102 divisions average over 25 bands each, six subdivisions have been established for each division. Each subdivision also begins with a sharp band. They are designated by the capital letters A to F. A particular band would then be referred to as 17B3; a break would be referred to as just to the left




**SALIVARY CHROMOSOMES AND GONIAL CHROMOSOMES COMPARED**

**Figure 4**

Drawings of chromosome 4 of *D. melanogaster* and, on the same scale, of the entire group of gonial chromosomes at metaphase. In this gonial group the paired fourth chromosomes are represented by the small black dots, in which no structural details can be seen under the highest magnification, in striking contrast to the wealth of detail visible in the salivary chromosomes.

of 42C2 and a section as extending from 36A1 (included) to 38B1 (not included). The bands are not given definitive numbers on the maps because those numbers would change from year to year as our knowledge of the banding becomes more detailed and small bands are seen which were previously missed. No such change should be necessary for the divisions and subdivisions, since each begins with a sharp band already well established.

### Landmarks of the Chromosomes

The "segmentations" observable in the chromosomes do not offer sharp enough boundaries for a serviceable reference system, though certain ones are very convenient landmarks for recognizing chromosomes in a tangle. Among the natural landmarks which should be learned first are the "puff" in 2B, the "four brothers" in 9A, the "weak spot" in 11A, the two "chains" in 15, the "turnip" in 16 and the "offset" in 19E. The huge lightly staining nucleolus of the salivary gland cell nucleus is attached to the base of the X at the bands in 20C and D. Diagnostic of 2L are the "dog-collar" in 21CD, the "shoe-buckle" of 25A, the "shield" in 30A, the "goose-neck" in 31BF, the "spiral loop" of 32-35, a "turn-back" in 36, and the "basal loop" in 37-39. One recognizes 2R by its thick "onion" base and "huckleberry" tip. Three L has a "barrel" at 61CF, a "ballet skirt" in 68BC, "chinese lanterns" in 74-75 and "graded capsules" in 79CDE. Finally, 3R, the longest limb, has a large clear "cucumber" base 81-83D, a "duck's head" at 89E to 91A (frequently breaks at junction of 89D and E) and a "goblet" tip. Most of the apparent segmentation is inconstant and is due to the fact that the maternal and paternal part-

ners are sometimes seen side by side, as in section 25, and sometimes superimposed, as in section 24.

The free end of every chromosome limb presents a characteristically narrowed terminal region, seen especially clearly in the maps at 1A, 21A, 60F, but present also in part in 61A, 100F and 102F. It is suggested that this narrowed region represents a lag of one division in the gene-strings, somehow due to the terminal position occupied but not due to special properties of those particular genes.

### Size and Structure of the Chromosomes

The average lengths of the moderately stretched salivary chromosomes as drawn are: 1 = X = 220 μ; 2 = 215 + 245 = 460 μ; 3 = 210 + 275 = 485 μ; 4 = 15 μ; total = 1,180 μ. Thus the total length of the moderately stretched salivary chromosomes is approximately 150 times that of the total length of the gonial chromosomes, which is 7.5 μ. Individual salivary chromosomes with lengths exceeding 180 times normal have been measured. These excessively stretched chromosomes or portions show most of the stretching in the hyaline zones between the dark "bands." The heavier cross-bands or capsules retain their shape as broad firm discs while the material between may stretch into a narrow cord ten times its normal length but still showing its compound nature.

As deduced independently by Dr. Koltzoff (*Science*, Oct. 5, 1934) and myself (in press in U. S. S. R. since June, 1934) the large size of the salivary chromosomes is partly due to their being compound structures. Each of the fused maternal and paternal homologues consists of eight chromonemata or gene-strings derived from the corresponding chromosome of the gamete by successive divisions without complete separation of the division products. The cable of 8 + 8 strands shows its structure most clearly in the less heavy cross-bands in which 16 individual dots, vesicles or small capsular units may be seen.

### The Relation of the Genes to the Bands

My inference as to the relation of the genes to the structures seen in the salivary chromosomes is that each of the faint cross-bands made up of 8 + 8 dots (see last line of 102D), dashes (middle line of 102E) or vesicles (first line of 102E) corresponds to one locus with 8 maternal and 8 paternal sister genes. I suppose that each of the sister genes of a locus is enclosed within one of the vesicular units, the gene itself being small and unstained but the walls of the enclosing capsular or clam-shell structure being visible from a deposit of chromatin. If the bipolar deposits of chromatin are somewhat heavier they run together at their edges to form two thin discs, one on each side of the plane of the genes. This hypothesis is in line with the observation of large numbers of thin bands (like those in 102A) which in edge view are wavy close doublets like the two halves of a split pea-pod. With still heavier deposits of chromatin, in amount characteristic for each locus, heavy obscurely double and somewhat nodulated bands result (like the first in 102B).

But besides these "thin-walled" structures there are "heavy-walled" structures (such as the first in 4F) which seem better interpreted as compound bands made up of two bands more or less united at their edges to give a "heavy-walled capsule." In very good preparations some of these heavy-walled capsules can be seen as two parallel bands not united at their edges. Certain of the heavy-walled capsules would seem to correspond to three loci, since they enclose between them a line of dots or dashes (4A1; 6A1; 7C1, etc.).

A count of the distinctly seen "lines" in the salivary chromosomes gave approximately 725 for X, 1,320 for 2, 1,450 for 3 and 45 for 4, totaling 3,540 lines. But a considerable proportion of the above lines seem incipiently double

or with an indistinct split. A count of "bands" or "loci," on the basis that each line of dots, dashes or vesicles and also each pair of closely approximated thin or moderately heavy lines represents one locus, while each heavy-walled capsule represents two (or three) loci, yielded: for X, 537; for 2, 1,032; for 3, 1,047, and for 4, 34 "bands." The total of 2,650 bands is in good agreement with calculations of 1,500 to 3,000 genes for the animal.

### Direct and Reversed Repetitions of Series of Bands

A structural feature of the highest theoretical importance is shown clearly in the "loops" and the "turn-back" in the basal half of 2L. Great difficulty was encountered at first in studying the base of 2L because of the troublesome kinks and coils. But in some cells the spiral coils were found stretched nearly straight and then it was observed that the edges of certain bands had been fused together and had been stretched out into connecting threads. The connected bands were found to match morphologically, 32F to 33C with 34F to 35C, in a whole series of repeated bands. The spiral loop was due to synapsis between homologous series of bands in two different positions in the same chromosome. Similarly, the basal loop in 37 and 38-39, when stretched out, revealed connecting strands linking (for example) the bands in 37EF and 38A with bands of identical morphology in 39CDE. Further study showed that all the bands in 37 were at least roughly matched in 38E to 39E (bracketed).

The "turn-back" in 36 with side by side fusion and the "shield" in 30A show reversed series forming sections symmetrical about their center points. It is significant that the bands in 30A show fusion along the surface of the chromosome to give a giant capsule. This suggests that most of the larger capsules, such as the four-banded ones in 25A and in 56F, are symmetrical reversed repeats. The reversed symmetries around "weak spots" 3C and 11A may be further examples of duplications of sections of homologous bands.

How complicated a structure could be built by successive direct and reversed repeats is perhaps illustrated by the reversed symmetry with capsule tendency which centers around 33B and which is itself a part of the series repeated in direct sequence in 34-35.

The thick-walled capsules like that in 21E, some of which seem to enclose a line between them, as in 21D, may represent single-band or short repeats. There are also large numbers of pairs of lines of equal intensity, like that in 27E, the members of which are definitely separate. Perhaps some local process, such as unequal crossing-over, may have to be invoked to account for them.

### The Role of Duplications in the Evolution of Chromosomes and the Initiation of Species

For the long repeats, both direct and reversed, an origination as duplications through the process of translocation would seem an adequate explanation. In my first report on duplications at the 1918 meeting of the A. A. A. S., I emphasized the point that the main interest in duplications lay in their offering a method for evolutionary increase in lengths of chromosomes with identical genes which could subsequently mutate separately and diversify their effects. The present demonstration that certain sections of normal chromosomes have actually been built up in blocks through such "repeats" goes far toward explaining species initiation. For the duplication of sections of genes is known in Drosophila to cause many slight poorly-defined differences in all parts of the duplicant type.

OPEN ACCESS Freely available online

PLOS ONE

# Genetic Organization of Interphase Chromosome Bands and Interbands in *Drosophila melanogaster*

**Igor F. Zhimulev[1,2]*, Tatyana Yu. Zykova[1], Fyodor P. Goncharov[1], Varvara A. Khoroshko[1], Olga V. Demakova[1], Valeriy F. Semeshin[1], Galina V. Pokholkova[1], Lidiya V. Boldyreva[1], Darya S. Demidova[1,2], Vladimir N. Babenko[3], Sergey A. Demakov[1], Elena S. Belyaeva[1]**

**1** Institute of Molecular and Cellular Biology of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia, **2** Novosibirsk State University, Novosibirsk, Russia, **3** Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

## Abstract

*Drosophila melanogaster* polytene chromosomes display specific banding pattern; the underlying genetic organization of this pattern has remained elusive for many years. In the present paper, we analyze 32 cytology-mapped polytene chromosome interbands. We estimated molecular locations of these interbands, described their molecular and genetic organization and demonstrate that polytene chromosome interbands contain the 5′ ends of housekeeping genes. As a rule, interbands display preferential "head-to-head" orientation of genes. They are enriched for "broad" class promoters characteristic of housekeeping genes and associate with open chromatin proteins and Origin Recognition Complex (ORC) components. In two regions, 10A and 100B, coding sequences of genes whose 5′-ends reside in interbands map to constantly loosely compacted, early-replicating, so-called "grey" bands. Comparison of expression patterns of genes mapping to late-replicating dense bands vs genes whose promoter regions map to interbands shows that the former are generally tissue-specific, whereas the latter are represented by ubiquitously active genes. Analysis of RNA-seq data (modENCODE-FlyBase) indicates that transcripts from interband-mapping genes are present in most tissues and cell lines studied, across most developmental stages and upon various treatment conditions. We developed a special algorithm to computationally process protein localization data generated by the modENCODE project and show that *Drosophila* genome has about 5700 sites that demonstrate all the features shared by the interbands cytologically mapped to date.

**Citation:** Zhimulev IF, Zykova TY, Goncharov FP, Khoroshko VA, Demakova OV, et al. (2014) Genetic Organization of Interphase Chromosome Bands and Interbands in *Drosophila melanogaster*. PLoS ONE 9(7): e101631. doi:10.1371/journal.pone.0101631

**Editor:** Igor V. Sharakhov, Virginia Tech, United States of America

**Received** March 17, 2014; **Accepted** June 9, 2014; **Published** July 29, 2014



**Data Availability:** The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

**Funding:** This research was funded by grant of Russian Scientific Foundation #14-14- 00934(http://www.rscf.ru/sites/default/files/docfiles/Spisok_pobediteley.pdf). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

* Email: zhimulev@mcb.nsc.ru

## Introduction

*Drosophila* polytene chromosomes have served as the best available model of eukaryotic interphase chromosome. They are prominent for their banding pattern formed by dark transverse stripes (called bands), which encompass large chunks of chromatin material. These bands alternate with fine, lighter-colored stripes that have less material and are more loosely packed. Such light-transparent structures between bands are known as interbands.

Genetic organization of bands and interbands defined as the pattern that sets positioning of genes and genetic features relatively to the structural elements of a chromosome, is still largely elusive. This is due to the fact that despite the availability of the *Drosophila* genome, methods to even approximately map band/interband borders on a physical map are still lacking.

Yet, many interesting hypotheses regarding the genetic organization of bands and interbands in polytene chromosomes have been put forth. Some of the points of these hypotheses were experimentally validated, so we believe it is important to consider them below.

Genes were proposed to reside in interbands [1], or bands (1–2 genes/band) [2–5]. Also, bands were proposed to contain many structural genes transcribed coordinately and as polycistronic messages [6]. In some models, band and interband were considered to form a single genetic unit, where one part of the gene was embedded in a band, and the other part mapped to an interband [7–12]. The Paul model [8] is of special interest. The author considered interband regions as essentially polymerase-binding sites, so transcription would progress into band regions from initiation sites which were likely situated near the band-interband junctions.

Very interesting conclusions were made regarding the general meaning of banding pattern: bands were regarded as hosting inactivated genes, interbands were represented by the genes in a steady state of activity [13–15]; in other models, interbands were believed to contain constantly active housekeeping genes [16–17]. Further details on various banding pattern models are available in [18].

Several recent technological advances have dramatically moved forward our understanding of polytene chromosome organization. First, efforts of the modENCODE project have produced genome-

wide profiling data for many proteins that specifically localized to bands or interbands in interphase chromosomes ([19], for review).

Secondly, using genome-wide DamID mapping of 53 chromosomal proteins and histone modifications Filion et al. [20] have generated a map of *Drosophila* chromatin landscape and demonstrated that the genome can be segmented into five main chromatin types. Conditionally named "BLUE" (Pc-dependent repression) and "BLACK" (repression mechanism not defined) chromatin types associated with repressed chromatin, "YELLOW" chromatin contained ubiquitously expressed genes, whereas "RED" chromatin harbored active genes with more complex expression patterns. "GREEN" chromatin type was defined by enrichment of heterochromatin-specific proteins HP1 and Su(var)3–9 ([21], for discussion). More refined analysis of modENCODE data resulted in description of many more chromatin states [22]. Significant proportion of genome sequence is known to map to a special class of polytene chromosome bands, called intercalary heterochromatin (IH) [23]. These chromosome regions associate with BLACK chromatin proteins (H1, SUUR, LAM, D1) and range from 100 to 700 kb in length. Here, DNA replicates late and compared to the genome average, these regions have lower gene density [24,25]. Genomic localization of proteins that constitute repressed chromatin can thus be used as a marker to establish the molecular position of IH [26].

Third, we recently developed an approach to simultaneously map the interband material on polytene chromosomes and in the genome using transposon insertion tags. This allows exact localization of insertion sites both on cytological and physical maps as well as precise identification of sequences around the transposon integration sites.

Using this approach, we describe protein composition and other chromatin parameters in 12 DNA sequences corresponding to polytene chromosome interbands. They display general features of open chromatin: low nucleosome density, histone H1 dips, association with TSS-specific proteins such as RNA polymerase II, various transcription factors, nucleosome remodeling proteins - NURF, ISWI, WDS, interband-specific proteins (CHRIZ/CHROMATOR, CHRIZ hereafter), proteins of origin recognition complexes (ORC). Moreover, they show clustering of DNaseI hypersensitive sites (DHS) (see for discussion [27,28]).

Based on these data, we subdivided all polytene chromosome bands into two contrasting groups: loosely compacted early-replicating, so-called "grey" bands and dense late-replicating compact bands ("black" IH bands). They differ in many aspects of their protein and genetic make-up, as well as in DNA compactization [27].

Previously, we showed that polytene chromosomes and interphase chromosomes from dividing cells display identical organization. Namely, interbands from polytene chromosomes and the corresponding DNA sequences from cell line chromosomes share similar features in terms of localization of open chromatin-type proteins. Consequently, banding pattern appears as a fundamental organization principle of interphase chromosomes. In both types of chromosomes, homologous interbands and bands have identical physical borders and length; importantly, they also associate with identical sets of proteins [19,26,27]. Hence, the notion of an interband defined as a decondensed region in the context of polytene chromosomes is also applicable to other types of interphase chromosomes. In other words, the term "interband" should be viewed as an equivalent of a constantly decondensed region in the context of any interphase chromosome. Accordingly, hereafter we use this wider definition of an interband.

In the present work, using various cytological approaches, we first characterized a new set of precisely mapped interbands, and then processed the modENCODE data on localization of active chromatin proteins using a custom-designed computation model. This analysis suggests that interphase chromosome interbands contain constantly active promoter regions of ubiquitously active genes. Coding sequences of these genes, at least in two regions studied, map to adjacent loosely compacted early-replicating "grey" bands. In contrast, densely packed, late-replicating bands of polytene chromosomes appear to preferentially harbor tissue-specific genes.

## Results

### Mapping interbands in polytene chromosomes and on a physical map

To analyze the interbands' protein make-up and to explore their molecular organization, these structures must be accurately mapped on both cytological, electron microscopy (EM) and physical maps. In this study, we present the molecular-genetic analysis of a set of interbands (32 in total), which we believe were unambiguously identified at the cytology level; 21 of these interbands were mapped earlier [29–33].

A group of 11 interbands was characterized in detail here, by comparing Bridges' polytene maps, EM data, modENCODE protein localization data and mapping of IH regions [26]. This group comprises the interbands from regions 7F, 19E, 21D, 35D, 56A, 58A, 70A, and 100B (see below and Text S1 [35–38]).

For illustrative purposes, below we provide detailed description of mapping data for interband regions found in 7F1-2 and 100B.

In the region 7F, two condensed and late-replicating bands 7F1-2 and 7F3-4 [34] flank a thin and well-defined interband, which is clearly observable both on light microscopy [35] and EM maps. When performing high-resolution analysis of chromosome banding pattern, C. Bridges never reported any additional minibands between these bands (Fig. 1A). Likewise, upon EM analysis of this region, we also observed no additional minibands (the interband of interest is marked by an arrow in Fig. 1 B). This interband is clearly decorated by an interband-specific protein CHRIZ (arrows in Fig. 1C-E), and FISH analysis indicates it harbors the 5′-end of the *Nrg* gene (arrows in Fig. 1F, G).

In the region 100B, polytene maps [39] show two doublets, 100B1-2 and 100B4-5, as well as a very faint band 100B3 in between (Fig. S1A, D). Thus, this region encompasses two interbands, one proximal and one distal to 100B3. DNA material between the bands 100B3 and 100B4-5 hosts the 5′-end of *dco* gene (Fig. S1E, F), whereas the region between 100B1-2 and 100B3 harbors the 5′-end of the gene *l(3)03670* (See Fig. S3, below). Both interbands appear CHRIZ-positive: the region demonstrates two stripes, one located in the interband 100B1-2/100B3 and the other in the 100B3/100B4-5 interband (Fig. S1B, C).

### Molecular and genetic organization of interbands and bands in *Drosophila* chromosomes

**Algorithm to identify genomic regions enriched in interband chromatin proteins.** Earlier we reported high similarity of banding patterns in both polytene and non-polytene diploid cells [19,27]. Out of the proteins mapped by modENCODE (modENCODE Consortium, 2010) in Kc, S2 and BG3 cell lines, we identified a set of proteins enriched in the regions of 12 reference P-element insertions in the interband regions of polytene chromosomes [27,28].

In this study, we aimed to partition the whole genome into discrete chromatin states defined by the local enrichment of "open chromatin" proteins that are found predominantly in interbands.

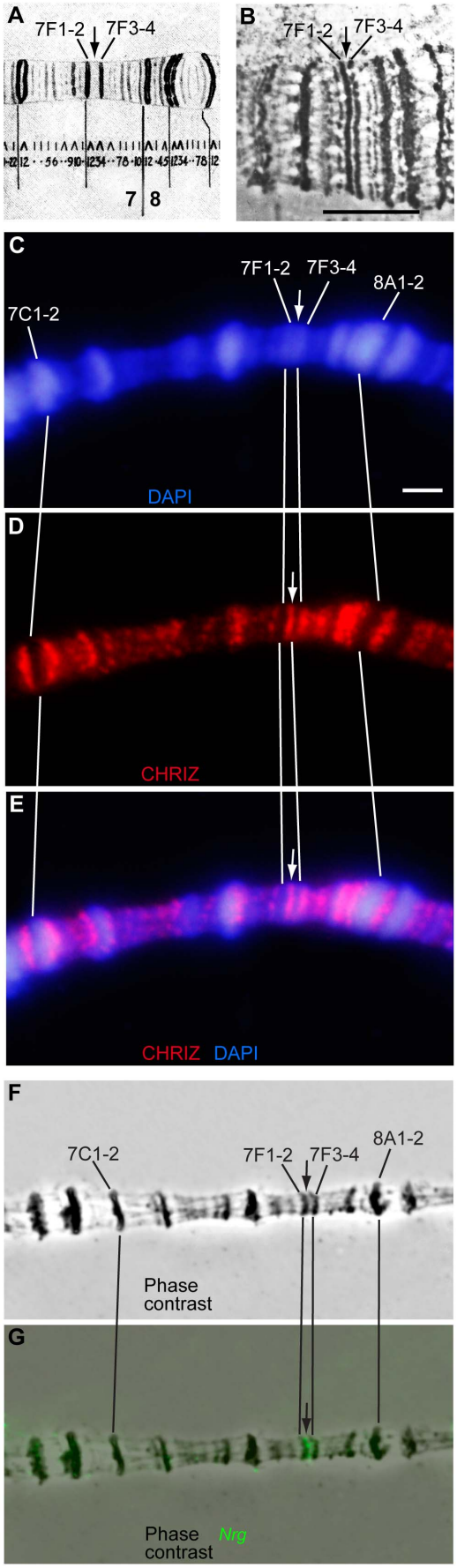



**Figure 1. Cytological mapping of an interband (arrow) in the 7F1-4 region of the X chromosome. A, B –** Fragment of the revised Bridges' map of *D. melanogaster* X chromosome [35] (**A**) and EM image of the region (**B**). Bar corresponds to 2 mkm. **C-E –** immunofluorescence localization of CHRIZ (red) in the interband 7F1-2/7F3-4. DNA is counterstained with DAPI (blue). **F-G –** FISH localization of the DNA probe containing the 5′-fragment of *Nrg* (green in **G**).
doi:10.1371/journal.pone.0101631.g001

To do so, we developed an algorithm that would allow us to assign molecular coordinates to the regions corresponding to interbands of polytene chromosomes, based on the localization of the above-mentioned "interband"-specific protein markers. Additionally, in contrast to the whole-genome approaches [20,22], our chromatin clustering analysis is based on the combination of protein profiling data across four cell lines: S2, Kc, BG3, Clone 8 using fly modENCODE non-histone protein dataset as a source (modENCODE Consortium, 2010). This was done in order to identify the genomic regions that are generally co-occupied by most of the proteins analyzed.

We first performed correlation analysis of protein localization data generated by the fly modENCODE project, and then proceeded to hierarchical clustering of proteins using the X chromosome as the best characterized chromosome in *Drosophila* genome (dendrogram in Fig. S2). Stopping rule $a_{i+1} \geq a+2s_a$ was used to define the appropriate sensitivity threshold [40], which produced four major classes of proteins. We chose to focus on just one class comprising 12 chromatin proteins, many of which overlapped with the set of interband proteins studied previously (RNA polymerase II, CHRIZ, dMi-2, NURF301, WDS) [27,28,41]. Additionally, the cluster included ISWI, JIL-1, MLE, MOF, MRG15, MSL-1 and MBD. Within this cluster, we observed the highest correlation between datasets (see the green frame on the Fig. S2). In order to identify the DNA regions where 12 selected chromatin proteins would preferentially co-localize, we applied principle component analysis (PCA). Closer inspection of the two first principle components (PC1 and PC2) covering 62.4% of the sample variance showed that they scored high in the 12 interband regions described previously [27], particularly, in gene promoter regions (PC1) and in decompacted chromatin regions studied (PC2). To define the tentative interband borders based on PC1 and PC2, we proceeded to the HMM (hidden Markov model) analysis.

Specifically, we performed a series of PC1 and PC2 clustering and allowed the number of states to vary from 2 to 15. When conditioning that all 12 interbands mapped previously consistently group together, a 4-state model was produced. Furthermore, we obtained an independent estimate of clustering quality using Calinski-Harabasz criterion, which similarly returned 4 states (further details are provided in Materials and Methods).

Thus, the genome was partitioned into 4 states. Of these, the first state, which we named "cyan", included all of the experimentally characterized interbands. Further, we identified three more chromatin states which differed in their protein ensembles. Blue chromatin is enriched in RNApolII, although not as high as cyan chromatin is. Notably, blue chromatin is not associated with CHRIZ. Next chromatin state, named magenta, is completely devoid of interband-specific proteins. As for the green-state chromatin, it differs from the above three states in having no obvious protein specificity; hence it is not considered here in detail and these data will be published elsewhere. To summarize, the whole body of chromatin turned out to be divided into 4 states that differed in associated proteins.

As is shown in Fig. 2, localization of proteins that largely define these states varies significantly between different experiments (for

instance, compare CHRIZ WR.S2 and CHRIZ BR.KC profiles). Our mathematical pipeline processes the regions occupied by interband-enriched proteins so that their positions are averaged across the experiments and so the coordinates for localization region are produced that fit best all the individual enrichment profiles. The coordinates thus obtained define the borders of cyan chromatin state. Its span and coordinates on the physical map are used to conditionally define localization of DNA sequences that we attribute to interbands (Fig. 2, Figs. S3–S5).

The fraction of the *Drosophila* genome occupied by the chromatin states identified and the number of fragments are: cyan – 12.7% (5674 fragments), blue – 16.8% (4006 fragments), green – 22.5% (8903 fragments) and magenta – 48.0% (5148 fragments). The sizes of the cyan chromatin fragments range between 0.2 and 39.6 kb (2.7 kb average), blue - 0.2-46.8 (4.9 kb average), green - 0.2-46.8 kb (3.1 kb), magenta - 0.2-82.8 kb (11.2 kb average).

Genome browser-compatible tracks showing the positions of all four chromatin states can be found in the File S1.

**Protein and genomic features in bands and interbands of polytene chromosomes.** In our downstream analysis, we used a set of 32 interbands described above and whose positions on the cytology map were established with maximum accuracy. This set of interbands is also molecularly well-characterized. Therefore, it can be used for fine analysis of interbands, i.e. for mapping of proteins and functional chromatin elements enriched in these interbands. For each of these interbands, detailed maps of associated proteins and other open chromatin features were constructed; Fig. 2 provides an example of such maps.

Interband 7F1-2/7F3-4 shows nearly perfect overlap between the FISH signal from the 5′-*Nrg* probe, cyan state, various promoter types identified by Hoskins et al. [42], and active chromatin marks, such as CHRIZ, RNAPolII, nucleosome remodelers WDS, ISWI, NURF301, H1 dips, promoter-enriched H3K4me3 [50], DHS and red chromatin state (state 1 as defined by Kharchenko et al. [22]). Furthermore, of the entire 7F region only interbands display significant enrichment for ORC2 and NSL complexes.

Very similar trends are clearly observed for all other interbands as well (Figs. S3–S5): although protein profiles do show minor variability, overall most of the protein-enriched regions fall within the borders of cyan chromatin (red dashed line in Fig. 2, Figs. S3–S5).

All 32 interbands studied here display similar organization. First of all, 100% of interbands harbor cyan chromatin fragments, 5′-UTRs of genes, they show low nucleosome density and overlap with the positions of state 1 chromatin (red in 9-state model by Kharchenko et al. [22]) (Fig. 3A, B). Vast majority of interbands also display a number of features characteristic of transcriptionally active regions, namely broad-class promoters, H1 dips, DHS, RNApolII (Fig. 3), CHRIZ and BEAF-32. In all the cytologically defined interbands, we observed enrichment for replication complex components (ORC2) (Fig. 3).

Our set of interbands displays clear enrichment for NSL complex proteins, which have been reported to specifically associate with promoters of multiply active genes [48,49]. If one compares NSL localization data from *Drosophila* salivary gland polytene chromosomes with the interband-mapping data, the overlap is nearly perfect (see Fig. 3B).

What would be the controls for whether the interband fragments have been correctly identified and whether the whole idea of mapping the interbands on a physical map is robust? We believe one of the solutions would be to do a reverse experiment, i.e. to perform polytene chromosome mapping of transposons that we determined to have landed into cyan chromatin fragments.

We selected three matching transposon lines: *w* P{EP}G400*, $y^1$ $w^{67c23}$ *P{EPgy2}Hsp60*$^{EY01572}$ and $y^1$ *P{EPgy2}EY09320* $w^{67c23}$ located in cyan regions. FISH analysis of polytene chromosomes from these transgenic stocks using *white* DNA as a probe (as all these transposons are *white*-marked) shows that for *w* P{EP}G400* the FISH signal maps immediately proximal to 10A1-2, i.e. it is located in the adjacent interband (Figure S6A, D, G). The insertion of *P{EPgy2}Hsp60*$^{EY01572}$ lies a little proximal, in the interband 10A3/10A4-5. Accordingly, FISH signal is also found more proximal, i.e. in the polytene chromosome interband 10A3/10A4-5. In this case, there is a small gap between the band 10A1-2 and the FISH signal (Figure S6B, E, H). Insertion of *P{EPgy2}EY09320* $w^{67c23}$ maps to the interband 10A7/10A8-9 of a physical map, which is consistent with the FISH signal localization in the same polytene chromosome interband (Figure S6 C, F, I).

Using EM analysis of these regions, we found three novel bands in exactly the interband regions we expected (arrows in Fig. 4C-E). This serves as independent and very important evidence arguing in favor of correct identification of interbands, based on the protein localization data in the interbands from the 9F13 to 10B3 region. It must be emphasized that mapping of interbands followed by transposon localization stemmed from protein localization data in interphase chromosomes of mitotically active cells. This allowed tracking the transposon insertions into these same interbands, yet in the context of polytene chromosomes.

These data are consistent with the idea that cyan chromatin fragments correspond to interband DNA, moreover the chromatin features such as ORC2 and DHS preferentially map to interbands as well. Therefore, it is straightforward to analyze the distribution of chromatin features such as ORC2, H1 dips and DHS (the features that were not used to compile the 4-state model) across the genome. Results of this genome-wide analysis are presented in Fig. 5 and show that in Kc cells 85.6% DHS, 91.4% ORC2 and 46.9% sites of histone H1 dips overlap with cyan chromatin. Similar numbers are observed for other cell lines (S2 and BG3) as well as for larval salivary glands (data not shown).

**Genes vs bands and interbands.** As was mentioned above, in polytene chromosomes there are three basic structural types: interbands and two types of bands – large IH bands composed of tightly compacted late-replicating material, and loosely compacted small bands replicating early. High-resolution mapping of chromosomal structures presented in this study makes it possible to match the positions of genes and chromosomal structures on a scale of the physical map. One such comparison for the region 10A is illustrated in Fig. 6.

In polytene chromosomes, two prominent late-replicating IH bands 10A1-2 and 10B1-2 flank a group of six faint loosely compacted bands that are barely detectable under the light microscope. They are depicted grey on Bridges' map [35], and look similarly grey on the EM map (see Fig. 4). Thus, this region encompasses 7 interbands. The schematic figure of this genomic region clearly shows 6 alternating pairs of cyan/blue chromatin states (Fig. 6), much like the number of bands and interbands. In genetic terms, cyan chromatin corresponds to the 5′-regulatory part of the gene (alternatively, small genes are entirely engulfed by cyan chromatin). Given that cyan chromatin is a defining feature of interbands, then blue chromatin can only map to the space between two interbands, i.e. to the neighboring loosely compacted grey bands. On the molecular map, these interleaving blue and cyan chromatin types perfectly mirror the pattern of alternating bands and interbands on the cytology map (Fig. 6).

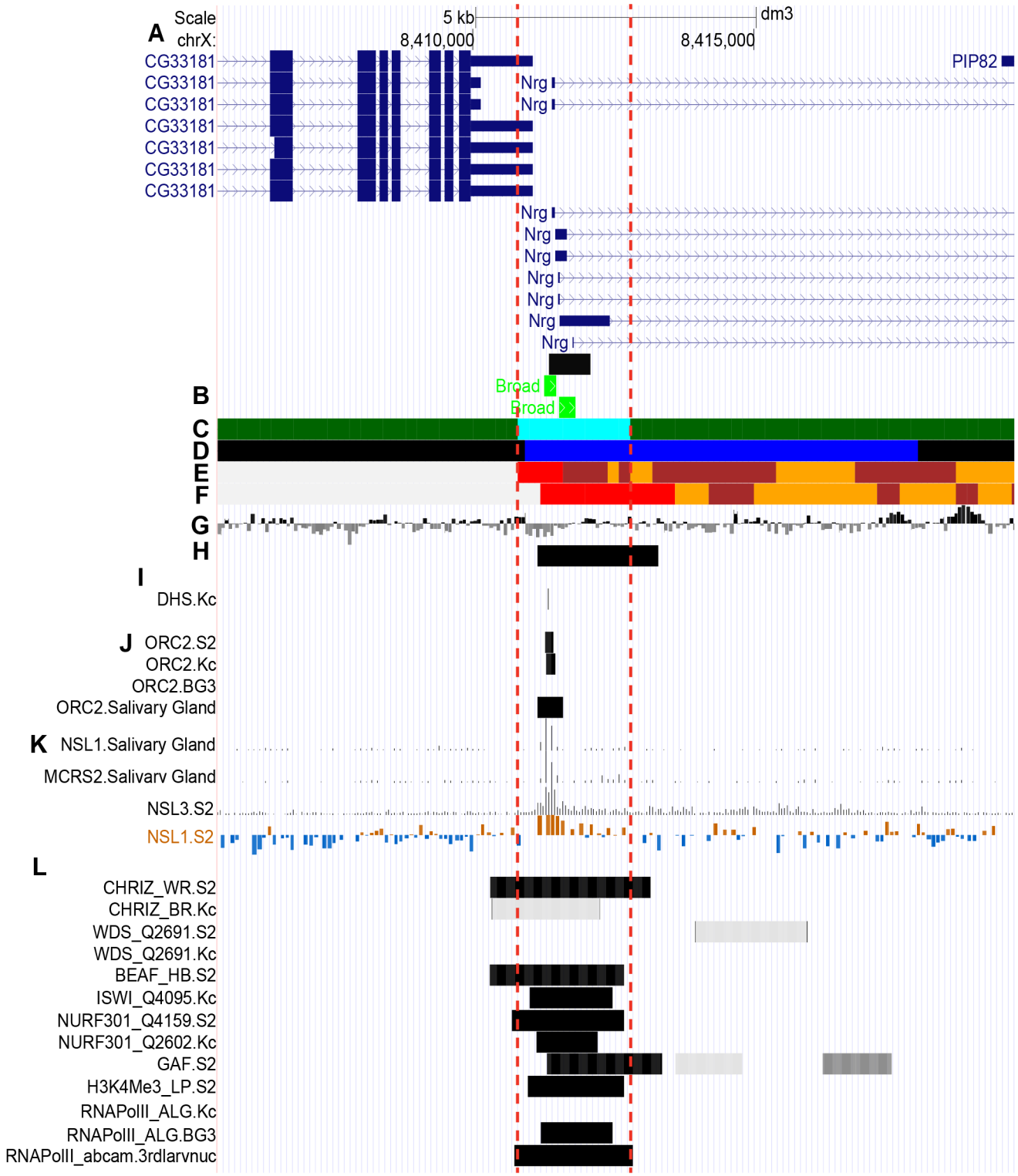



**Figure 2. Localization of proteins and genomic features (fly modENCODE) in the interband 7F1-2/7F3-4.** Vertical red dashed lines delimit the edges of cyan state chromatin in the region which conditionally reflects the location of the interband. **A –** gene map (RefSeq Genes). **B** – Localization of broad class promoters according to Hoskins et al. (2011) [42] (light green rectangles) and a FISH probe (black bar). **C** – Localization of 4-state chromatin types according to the algorithm developed in this paper. Only cyan and green chromatin types map to this genomic region. **D –** Five-state chromatin types in Kc cells by Filion et al. [20]. **E** - 9-chromatin states in S2 cells by Kharchenko et al. [22]. Chromatin state 1 is shown red. **F** - 9-chromatin states in BG3 cells by Kharchenko et al. [22]. Chromatin state 1 is shown red. **G –** Nucleosome density according to Henikoff et al. [43]. Peaks above the axis reflect high density and those below axis denote low nucleosome density. **H** – Localization of histone H1 dips in Kc cells by Braunschweig et al. [44]. Black horizontal bars indicate the genomic regions with low histone H1 binding. **I** - DNAse I hypersensitivity sites (high magnitude DHS - vertical lines) in S2, BG3, and Kc cells by Kharchenko et al. [22]. **J** - ORC2-binding sites in S2, BG3, Kc cells and salivary glands by

Eaton et al. [45], Sher et al. [46]. **K –** Enrichment profiles of NSL complex components: NSL1 binding profile from salivary glands by Raja et al. [47], NSL3 in S2 cells by Lam et al. [48], NSL1 in S2 cells by Feller et al. [49]. **L –** Enrichment regions of various proteins specific for interbands and active chromatin (fly modENCODE). The list of interband-specific proteins is taken from Demakov et al. [28] and Vatolina et al. [27]. (See text for more detailed explanations). list of interband-specific proteins is taken from [27–28,50–53]. (See text for more detailed explanations).
doi:10.1371/journal.pone.0101631.g002

Yet another region, 100B, may serve to further illustrate of this trend. As was mentioned above, this region harbors a group of two interbands and a loosely compacted very faint band 100B3 (Figs. S1) barely visible under EM. These structures are flanked by late-replicating bands 100B1-2 and 100B4-5 [34] (green fragments on Fig. S3). The two interbands distal and proximal to 100B3 display all the features characteristic of interbands. On a physical map, the coding sequence of *dco* gene of about 3 kb maps between these

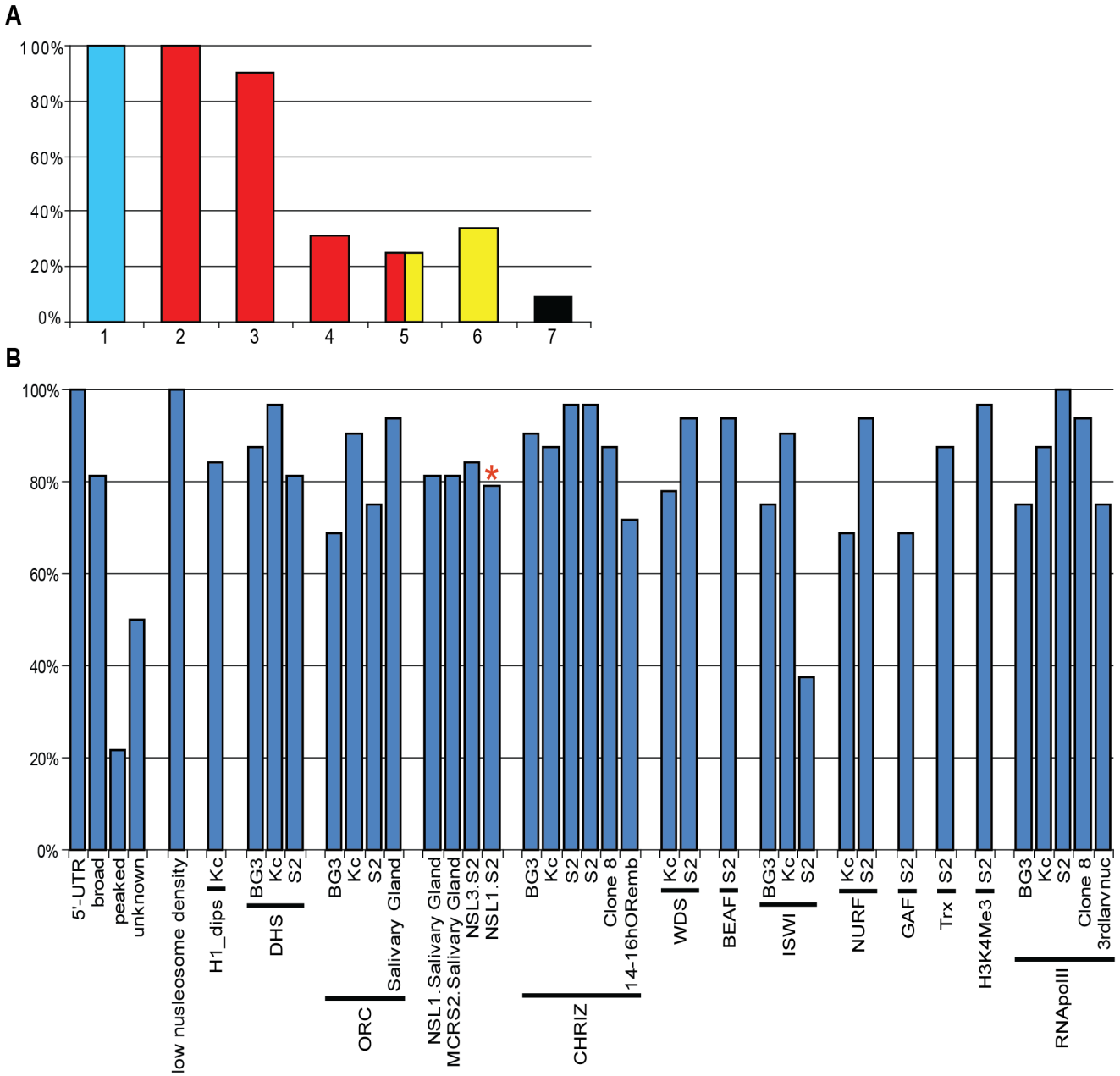



**Figure 3. Positions of protein enriched regions, chromatin states and other genomic features in 32 cytologically defined interbands.** A – Frequencies of chromatin states in select 32 interbands: cyan (1-this paper), state 1 (red chromatin) in 9 state chromatin model in BG3 (2) and state 1 (red chromatin) in S2 cells (3-according to [22], RED, YELLOW, RED/YELLOW and BLACK/BLUE chromatin types (4–6) according to Filion et al. [20]. B – Occurrence of various chromatin states, proteins and other genomic features in interbands. X axis - proteins or genomic features found in different cell cultures. Y axis shows a fraction of interbands demonstarting these characteristics. *Since some of the data were originally missing from the analysis, the frequencies are presented for a set of just 19 interbands.
doi:10.1371/journal.pone.0101631.g003

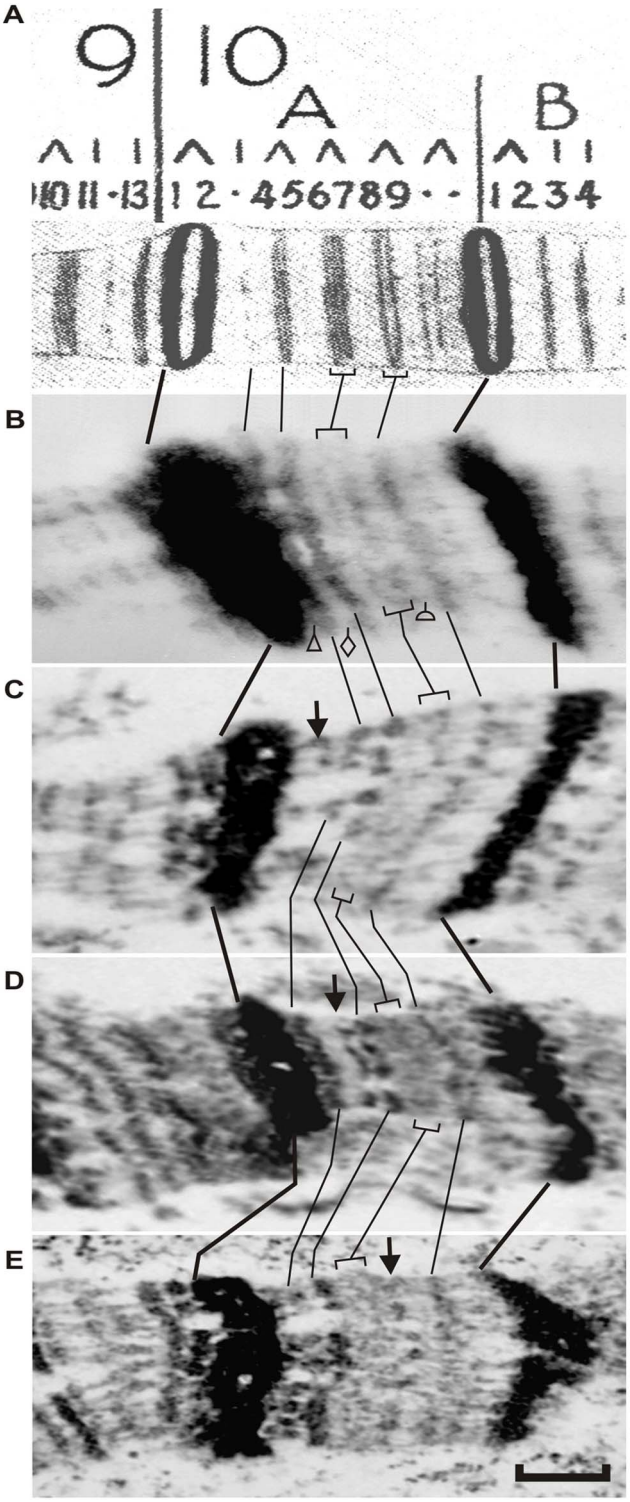



**Figure 4. EM mapping of transposon insertions in the predicted interbands, 10A1-2/10A3 (C), 10A3/10A4-5 (D), and 10A7/10A8-9 (E) in salivary gland polytene chromosome X.** A – Fragment of the Bridges' map of the X chromosome [35]; B – Electron micrograph showing morphology of the region 10A in wild-type X chromosome; C - E – polytene chromosomes in *Drosophila* lines containing w* P{EP}G400, $y^1w^{67c23}$P{EPgy2}Hsp60$^{EY01572}$ and $y^1$P{EPgy2}EY09320$w^{67c23}$ transposons, respectively. Location of the insertion site of P{EP}G400 transposon on the chromosome map of wild type is indicated by triangle in wild type chromosome (B) and by an arrow in the insertion line (C). For insertion $y^1w^{67c23}$P{EPgy2}Hsp60$^{EY01572}$, the integration site is shown as diamond (B) and arrow (D). The same for insertion of $y^1$P{EPgy2}EY09320$w^{67c23}$: semicircle (B) and arrow (E). Bar corresponds to 1.5 mkm.
doi:10.1371/journal.pone.0101631.g004

two interbands, hence it likely corresponds to the miniature band 100B3. Thus, in the regions where interbands alternate with grey bands, the interbands generally tend to comprise 5′-ends of the genes and correspond to cyan chromatin state, whereas grey bands harbor gene coding sequences. This observation forms the basis of a hypothesis that in the context of interphase chromosome, many other regions may share the same pattern of genetic organization (interband – 5′-end, loose grey band – gene coding sequence).

So, the genes mapping to grey bands/interbands, in fact, may occupy two polytene chromosome structures: interband hosts the 5′-end of a gene encompassing its regulatory part, the first exon and intron, whereas the structural part of the gene is found in the neighboring loosely compacted grey band. In this respect, localization of cyan fragments serves as a marker of 5′-ends of genes and interbands, while blue chromatin may correspond to loosely compacted grey bands and coding parts of the genes. Thus, loosely compacted grey bands and the adjoining interbands can be viewed as linked structures: gene promoter locates to the interband, gene coding region resides in the adjacent grey band.

In the set of 32 interband/grey bands studied here, we identified 65 genes (Table S1). Late-replicating IH bands are dramatically different from the loosely compacted grey bands in that they comprise densely packed chromatin: they replicate late, they lack typical interband- or grey band-specific proteins, and instead associate with SUUR, D1, lamin B, histone H1 and other proteins characteristic for BLACK chromatin [20,23,33]. 238 genes were found in the late-replicating IH bands 7F1-2, 7F3-4, 10A1-2, 10B1-2, 19E1-4, 21D1-2, 21E1-2, 35D1-4, 56A1-5, 58A-B1-2, 70A1-5 and 100B1-2 – 100B4-5 located next to the 12 reference interbands (Table S1).

**Preferential integration of P-elements in interbands as a feature of open chromatin.** Our earlier analysis of 12

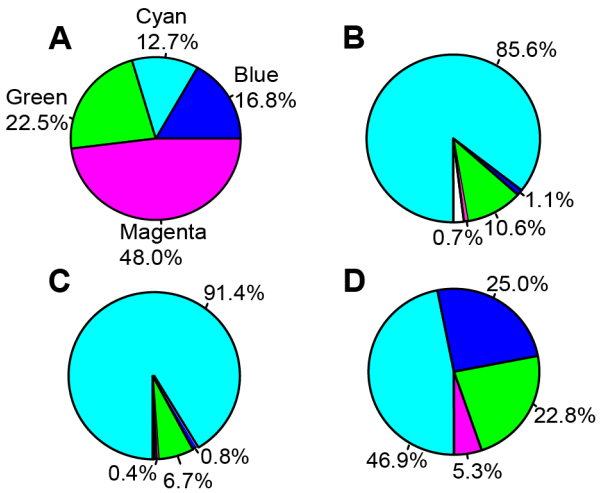



**Figure 5. Pie charts showing proportions of four chromatin types in the genome (A), genome-wide distribution of DHS (B), ORC2 sites (C) and histone H1 dips (D) in these chromatin types in the *Drosophila melanogaster* genome (Kc cells).** Chromatin types are color-coded according to their names: cyan, blue, magenta, green. Unshaded sectors denote absence of the data in modENCODE: 2% (B) and 0.7% (C).
doi:10.1371/journal.pone.0101631.g005

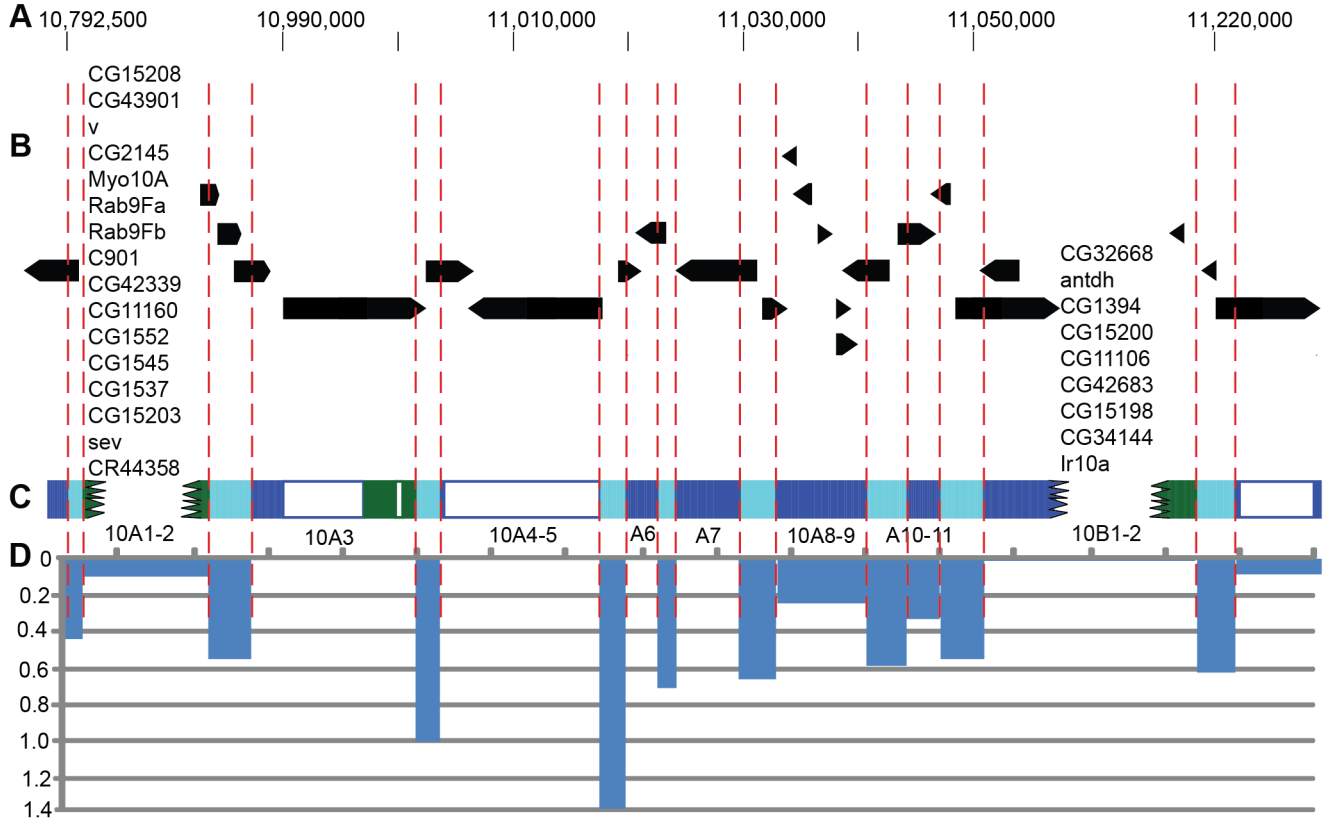



**Figure 6. Comparison of the banding pattern and gene positions in the region 9F13 – 10B3 of the X chromosome drawn to the same scale.** A – Genomic coordinates. B – positions of genes (RefSeq genes) on the physical map. Genes in the bands 10A1-2 and 10B1-2 are listed randomly and the extents of these bands are depicted as a jagged line. C – location of four-state chromatin types described in this study. Vertical red dashed lines delimit the borders of bands (labeled below) and interbands according to the borders of cyan fragments. Unshaded areas reflect absence of the data in modENCODE. D – frequencies of P-element transposon insertions (per 1 kb) averaged across the span of a band or an interband.
doi:10.1371/journal.pone.0101631.g006

interband regions from polytene chromosomes [32] reported that in the *Drosophila* genome, P-elements preferentially integrate into interbands. Genome-wide analysis was indicative of predominant integration of P-element transposons into replication origins (ORC-positive regions) [54], which in turn tend to largely locate to interbands [26,27].

We further confirm this observation, as we compare the insertions of transposons and chromosomal structures. Clearly, the interbands serve as the hotspots of transposon insertions, which is evident on the molecular and genetic maps of all interbands studied to this end (Fig. 6D).

When this analysis is performed genome-wide, and localization data for all the P-element insertions referenced in the FlyBase (38,888 insertions) are used, 78.3% of all insertions map to the cyan chromatin states, which constitutes only 12.7% of the genome sequence (see above). This translates into 6.2-fold higher frequency compared to the random distribution of insertions. Importantly, we observe a pronounced decrease in insertion frequencies in other chromatin types. P-element based transposons are 9.1-fold less likely to land in magenta chromatin (6.0 and 2.0-fold for blue and green, respectively (Fig. 7B) ($X = 108327.3$, p value $<2.2e-15$, Mann-Whitney test).

P-elements are known to typically transpose in diploid germline cells. On the other hand, we and others have observed that P-element transposons tend to insert into open chromatin regions [32]. Consequently, one may speculate that regions appearing as

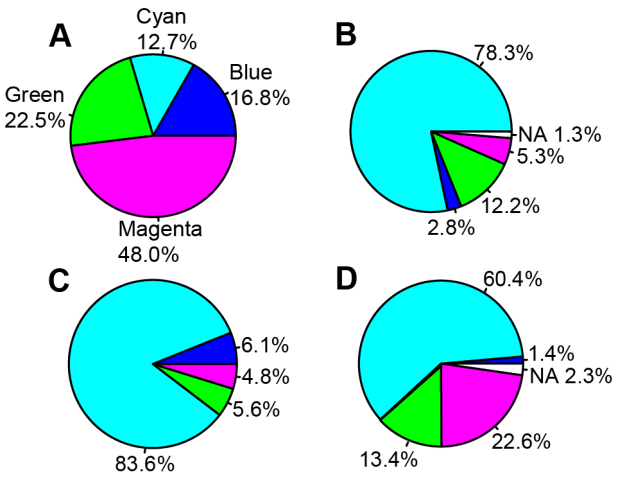



**Figure 7. Pie charts showing frequencies of various genomic features mapping to the different four-state chromatin types (A): P-element insertions (B), localization of "broad" type promoters (C) and ''head-to-head'' arrangement of genes (D) in the *Drosophila melanogaster* genome.** (see text for further explanations).
doi:10.1371/journal.pone.0101631.g007

interbands in the context of polytene chromosomes should be similarly found in an "open chromatin" state in the germline.

**Interbands correspond to 5′ regions of ubiquitously active genes.** Distinct chromatin organization observed for genes residing in three basic types of polytene chromosome structures selected for this analysis (IH bands, interbands and loosely compacted grey bands) is suggestive of their specific organization and function. We compared expression patterns of genes found in the late-replicating IH bands vs those located in 32 interbands/grey bands (see the list in Table S1). Quantitation of gene expression across 8 larval and 17 adult organs has been reported in Chintapalli et al. (FlyAtlas) [55]. We observed that gene activity varied depending on which structure the gene mapped to.

In the region 9F13 – 10B3, interband/grey band genes (genes within the interband 9F13/10B1-2, all genes between 10A1-2 and 10B1-2, and in interband 10B1-2/10B3) are active across almost all tissues analyzed. At the same time, expression pattern of genes found in large IH bands 10A1-2 and 10B1-2 is much more restricted (Fig. 8). Similar trends were also observed when comparing gene activity in other interband/grey bands vs IH bands mapped in this study (Fig. 9). On average, the transcripts from genes that overlap with the studied interbands are likelier to be present in more tissues, as compared to the transcripts mapping to IH bands, - both in larvae and adults (p-value $\leq 2.2e^{-16}$ Mann-Whitney test).

To compare the magnitude of transcription for the genes found in IH bands and 32 interbands (see the list of genes in Table S1), we calculated average expression levels for each gene present on the microarray (FlyAtlas Anatomical Expression Data) [55]. We estimate that interband-resident genes have 27-fold higher median expression (112 vs. 4.2) than those mapping to IH bands (Figure S7).

Next, we performed a more comprehensive genome-wide analysis of gene activity in cyan and magenta chromatin states. On average, median expression levels for genes in cyan fragments was 20.9 times higher than those in magenta chromatin (100.4 vs 4.8) (Figure S8).

Recently, many interesting datasets characterizing *Drosophila* transcriptome have been published, utilizing high-throughput RNA-seq technique (Gelbart, W.M., Emmert, D.B., 2013 FlyBase High Throughput Expression Pattern Data). For each gene, abundance of RNA transcripts was measured throughout development (whole organism, 30 developmental stages) [55], in different organs (29 dissected tissues) (modENCODE Tissue Expression Data), in cell lines of various origin (24 cell lines) (modENCODE Cell Line Expression Data), and upon different experimental treatments (21 treatments of whole animals with various toxins) (modENCODE Treatment Expression Data) (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3032933/).

Using these data, we confirmed that expression of 238 genes mapping to IH bands (see the list in the Table S1) is significantly lower (median expression is almost zero) than expression levels observed for 32 interband genes (median expression is 15 RPKM) (Figures S9–S10).

On a genome-wide scale, we observed maximum activity of genes whose 5′-ends mapped to cyan chromatin fragments (median expression values are around 11) in comparison with magenta fragments (median expression values are close to 0) (Figures S11–S12).

Taken together, these data are indicative of the constantly high transcription level of genes whose 5′-ends locate to the interbands or to the cyan chromatin state in the genome. These genes are active across the majority of cell types (i.e. they may be referred to as ubiquitously active or multiply active genes).

**Promoter architecture in genes residing in interbands and bands.** Recent high-resolution mapping of promoters active in the *D. melanogaster* embryos identified 12454 promoters in 8037 genes. It was shown that distribution of transcription start sites (TSS) within these promoters forms a complex continuum of shapes, and that promoters active in the embryo and adult have highly similar shapes in 95% cases. This led to the conclusion that these distributions are generally determined by static elements such as local DNA sequence, rather than by dynamic signals such as histone modifications. As it turned out, *Drosophila* promoters are characterized either by broad region of distributed TSSs, or by single TSS defining a discrete promoter. These patterns are consistent with a definition of "broad" and "peaked" promoter classes in the human and mouse genomes [42]. Peaked promoter shape is correlated with both temporal and spatial regulation of gene expression [42]. Of the 32 interband regions studied in the present paper, 26 encompass gene promoters of various architectures, yet nearly invariably including broad class promoters (81.2% of interbands); 6 interbands lack broad promoters, as they host either unknown or peaked promoter classes (18.8%). Thus, interbands generally include broad promoters.

On a genome-wide scale, 83.6% (6436) of broad-class promoters map to cyan (interband-enriched) chromatin fragments (Fig. 7C). This argues that broad promoters predominantly map to interband regions.

Interestingly, analysis of promoter architecture in the human and *Drosophila* genome showed high frequency of bi-directional promoters activating expression of genes arranged in a "head-to-head" configuration with less than 2000 base pairs of intervening sequence [57,58]. It is possible that such gene orientation contributes to enhancing gene transcription activation and elongation.

Of the genes located in 32 interbands listed in Table S1, bi-directional promoter architecture was observed in 13 cases (41% interbands). In the *Drosophila* genome, there are 3357 head-to-head oriented genes with less than 2000 base pairs between them (FlyBase). Of these, 5′-ends of 2027 genes (60.4%) map to cyan chromatin, which constitutes only 12.7% of the genome. In stark contrast, only 46 "bi-directional" genes (1.4%) are found within blue chromatin (16.8% genome). Green (22.5%) and magenta chromatin types (48.0%) encompass 448 (13.4%) and 760 "bi-directional" genes (22.6%), accordingly (Fig. 7D). Clearly, interbands are enriched for this particular class of genes in a head-to-head orientation, whereas all other chromatin types are significantly depleted for this feature (X-squared = 1208.28, p-value< 2.2e-16).

## Discussion

The present study aims at unraveling the genetic and functional organization of basic morphological features of interphase chromosomes. In the context of polytene chromosomes, these features display distinct degrees of chromatin packaging and comprise interbands, loosely compacted grey bands and dense IH bands. We attempted to correlate positions of gene elements, gene expression and the epigenetic state of underlying chromatin for these structures. To do so, we first had to accurately locate these morphological elements on the physical map of the genome. This allowed us to compare their positions with genetic and epigenetic maps, as well as with protein localization profiles, transcription profiles and other features of chromatin. So, we could relate

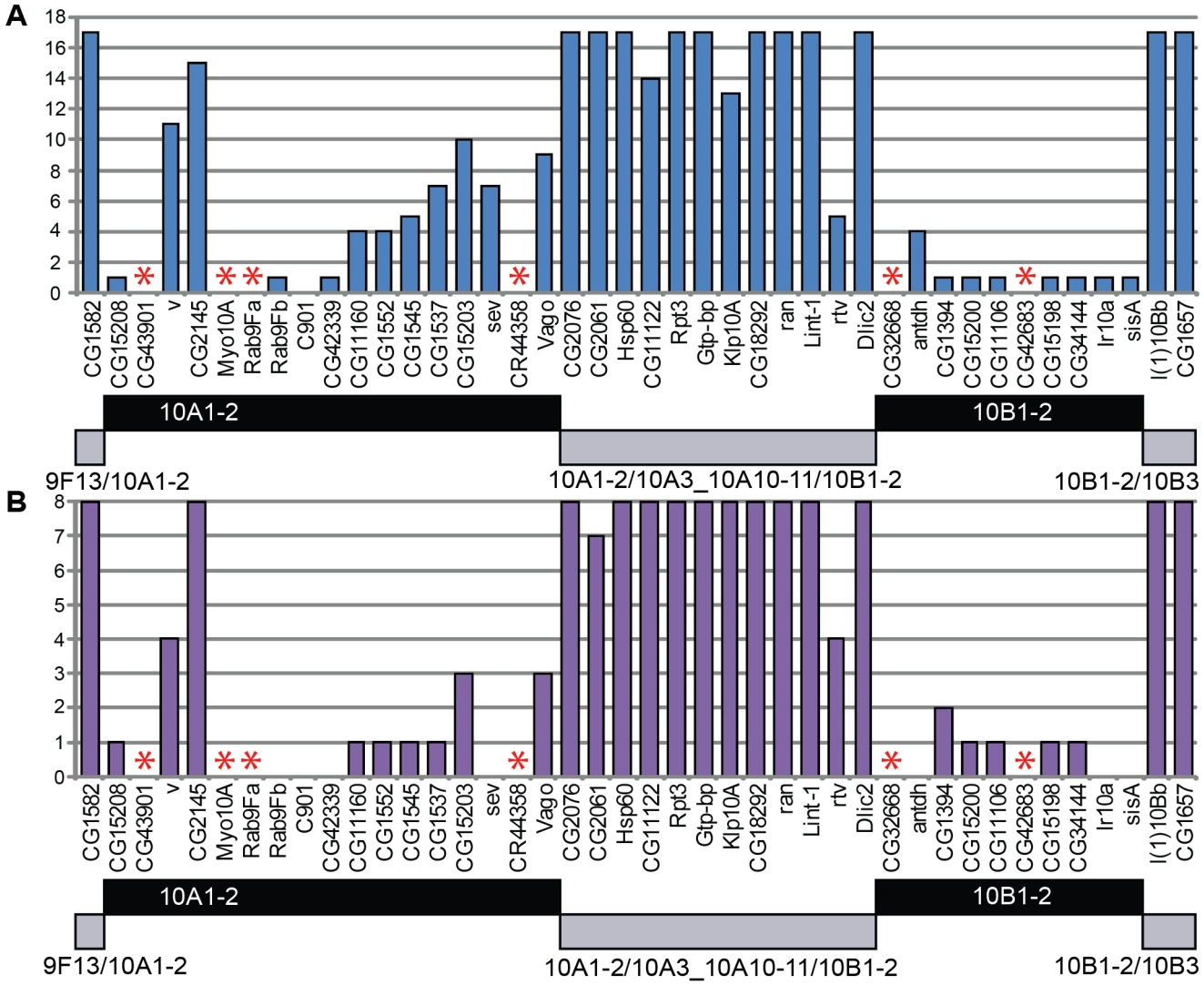



**Figure 8. Activity of ''band'' and ''interband'' genes located in 9F13 – 10B3 region.** Genes are listed below the *x* axis. The number of tissues where gene activity was found in adult flies (totally 17 tissues studied) (**A**) and larvae (totally 8 tissues studied) (**B**) is plotted on the *y* axis. Data on gene activity were taken from Chintapally et al. [55]. Horizontal bars below the *x* axis denote the extent of the bands 10A1-2 and 10B1-2 (black) and alternating interband/grey bands (grey), as well as two interbands (grey) on the very edges of the region. Asterisks – NO data.
doi:10.1371/journal.pone.0101631.g008

functional domains with the banding pattern of polytene chromosomes.

Dense black bands are the most prominent structures in polytene chromosomes. They are readily noticeable due to their highly compacted state, large size, lack of transcription, late replication in the S phase, and a tendency to form ectopic pairing with other bands and pericentric heterochromatin. In fact, black bands are in many regards very similar to pericentric heterochromatin, hence they were called IH [59,60]. In polytene chromosomes, IH bands frequently fail to complete replication during the S phase endocycles, and are therefore underreplicated. It has recently become clear that underreplication results from the absence of internal replication origins within IH and is dependent on SUUR protein, which maps to IH bands and modulates replication [34,61] by decreasing the rate of replication fork progression [46].

Underreplication regions showing lowered DNA copy number in polytene chromosomes were molecularly mapped [25,61,62]. This analysis established that IH bands encompass clusters of widely-spaced unique genes (i.e. genes with large intergenic regions, with 6–40 genes per IH band), and that they are generally quite large (100–600 kb) [24,25]. Combined with the data on localization of chromatin proteins [20,22], IH borders were precisely mapped for 60 IH regions, which enabled a more refined analysis of these structures [26].

Our data (present paper) and those of Filion et al. [20] indicate that IH bands are composed of tissue-specific genes showing low expression levels. One of the prominent features of IH regions is their evolutionary conservation, i.e. they tend to display conserved gene content and order throughout evolution, as has been demonstrated by microsynteny analysis in nine *Drosophila* species [63].

As compared to IH bands, it is far less trivial to provide accurate mapping for interbands and grey bands, because these regions are fully replicated and are much smaller. Yet, using a combination of EM, P-element tagging and FISH, we were able to unambiguously

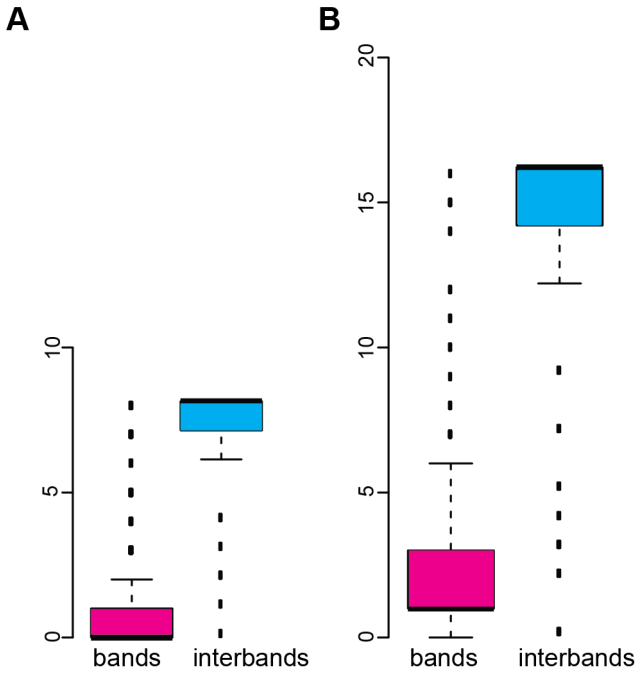



**Figure 9. Box-and-whiskers plot showing mean number of larval and imaginal tissues where the activity of "band" and "interband" genes was found.** Y axis: number of tissues profiled in larvae (A) and imago (B). Median positions are shown as thick black bars. Error bars (blue and red boxes) represent the data between 25 and 75 percentile (50% of data points). Dataset range between the largest and the smallest values is shown as thin horizontal bars. Dashed line represents outliers. To construct this plot, data from Chintapally et al. [55] were used.
doi:10.1371/journal.pone.0101631.g009

map the positions of 32 interbands. Using the data on the features of interband chromatin, we developed a mathematical model that defines four basic chromatin states in the drosophila genome. This model allows identification of interband regions chromosome-wide. Accordingly, the limits of the DNA sequences corresponding to interbands were defined as borders of cyan fragments.

With these data in hands, we proceeded to analyze the molecular and epigenetic organization of interbands. Interbands clearly displayed features of transcriptionally active regions: H3K4me3 histone modification, lower nucleosome density and histone H1 dips, presence of DHS, localization of RNA polymerase II and components of nucleosome remodeling complexes such as NURF, ISWI, WDS. One characteristic feature of interbands is that they are specifically bound by the chromodomain-containing CHRIZ protein [41,51,52,64]. CHRIZ associates with another interband-specific protein Z4, which directly binds DNA via its seven zinc fingers [52,65].

According to different estimates, there are 3500–5000 bands and interbands in *Drosophila melanogaster* polytene chromosomes [18]. Earlier, we predicted the existence of about 3500 interbands [28]. Here, we use an advanced model that takes into account more factors and hence is more accurate. We found 5674 cyan fragments each spanning 2.7 kb on average. Notably, both our previous and current estimates of interband numbers are very close to those obtained by cytology.

The major finding of our analysis of functional organization of interbands is that they typically encompass 5′-regions of multiply active genes (constitutively and actively transcribed). In a number of instances, we observe that short genes can be entirely engulfed by interbands (see Fig. 5), however in most cases the body of the gene is found in the adjacent loosely compacted grey band. Thus, the interband+grey band duo appears as a single functional unit for many multiply active genes, so this unit is heterogeneous in terms of compaction; likewise it shows non-uniform localization of protein markers. Whereas interbands are specifically decorated with CHRIZ, grey bands lack CHRIZ and instead they are enriched with RNApolII. CHRIZ can be speculated to provide the permanently open chromatin state to interbands, where it serves as a pioneer-factor recruiting other transcription components. It is also possible that the observed wide-spread transcription activity of interband regions results from the static physical properties of interband DNA, such as sequence-dependent DNA flexibility, which may create nucleosome-free regions at promoters. Such regions may serve as "entry points" to recruit proteins promoting further binding of transcription factors, chromatin remodelers, etc [66].

Our findings, therefore, resonate well with several early ideas regarding the interplay of structural and functional organization of banding pattern in polytene chromosomes. These include interband localization of multiply active genes, and the hypothesis of a single functional unit composed of band+interband (see Introduction).

Recently, there has been an avalanche of publications describing various types of domain organization in the genomes of eukaryotes [20,22,67,68]. So, our data can be conveniently compared with other genome-wide chromatin annotation projects. Figure S13 summarizes domain organization of a 400 kb fragment of the X chromosome encompassing various types of bands and accurately mapped interbands. This figure shows that two large domains, 189 and 170 kb long, correspond to polytene chromosome bands 10A1-2 and 10B1-2, and display features of intercalary heterochromatin (magenta-green chromatin states) (Figure S13 A-B). In between these late-replicating domains, there is a region composed of alternating interbands and grey bands (cyan and blue fragments) (Figure S13 B). When applied to this region, 5-state chromatin classification model by Filion et al. [20] produces very similar domains, - the important difference however is that regions of YELLOW chromatin (active gene transcription according to Filion et al. [20]) do not discriminate between small grey bands and interbands, nor between regulatory vs gene body parts (Figure S13 C).

Kharchenko et al. [22] performed genome-wide profiling of 18 histone modifications and constructed 9-state chromatin models. As is shown in Figure S13 D, E, in two contrasting cell lines (S2 and BG3), transcriptionally silent chromatin corresponds to the IH bands 10A1-2 and 10B1-2, whereas state 1 chromatin (active promoters and TSS) (shown red in the Figure S13 D, E) maps to the active genes and perfectly matches the cyan state of interbands, as defined by our analysis.

Using the modified Hi-C approach based on the ligation of chromatin fragments that are found in close proximity in cross-linked chromatin, high-resolution chromosomal contact maps were generated (reviewed in [21]). As it follows from this analysis, the entire genome is partitioned into a series of physical domains containing active and repressive epigenetic marks. These domains are delimited by boundaries demonstrating insulator binding, high DNaseI sensitivity and a set of specific proteins: CHRIZ and the active histone mark H3K4me3 [67,68]. The regions of interbands match well with the boundary sites that delimit the contacting domains identified via Hi-C (Figure S13 F, G). We compared localization of cyan chromatin and physical domains throughout the genome. Of 1100 boundary sites referenced in [67], 760 (69%)

map to interbands (cyan) and loosely compacted grey bands (blue chromatin).

Positions of interbands and loosely compacted grey bands display co-localization with clusters of multiply active genes [49,69] (Figure S13 H, I). According to our analysis, there are 12 genes nested between the 10A1-2 and 10B1-2, with their 5′-ends mapping to cyan chromatin. Of these, 5 genes were classified as housekeeping genes in Weber and Hurst [69].

Using the data from Chintapalli et al. [55], Feller et al. [49] also classified the genes as "housekeeping" or "differentially regulated" (Figure S13 I). Under this classification, the region 10A1-2 – 10B1-2 harbors 11 housekeeping genes, of which 9 genes correspond to our definition of a multiply active gene.

NSL complexes are reportedly regulators of multiply active genes and bind promoters demonstrating broad transcriptional pattern and nucleosome-free regions [48]. Positions of NSL-binding peaks match nicely the interband positions found in our study (Figure S13 K-M). All these comparisons further confirm the main conclusion of our work about interbands as sites of continuously active genes.

## Materials and Methods

### *Drosophila* stocks

*Drosophila* stocks containing insertions of transposons *w* P{EP}G400*, $y^1$ $w^{67c23}$ *P{EPgy2}Hsp60*$^{EY01572}$ and $y^1$ *P{EPgy2}EY09320* $w^{67c23}$ were used. The stocks were kindly provided by the Bloomington Drosophila Stock Center. Flies were raised on standard cornmeal–yeast–agar–molasses medium [70].

### Electron Microscopy

Salivary gland polytene chromosome squashes were prepared for electron microscopy analysis and examined as described earlier [31,70]. The 120–150 nm sections were cut using an LKB-IV (Sweden) ultratome and examined under a JEM-100C (JEOL, Japan) electron microscope at 80 kV.

### Fluorescence in situ hybridization (FISH)

Salivary glands were dissected in Ephrussi-Beadle solution, and then fixed in a 3:1 mixture of ethanol and acetic acid for 30 minutes at −20°C, squashed in 45% acetic acid, snap-frozen in liquid nitrogen and stored in 70% ethanol at -20°C. Fluorescence in situ hybridization (FISH) on polytene chromosomes was performed as described [71]. Random-primed labeling of DNA probes with biotin-16-dUTP or digoxigenin-11-dUTP (Roche) was done using Klenow enzyme. All the probes used in this study are described in the Table S2.

### Immunostaining of polytene chromosomes

Immunostaining was performed as described [41,52,72]. For CHRIZ detection, primary rabbit polyclonal anti-CHRIZ antibody (1:600 dilution) and secondary FITC-labeled goat anti-rabbit IgG-specific conjugates (Abcam, 1:200) were used.

### Genome-wide analysis of P-element insertion sites

To analyze the distribution of P-element insertions within the euchromatic part of the genome (X, 2L, 2R, 3R and 3L arms, as defined in FlyBase), we used insertion coordinates tagged as "transposable_element_insertion_site" from FlyBase release 5.50. We had to exclude 718 insertions (1.85% of the total sampling) from further analysis, as they mapped to multiple chromatin types or to the regions where the model failed to return a specific value.

### Analysis of promoter shape

Using R-language, we quantitatively described the distribution of promoter shapes [42] across chromatin types predicted by our model.

### Analysis of gene orientation

To analyze promoter orientation in pairs of adjacent genes, we used gene models from FlyBase release 5.50. Gene pairs were chosen so that no annotated gene bodies mapped between these genes. A custom R script (available upon request) was used to analyze the overlap of gene pairs with chromatin states in our model. We assumed that the chances of two genes being oriented "head-to-head", "tail-to-tail" and "head-to-tail" under random distribution would be 25%, 25% and 50%, respectively.

### Gene expression analysis

Two data sources were used for gene expression analysis. For processing the data from the FlyAtlas project [55], we used genomic coordinates from Affymetrix Drosophila Genome Tiling 2.0R Array release 5.33 (gene names and sequences from FlyBase 5.3 were used to design this array release). We also used FlyBase High Throughput Expression Pattern Data (Gelbart, W.M., Emmert, D.B., 2013, in FlyBase, http://flybase.org/static_pages/feature/previous/articles/2013_05/rna-seq_bulk.html, description: http://flybase.org/reports/FBrf0221009.html). Genomic coordinates of both expression datasets were converted so as to match the FlyBase release 5.50.

### Data processing and HMM clustering

Data from the fly modENCODE project were extensively used in our analyses and appropriate datasets are listed in the Table S3 (modENCODEFiles.csv). Custom scripts in R language [73] were used for data processing. To perform principal component analysis (PCA) on chromatin protein profiles, we used *prcomp* function in R. Since we analyzed pre-normalized log2-converted ChIP-chip data, no scaling was applied. Data clustering was done using *RHmm* library package in R [74]. To simplify comparisons and statistical analysis of log2-scaled ChIP-chip data, we subdivided the genome into consecutive 200 bp-long fragments using a non-overlapping sliding window method. Protein localization data were processed using HMM-based script with multivariate normal emission distributions, determined from the Baum–Welch algorithm, applying two first principal components to a set of 12 reference interband regions. Four states were chosen as appropriate, based on the Calinski-Harabasz criterion [75,76] applied for 14 models trained on X-chromosomal sequences with the number of states ranging from 2 to 15 (Fig. S14). Regions of enrichment for 12 interband regions were determined from the Viterbi path. Interbands were identified as one of these four HMM states (cyan), as all 12 P-element insertions previously mapped in interbands [32] were within the cyan domains. All scripts used to download and process modEncode datasets and to obtain the 4-state chromatin model can be found in the File S3 (hmm.tar.gz). The final file for loading our 4-state model data as a genome browser track is provided in the File S1 (hmm4.bed.gz).

## Supporting Information

**Figure S1 Cytological identification of the interbands in the 100B region of the *Drosophila melanogaster* 3R chromosome. A, D** – Comparison between Bridges revised map [39] (**A**) and Electron Microscopic map of the region 100A (**D**) (scale represents three micra). Increased part of the interbands proximally and distally to the 100B3 small grey band is shown in

the rectangle. **B, C** - Immunofluorescent localization of CHRIZ in the region. The interbands 100B1-2/100B3 and 100B3/100B4-5 are marked by asterisks. **E, F** - FISH localization of the DNA containing a fragment of the *dco* gene (arrow in **F**), phase contrast as a control (**C**).
(TIF)

**Figure S2 Heatmap showing correlation between protein distributions along the X chromosome.** Spearman correlation matrix between protein binding data on the X chromosome (ChIP-chip, modENCODE Consortium, 2010). Pairwise correlation values are presented and color-coded according to the color map shown on the bottom left. Spearman correlation distances are illustrated by the dendrogram on the left of the graph. The cluster of proteins to be analyzed in more detail was identified using the appropriate stopping rule [40] and is highlighted as a green frame.
(TIFF)

**Figure S3 Localization of proteins and genomic features (fly modENCODE) around the interbands 100B1-2 - 100B4-5.** Dashed red vertical lines show the edges of cyan state chromatin in the region which conditionally reflect the location of the interbands. **A –** gene map (RefSeq Genes). **B** – Localization of promoter broad, unknown and peacked types according to Hoskins et al. (2011) [42] (light green, red and blue rectangles) and probes for FISH (black). **C** – Localization of 4-state chromatin types according to the algorithm developed in this paper. Only cyan and green chromatin types map to this genome region. **D –** Five-state chromatin types in Kc cells by Filion et al. [20]. **E** - 9-chromatin states in S2 cells by Kharchenko et al. [22]. Chromatin state 1 is marked with red. **F** - 9-chromatin states in BG3 cells by Kharchenko et al. [22]. Chromatin state 1 is marked with red. **G –** Nucleosome density according to Henikoff et al. [43]. Peaks above the axis reflect high density and those below axis denote low nucleosome density. **H** – Localization of histone H1 dips in Kc cells by Braunschweig et al. [44]. Black horizontal bars indicate the genomic regions with low Histone H1 binding. **I** - DNAse I hypersensitivity sites (high magnitude DHS - vertical lines) in S2, BG3, and Kc cells by Kharchenko et al. [22]. **J** - ORC2-binding sites in S2, BG3, Kc cells and salivary glands by Eaton et al. [45], Sher et al. [46]. **K –** Enrichment profiles of NSL complex components: NSL1 binding profile from salivary glands by Raja et al. [47], NSL3 in S2 cells by Lam et al. [48], NSL1 in S2 cells by Feller et al. [49]. **L –** Enrichment regions of various proteins specific for interbands and active chromatin (fly modENCODE). The list of interband-specific proteins is taken from Demakov et al. [28] and Vatolina et al. [27].
(TIF)

**Figure S4 Localization of proteins and genomic features (modENCODE) around the interband 1A8/1B1-2.** Red dotted vertical lines are according to edges of cyan state chromatin in the region which conditionally reflect the location of the interbands. **A –** gene map (RefSeq Genes). **B** – Localization of pICon3C(1A) reference transposon, of promoter broad and unknown types according to Hoskins et al. (2011) [42] (light green and red rectangles). **C** – Localization of 4-state chromatin types according to the algorithm developed in this paper. Only cyan and green chromatin types map to this genome region. **D –** Five-state chromatin types in Kc cells by Filion et al. [20]. **E** - 9-chromatin states in S2 cells by Kharchenko et al. [22]. Chromatin state 1 is marked with red. **F** - 9-chromatin states in BG3 cells by Kharchenko et al. [22]. Chromatin state 1 is marked with red. **G –** Nucleosome density according to Henikoff et al. [43]. Peaks above the axis reflect high density and those below axis denote low nucleosome density. **H** – Localization of histone H1 dips in Kc cells by Braunschweig et al. [44]. Black horizontal bars indicate the genomic regions with low Histone H1 binding. **I** - DNAse I hypersensitivity sites (high magnitude DHS - vertical lines) in S2, BG3, and Kc cells by Kharchenko et al. [22]. **J** - ORC2-binding sites in S2, BG3, Kc cells and salivary glands by Eaton et al. [45], Sher et al. [46]. **K –** Enrichment profiles of NSL complex components: NSL1 binding profile from salivary glands by Raja et al. [47], NSL3 in S2 cells by Lam et al. [48], NSL1 in S2 cells by Feller et al. [49]. **L –** Enrichment regions of various proteins specific for interbands and active chromatin (fly modENCODE). The list of interband-specific proteins is taken from Demakov et al. [28] and Vatolina et al. [27].
(TIF)

**Figure S5 Localization of proteins and genomic features (modENCODE) around the interband 10A7/10A8-9.** Red dashed vertical lines are according to edges of cyan state chromatin in the region which conditionally reflect the location of the interbands. **A –** gene map (RefSeq Genes) and position of the reference transposon insertion P{EPgy2}EY09320 (arrow). **B** – Localization of promoter broad, peacked and unknown types according to Hoskins et al. (2011) [42] (light green blue and red rectangles). **C** – Localization of 4-state chromatin types according to the algorithm developed in this paper. Only cyan and green chromatin types map to this genome region. **D –** Five-state chromatin types in Kc cells by Filion et al. [20]. **E** - 9-chromatin states in S2 cells by Kharchenko et al. [22]. Chromatin state 1 is marked with red. **F** - 9-chromatin states in BG3 cells by Kharchenko et al. [22]. Chromatin state 1 is marked with red. **G –** Nucleosome density according to Henikoff et al. [43]. Peaks above the axis reflect high density and those below axis denote low nucleosome density. **H** – Localization of histone H1 dips in Kc cells by Braunschweig et al. [44]. Black horizontal bars indicate the genomic regions with low Histone H1 binding. **I** - DNAse I hypersensitivity sites (high magnitude DHS - vertical lines) in S2, BG3, and Kc cells by Kharchenko et al. [22]. **J** - ORC2-binding sites in S2, BG3, Kc cells and salivary glands by Eaton et al. [45], Sher et al. [46]. **K –** Enrichment profiles of NSL complex components: NSL1 binding profile from salivary glands by Raja et al. [47], NSL3 in S2 cells by Lam et al. [48], NSL1 in S2 cells by Feller et al. [49]. **L –** Enrichment regions of various proteins specific for interbands and active chromatin (fly modENCODE). The list of interband-specific proteins is taken from Demakov et al. [28] and Vatolina et al. [27].
(TIF)

**Figure S6 FISH localization of transposon insertions in the polytene chromosome interband regions 10A1-2/10A3 (A, D and G), 10A3/10A4-5 (D, E and H) and 10A7/10A8-9 (C, F and I). A - C –** overlay of FISH signal (green) and phase contrast. **D – F –** overlay of FISH signal (green), phase contrast and DAPI (blue). **G – I –** overlay of FISH signal (green) and DAPI (blue). Arrows point to the FISH signals in polytene chromosome regions.
(TIF)

**Figure S7 Box-and-whiskers diagram reflecting activity of genes located in the set of select 32 interbands (A) and intercalary heterochromatin bands (B).** To plot this diagram, data from Chintapalli et al. [55] were used. The list of tissues and organs is shown along the *X* axis. Mean mRNA expression ($\log_{10}$ scale) is shown on the *Y* axis. Median value is shown as thick black line; open boxes represent the data between 25 and 75 percentile (50% of data points). Whiskers extend to the most extreme data points which are no more than 1.5 times the

length of the box away from the box. Separate circles represent outliers.
(TIFF)

**Figure S8 Box-and whiskers diagram reflecting activity of genes located in the cyan (A) and magenta (B) chromatin types in the whole genome.** Labeling is the same as in the Figure S7.
(TIFF)

**Figure S9 Box-and-whiskers diagram showing mean expression of genes located in the 32 interband chromatin according to modENCODE RNA-seq data.** The list of datasets on temporal, tissue, treatment and cell line expression is given below the *x* axis. Y axis shows gene expression levels (RPKM in $\log_{10}$ scale) (according to [56] and S. Celniker group). Median value is shown as a thick black line; open boxes represent the data between 25 and 75 percentile (50% of data points). Whiskers extend to the most extreme data points which are no more than 1.5 times the length of the box away from the box. Separate circles represent outliers.
(TIFF)

**Figure S10 Box-and-whiskers diagram reflecting mean activity of genes located in IH bands according to modENCODE RNA-seq data.** Labeling is the same as in the **Figure S9**.
(TIFF)

**Figure S11 Box-and-whiskers diagram reflecting mean activity of genes located in the cyan type of chromatin in the whole genome according to modENCODE RNA-seq data.** Explanations as in the **Fig. S9**.
(TIFF)

**Figure S12 Box-and-whiskers diagram reflecting mean activity of genes located in magenta type of chromatin in whole genome according to modENCODE RNA-Seq data.** Explanations as in the **Fig. S9**.
(TIFF)

**Figure S13 Different types of genome organization domains described for the region 9F13 – 10B3. A -** Genomic coordinates around the 10A1-2 and 10B1-2 bands (red dashed lines). **B -** Track showing our four-state chromatin types. The span of 10A1-2 the 10A1-2 and 10B1-2 bands is indicated above the track. **C** - Map of five-state chromatin types in Kc cells by Filion et al. [20]. **D** - Position of 9 chromatin states in S2 cells by Kharchenko et al. [22]. **E** - Position of 9 chromatin states in BG3 cells by Kharchenko et a. [22]. **F** - Physical domains by Sexton et al. [67]. "Active" domain is shown in red. It indicates the transcriptionally inert IH band 10A1-2 and overlaps with theexpressed region (a series of grey bands and interbands) encompassing housekeeping genes. "Null" domain (shown in black) co-localizes with the transcriptionally silent band 10B1-2. **G** - Physical domains by Hou et al. [68]. **H** - Select genes from FlyBase which according to Weber, Hurst [69] are housekeeping. **I** - Select genes from FlyBase referenced as housekeeping in Feller et al. [49]. **J** - FlyBase genes. Genes that we classify as housekeeping are shown in red. **K** - NSL1 enrichment profile in S2 cells according to Feller et al. [49]. **L** - NSL1 enrichment in salivary glands according to Raja et al. [47]. **M** - NSL3 enrichment in S2 cells according to Lam et al. [48].
(TIF)

**Figure S14 Values of Calinski-Harabasz criterion at different numbers of states used for HMM clustering of *D. melanogaster* X chromosome sequences.**
(TIF)

**Text S1 Identification of a new set of interbands in *Drosophila melanogaster* polytene chromosomes.**
(DOC)

**Table S1 Lists of genes, whose 5′-ends map to the cytologically defined interbands and genes located completely in the intercalary heterochromatin bands.**
(XLS)

**Table S2 Coordinates and descriptions of probes selected for FISH mapping on polytene chromosomes (release 5.12).**
(DOC)

**Table S3 The list of ChIP-chip protein binding data sets used for clusterization (modENCODE Consortium, 2010).**
(XLS)

**File S1 Track of 4-state chromatin model developed in this study.**
(GZ)

**File S2 Track of head-to-head oriented genes.**
(GZ)

**File S3 Scripts in R language for building 4-state chromatin model.**
(GZ)

## Acknowledgments

The authors are grateful to Drs. L. Cherbas, P. Cherbas and J. Johansen for invaluable discussions and advice. We are also indebted to the Oxford University Press Copyright Clearance Center for granting the permission to reproduce the fragments of C. Bridges' revised maps of polytene chromosomes of *Drosophila melanogaster*.

## Author Contributions

Conceived and designed the experiments: IFZ TYZ OVD LVB VFS GVP SAD ESB. Performed the experiments: OVD VFS GVP LVB SAD. Analyzed the data: IFZ TYZ OVD FPG VAK LVB DSD VNB SAD ESB. Contributed reagents/materials/analysis tools: IFZ OVD VFS GVP LVB SAD. Contributed to the writing of the manuscript: IFZ TYZ FPG OVD LVB SAD ESB.

## References

1. Koltzoff NK (1934) The structure of the chromosomes in the salivary glands of *Drosophila*. Science 80: 312–313.
2. Bridges CB (1935) Salivary chromosome maps with a key to the banding of the chromosomes of *Drosophila melanogaster*. J Hered 26: 60–64.
3. Beermann W (1965) Structure and function of interphase chromosomes. In: Geerts SJ, ed. Genetics today. Vol. 2. XI Intern. Congress Genet. Oxford, London, Edinburgh, New York, Paris, Frankfurt: Pergamon Press. pp 375–384.
4. Beermann W (1967) Gene action at the level of the chromosome. In: Brink RA, Styles ED, eds. Heritage from Mendel. Madison, London: University of Wisconsin Press. pp 179–201.
5. Lefevre G Jr (1971) Salivary chromosome bands and the frequency of crossing over in *Drosophila melanogaster*. Genetics 67, 497–513.
6. Davidson EH, Britten RJ (1973) Organization, transcription, and regulation in the animal genome. Q Rev Biol. 48(4): 565–613.
7. Crick F (1971) General model for the chromosomes of higher organisms. Nature 234: 25–27.
8. Paul J (1972) General theory of chromosomes structure and gene activation in eukaryotes. Nature 238: 444–446.

9. Ashburner M (1972) The hormonal control of gene activity in polytene chromosomes of *Drosophila melanogaster*. International Congress Series N 273, Endocrinology, 315–318.
10. Sorsa V (1975) A hypothesis for the origin and evolution of chromomere DNA. Hereditas 81: 77–84.
11. Sorsa V (1984) Electron microscopic mapping and ultrastructure of *Drosophila* polytene chromosomes. In: King RC, Akai H, eds. Insect Ultratructure, Vol. 2. New York: Plenum Press. pp. 75–107.
12. Sokoloff S (1977) Replication of polytene chromosomes in the dipteran Phryne. Disser. Grad. Doktors Naturwiss. Univers. Tubingen.
13. Kosswig C, Sengun A (1947) Intraindividual variability of chromosome IV of Chironomus. J Hered 38: 235–239.
14. Fujita S (1965) Chromosomal organization as a genetic basis of cytodifferentiation in multicellular organisms. Nature 206(985): 742–744.
15. Gersh ES (1975) Sites of gene activity and of inactive genes in polytene chromosomes of diptera. J Theor Biol 50(2): 413–28.
16. Speiser Ch (1974) Eine Hypothese uber die functionelle Organization der Chromosomen hoherer Organismen. Theoret Appl Genet 44: 97–99.
17. Zhimulev IF, Belyaeva ES (1975) Proposals to the problem of structural and functional organization of polytene chromosomes. Theor. Appl. Genet. 45, 335–340.
18. Zhimulev IF (1999) Genetic organization of polytene chromosomes. Adv Genet 39: 1–589.
19. Zhimulev IF, Belyaeva ES, Vatolina TY, Demakov SA (2012) Banding patterns in *Drosophila melanogaster* polytene chromosomes correlate with DNA-binding protein occupancy. Bioessays 34(6): 498–508.
20. Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J (2010) Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. Cell 143(2): 212–24.
21. White R (2012) Packaging the fly genome: domains and dynamics. Brief Funct Genomics 11(5): 347–55.
22. Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, et al. (2011) Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. Nature 471(7339): 480–5.
23. Pindyurin AV, Moorman C, de Wit E, Belyakin SN, Belyaeva ES, et al. (2007) SUUR joins separate subsets of PcG, HP1 and B-type lamin targets in *Drosophila*. J Cell Sci 120(Pt 14): 2344–51.
24. Belyakin SN, Christophides GK, Alekseyenko AA, Kriventseva EV, Belyaeva ES, et al. (2005) Genomic analysis of *Drosophila* chromosome underreplication reveals a link between replication control and transcriptional territories. Proc Natl Acad Sci U S A 102(23): 8269–74.
25. Belyakin SN, Babenko VN, Maksimov DA, Shloma VV, Kvon EZ, et al. (2010) Gene density profile reveals the marking of late replicated domains in the *Drosophila melanogaster* genome. Chromosoma 119(6): 589–600.
26. Belyaeva ES, Goncharov FP, Demakova OV, Kolesnikova TD, Boldyreva LV, et al. (2012) Late replication domains in polytene and non-polytene cells of *Drosophila melanogaster*. PLoS One 7(1):e30035.
27. Vatolina TYu, Boldyreva LV, Demakova OV, Demakov SA, Kokoza EB, et al. (2011) Identical functional organization of cell line interphase and polytene chromosomes in *Drosophila melanogaster*. PloS ONE 6(10), e25960.
28. Demakov SA, Vatolina TY, Babenko VN, Semeshin VF, Belyaeva ES, et al. (2011) Protein composition of interband regions in polytene and cell line chromosomes of *Drosophila melanogaster*. BMC Genomics 12: 566.
29. Zhimulev IF, Belyaeva ES (1975) 3H-uridine labelling patterns in the *Drosophila melanogaster* salivary gland chromosomes X, 2R and 3L. Chromosoma 49: 219–231.
30. Sorsa V, Sorsa M (1968) Ideas on the linear organization of chromosomes revived by electron microscopic studies of stretched salivary chromosomes. Ann Acad Sci Fenn Ser A IV Biol 134: 1–16.
31. Semeshin VF, Zhimulev IF, Belyaeva ES (1979) Electron microscope autoradiographic study on transcriptional activity of *Drosophila melanogaster* polytene chromosomes. Chromosoma 73: 163–177.
32. Vatolina TYu, Demakov SA, Semeshin VF, Makunin IV, Babenko VN, et al. (2011) Identification and molecular genetic characterization of the polytene chromosome interbands in *Drosophila melanogaster*. Rus J Genet 47(5): 521–532.
33. Kozlova TYu, Semeshin VF, Tretyakova IV, Kokoza EB, Pirrotta V, et al. (1994) Molecular and cytogenetical characterization of the 10A1-2 band and adjoining region in the *Drosophila melanogaster* polytene X chromosome. Genetics 136: 1063–1073.
34. Zhimulev IF, Belyaeva ES, Makunin IV, Pirrotta V, Volkova EI, et al. (2003) Influence of the SuUR gene on intercalary heterochromatin in Drosophila melanogaster polytene chromosomes. Chromosoma 111: 377–398.
35. Bridges CB (1938) A revised map of the salivary gland X-chromosome of *Drosophila melanogaster*. J Hered 29: 11–13.
36. Bridges PN (1941) A revised map of the left limb of the third chromosome of *Drosophila melanogaster*. J Hered 32: 64–65.
37. Bridges PN (1942) A new map of the salivary gland 2L-chromosome of *Drosophila melanogaster*. J Hered 33: 403–408.
38. Bridges CB, Bridges PN (1939) A new map of the second chromosome: a revised map of the right limb of the second chromosome of *Drosophila melanogaster*. J Hered 30: 475–476.
39. Bridges PN (1941) A revision of the salivary gland 3R chromosome map of *Drosophila melanogaster*. J Hered 32: 299–300.
40. Mojena R (1977) Hierarchical grouping methods and stopping rules: an evaluation. The Computer Journal 20: 359–363.
41. Gortchakov AA, Eggert H, Gan M, Mattow J, Zhimulev IF, et al. (2005) Chriz, a chromodomain protein specific for the interbands of *Drosophila melanogaster* polytene chromosomes. Chromosoma 114(1): 54–66.
42. Hoskins RA, Landolin JM, Brown JB, Sandler JE, Takahashi H, et al. (2011) Genome-wide analysis of promoter architecture in Drosophila melanogaster. Genome Res 21(2):182–92.
43. Henikoff S, Henikoff JG, Sakai A, Loeb GB, Ahmad K (2009) Genome-wide profiling of salt fractions maps physical properties of chromatin. Genome Res 19(3):460–9.
44. Braunschweig U, Hogan GJ, Pagie L, van Steensel B (2009) Histone H1 binding is inhibited by histone variant H3.3. Embo J 28: 3635–3645.
45. Eaton ML, Prinz JA, MacAlpine HK, Tretyakov G, Kharchenko PV, et al. (2011) Chromatin signatures of the *Drosophila* replication program. Genome Res 21(2):164–74.
46. Sher N, Bell GW, Li S, Nordman J, Eng T, et al. (2012) Developmental control of gene copy number by repression of replication initiation and fork progression. Genome Res 22(1):64–75.
47. Raja SJ, Charapitsa I, Conrad T, Vaquerizas JM, Gebhardt P, et al. (2010) The nonspecific lethal complex is a transcriptional regulator in *Drosophila*. Mol Cell 38(6):827–41.
48. Lam KC, Mühlpfordt F, Vaquerizas JM, Raja SJ, Holz H, et al.(2012) The NSL complex regulates housekeeping genes in *Drosophila*. PLoS Genet 8(6):e1002736.
49. Feller C, Prestel M, Hartmann H, Straub T, Söding J (2011) The MOF-containing NSL complex associates globally with housekeeping genes, but activates only a defined subset. Nucleic Acids Res 40(4): 1509–22.
50. Petesch SJ, Lis JT (2012) Overcoming the nucleosome barrier during transcript elongation. Trends Genet 28(6): 285–94.
51. Rath U, Wang D, Ding Y, Xu YZ, Qi H, et al. (2004) Chromator, a novel and essential chromodomain protein interacts directly with the putative spindle matrix protein skeletor. J Cell Biochem 93(5): 1033–47.
52. Eggert H, Gortchakov A, Saumweber H (2004) Identification of the *Drosophila* interband-specific protein Z4 as a DNA binding zinc-finger protein determining chromosomal structure. J Cell Sci 117: 4253–4264.
53. Zhang P, Lee H, Brunzelle JS, Couture JF (2012) The plasticity of WDR5 peptide-binding cleft enables the binding of the SET1 family of histone methyltransferases. Nucleic Acids Research 1–10.
54. Spradling AC, Bellen HJ, Hoskins RA (2011) *Drosophila* P elements preferentially transpose to replication origins. Proc Natl Acad Sci U S A 108(38): 15948–53.
55. Chintapalli VR, Wang J, Dow JA (2007) Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. Nat Genet 39(6): 715–20.
56. Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, et al. (2011) The developmental transcriptome of *Drosophila melanogaster*. Nature 471(7339): 473–479.
57. Wakano C, Byun JS, Di LJ, Gardner K (2012) The dual lives of bidirectional promoters. Biochim Biophys Acta. 2012 1819(7):688–93.
58. Babenko VN, Matvienko VF (2013) P elements comprising mini-white marker gene suppression features in intergenic regions of D. melanogaster genome. Tsitologiia. 2013;55(3):181–4.
59. Kaufmann B (1939) Distribution of induced breaks along the X-chromosome of *Drosophila melanogaster*. Proc Natl Acad Sci U S A 25: 571–577.
60. Belyaeva ES, Andreyeva EN, Belyakin SN, Volkova EI, Zhimulev IF (2008) Intercalary heterochromatin in polytene chromosomes of Drosophila melanogaster. Chromosoma 117: 411–418.
61. Makunin IV, Volkova EI, Belyaeva ES, Nabirochkina EN, Pirrotta V, et al. (2002) The *Drosophila* Suppressor of Underreplication protein binds to latereplicating regions of polytene chromosomes. Genetics 160: 1023–1034.
62. Nordman J, Li S, Eng T, Macalpine D, Orr-Weaver TL (2011) Developmental control of the DNA replication and transcription programs. Science 21(2):175–81
63. Andreyenkova NG, Kolesnikova TD, Makunin IV, Pokholkova GV, Boldyreva LV, et al. (2013) Late replication domains are evolutionary stable in Drosophila genome PloS ONE under review.
64. Yao C, Ding Y, Cai W, Wang C, Girton J, et al. (2012) The chromodomain-containing NH2-terminus of Chromator interacts with histone H1 and is required for correct targeting to chromatin. Chromosoma 121(2):209–20.
65. Ding Y, Yao C, Lince-Faria M, Rath U, Cai W, et al. (2009) Chromator is required for proper microtubule spindle formation and mitosis in *Drosophila*. Dev Biol. 334(1):253–63.
66. Miele V, Vaillant C, d'Aubenton-Carafa Y, Thermes C, Grange T. (2008) DNA physical properties determine nucleosome occupancy from yeast to fly. Nucleic Acids Res. 36(11):3746–56.
67. Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, et al. (2012) Three-dimensional folding and functional organization principles of the *Drosophila* genome. Cell 148(3): 458–72.
68. Hou C, Li L, Qin ZS, Corces VG (2012) Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. Mol Cell 48(3): 471–84.
69. Weber CC, Hurst LD (2011) Support for multiple classes of local expression clusters in *Drosophila melanogaster*, but no evidence for gene order conservation. Genome Biol 12(3): R23.

70. Semeshin VF, Belyaeva ES, Shloma VV, Zhimulev IF (2004) Electron microscopy of polytene chromosomes. Methods Mol Biol 247: 305–24.
71. Ashburner M, Golic KG, Hawley RS (2005) *Drosophila*: A Laboratory Handbook. 2nd Edition. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
72. Kolesnikova TD, Semeshin VF, Andreyeva EN, Zykov IA, Kokoza EB, et al. (2011) Induced decondensation of heterochromatin in *Drosophila melanogaster* polytene chromosomes under condition of ectopic expression of the *Supressor of underreplication* gene. Fly (Austin) 5(3): 181–190.
73. R Development Core Team (2011) R: A Language and Environment for Statistical Computing. Vienna, Austria. Available: http://www.R-project.org.
74. Taramasco O, Bauer S (2013) RHmm: Hidden Markov Models simulations and estimations. Available: http://CRAN.R-project.org/package = RHmm.
75. Caliński T, Harabasz J (1974) A dendrite method for cluster analysis. Communications in Statistics 3: 1–27. doi:10.1080/03610927408827101.
76. Milligan GW, Cooper MC (1985) An examination of procedures for determining the number of clusters in a data set. Psychometrika 50: 159–179. doi:10.1007/BF02294245.

# Forty years of the 93D puff of Drosophila melanogaster

**Article** *in* Journal of Biosciences · August 2011

DOI: 10.1007/s12038-011-9078-1 · Source: PubMed



**1 author:**

Subhash Chandra Lakhotia
Banaras Hindu University
**273** PUBLICATIONS **2,738** CITATIONS

SEE PROFILE

# Forty years of the 93D puff of *Drosophila melanogaster*

SUBHASH C LAKHOTIA

*Cytogenetics Laboratory, Department of Zoology, Banaras Hindu University, Varanasi 221 005, India*

*(Fax, +91-542-2368457; Email, lakhotia@bhu.ac.in)*

The 93D puff of *Drosophila melanogaster* became attractive in 1970 because of its singular inducibility by benzamide and has since then remained a major point of focus in my laboratory. Studies on this locus in my and several other laboratories during the past four decades have revealed that (i) this locus is developmentally active, (ii) it is a member of the heat shock gene family but selectively inducible by amides, (iii) the *93D* or *heat shock RNA omega* (*hsrω*) gene produces multiple nuclear and cytoplasmic large non-coding RNAs (hsrω-n, hsrω-pre-c and hsrω-c), (iv) a variety of RNA-processing proteins, especially the hnRNPs, associate with its >10 kb nuclear (hsrω-n) transcript to form the nucleoplasmic omega speckles, (v) its genomic architecture and hnRNP-binding properties with the nuclear transcript are conserved in different species although the primary base sequence has diverged rapidly, (vi) heat shock causes the omega speckles to disappear and all the omega speckle associated proteins and the hsrω-n transcript to accumulate at the 93D locus, (vii) the hsrω-n transcript directly or indirectly affects the localization/stability/activity of a variety of proteins including hnRNPs, Sxl, Hsp83, CBP, DIAP1, JNK-signalling members, proteasome constituents, lamin C, ISWI, HP1 and poly(ADP)-ribose polymerase and (viii) a balanced level of its transcripts is essential for the orderly relocation of various proteins, including hnRNPs, RNA pol II and HP1, to developmentally active chromosome regions during recovery from heat stress. In view of such multitudes of interactions, it appears that large non-coding RNAs like those produced by the *hsrω* gene may function as hubs to coordinate multiple cellular networks and thus play important roles in maintenance of cellular homeostasis.

[Lakhotia SC 2011 Forty years of the 93D puff of *Drosophila melanogaster*. *J. Biosci.* **36** 399–423] **DOI** 10.1007/s12038-011-9078-1

## 1. Genesis of the 93D puff

The 93D puff has occupied a major part of my research career. I became fascinated by the 93D puff sometimes in 1969 and this fascination has increased since then. A personalized narrative of how the story of 93D puff developed during the past 40 years is presented on the occasion of completion of 65 years of my life.

Following the completion of a master's degree in zoology and comparative anatomy with specialization in parasitology, I entered the newly established genetics laboratory of Dr AS Mukherjee in 1967 for doctoral studies. Dr Mukherjee, a young and remarkably stimulating teacher, joined the Department of Zoology at the University of Calcutta (now University of Kolkata) in 1965 after his PhD under Prof Curt Stern (University of California Berkeley, USA) and post-doctoral work with Prof W Beermann (Max-Planck Institute, Tubingen, Germany). I opted to undertake research on dosage compensation of the X-linked genes of *Drosophila*. Mukherjee and Beermann (1965) had suggested that dosage compensation operates through hyperactivation of the X-linked genes in somatic cells of males such that the transcriptional activity of the single X-chromosome in male equaled that of the two X-chromosomes in female somatic cells. The dual expertise of Dr Mukherjee in *Drosophila* genetics and in polytene chromosome cytology was a great starting point for me. My studies, besides reconfirming the hyperactivity (increased rate of transcription and faster completion of a polytene replication cycle of the X-chromosome in salivary glands of male larvae; *see* Lakhotia and Mukherjee 1969, 1970a, b), also showed that the hyperactive state of the single X-chromosome in males was cell autonomous (Lakhotia and Mukherjee 1969). Further, in agreement with the proposal of Muller and Kaplan (1966) that, unlike the whole X-chromosome inactivation in female mammals (Lyon 1961),

**Keywords.** Heat shock; hnRNP; *hsrω*; non-coding RNA, omega speckles

the different X-linked genes in *Drosophila* were individually regulated for dosage compensation ('piecemeal' regulation), it was shown that the hyperactive state of the male X-chromosome did not spread to an inserted autosomal segment (Lakhotia 1970).

The evidence for the hyperactive state of the male X-chromosome was based entirely on its morphology, its faster completion of a replication cycle in polytene cells and relatively greater incorporation of $^{3}$H-uridine along its length compared with that on the autosomes in the larval salivary gland polytene chromosomes of *Drosophila melanogaster*. We reasoned that demonstration of a greater sensitivity of the rate of RNA synthesis on the male X-chromosome to transcriptional inhibition would more firmly establish the hyperactive-X model of Mukherjee and Beermann (1965). In the 1960s, actinomycin D was the commonly used inhibitor of eukaryotic transcription. However, its primary target was nucleolar (ribosomal) rather than chromosomal transcription, which was inhibited only at higher dosages of the inhibitor (Perry and Kelley 1968). α-Amanitin was yet to be identified as a potent inhibitor of chromosomal or RNA-pol-II-dependent transcription (Lindell *et al.* 1970). Since actinomycin would inhibit chromosomal transcription only at a higher concentration, I avoided its use for testing the predicted greater sensitivity of the male X-chromosome transcription. A search of the literature revealed that some benzimidazole derivatives and benzamide (BM) were already used (Sirlin and Jacob 1964) to specifically inhibit chromosomal transcription, without much effect on the nucleolar transcription, in polytene nuclei of a *Smittia*, a member of the Chironomidae family. Therefore, these inhibitors appeared attractive for our experiments. Because of the high cost of benzimidazole derivatives, I decided to try BM to see if it primarily blocked chromosomal transcription in *Drosophila* polytene nuclei as well, and if the X-chromosomal transcription in male nuclei was indeed more sensitive to such inhibition. A brief treatment of salivary glands with 1.3 mg/ml (~10 mM) concentration of BM did inhibit uptake of $^{3}$H-uridine at chromosomal sites without affecting the nucleolar incorporation and, as expected on the hyperactive-male-X model of dosage compensation (Mukherjee and Beermann 1965), it did reveal the X-chromosomal transcription in male cells to be more sensitive to the inhibition (Lakhotia 1970). However, a completely unexpected finding was that in the BM-treated polytene nuclei, one of the chromosome sites was highly puffed and showed several-fold higher incorporation of $^{3}$H-uridine amidst the otherwise poorly labelled polytene chromosome arms (figure 1). This cytological region, visible as a large puff with very high incorporation of $^{3}$H-uridine, was identified on the maps of polytene chromosomes prepared by Bridges as the 93D cytogenetic region on the right arm of chromosome 3. This serendipitous observation was the starting point for my fascination of the 93D puff (Lakhotia and Mukherjee 1970b).

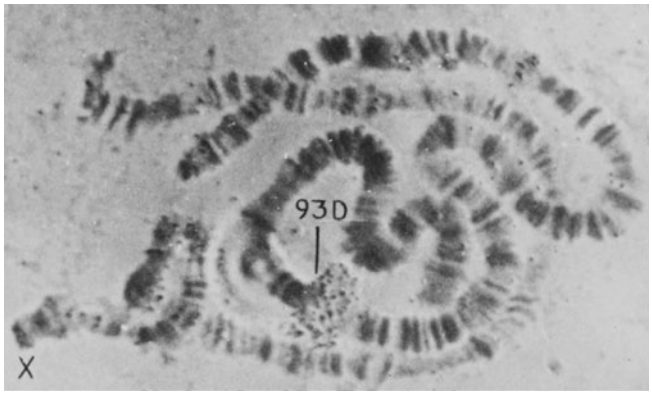



**Figure 1.** Specific induction of 93D puff by 10 min treatment of salivary glands of *D. melanogaster* larva (male) with 1.3 mg/ml benzamide at 24°C. While most of the chromosomal incorporation of $^{3}$H-uridine, visualized as small black granules in the autoradiogram, is blocked, that at the 93D site, which also forms a large puff, is enhanced. Image reproduced from Lakhotia (1970).

Classical studies in the 1960s utilizing polytene chromosomes provided much of the direct evidence for correlation of chromatin organization, gene activity and transcription. An elegant study (Edstrom and Beermann 1962) using microdissected Balbiani rings and chromosomes to analyse the base composition of DNA and RNA isolated from specific regions provided evidence that one of the two DNA strands of a given Balbiani ring produces its distinct transcript. This and other evidences then available established that puffs are sites of transcription and, therefore, of gene activity (reviewed by Berendes and Beermann 1969). However, a similar direct correlation between a given puff and the expected protein product was still not available in the 1960s, except for indirect evidence from studies like the observed specific difference in puffing of a given chromosome region and its correlation with the presence or absence of a secretory protein in salivary glands of two related species of *Chironomus* (Grossbach 1969). Against this background, the singular induction of the 93D puff in larval salivary glands of *D. melanogaster* following brief treatment with BM appeared to provide a unique and exciting opportunity to examine the correlation between a puff, its RNA transcript and the final protein product. Consequently, I decided to investigate the 93D puff further as and when I had the opportunity to take up my independent research programme.

By the end of my doctoral research, besides the desire to study the 93D puff in greater detail, I had interests in the processes of replication and transcription at chromosomal level and in heterochromatin. I did not, however, initiate work on the 93D puff for a few years after completing my doctorate in 1970. After the fly experience, I decided to gain some

experience in mammalian cytogenetics and joined Dr SRV Rao's laboratory at the Department of Zoology, Delhi University, as a post-doctoral fellow, where I participated in the description of karyotypes of a few as yet chromosomally uncharacterized rodent species. I took up a faculty position at the Department of Zoology of Burdwan University (Burdwan) in 1971 but moved to Gujarat University (Ahmedabad) in June 1972. As a legacy of my doctoral research, I was interested in knowing if replication and transcription could occur simultaneously at a given chromosome region and had imagined that electron microscope (EM) autoradiographic studies on polytene chromosomes doubly labelled with $^{14}$C-thymidine and $^{3}$H-uridine would be ideal to address this question. While at Delhi in 1970–1971, I met Dr J Jacob of the Institute of Animal Genetics, Edinburgh, who had considerable experience in EM autoradiography and who also happened to be one of the authors of the paper (Sirlin and Jacob 1964) that introduced me to BM. During one of our discussions, he indicated that if I could get a fellowship to come to Edinburgh, I could use his EM facility to examine the relationship between transcription and replication. Keeping this in mind, I went to the Institute of Animal Genetics (Edinburgh) in November 1972, having been selected for the Overseas Scholarship of the Royal Commission for the Exhibition of 1851 (UK). Although the experimental design for examining simultaneous transcription and replication at a given chromosome region at the EM level did not work out, during the 11 months' stay at Edinburgh, I obtained strong evidence for (i) active transcription in the classical chromocentric heterochromatin in polytene nuclei of *D. melanogaster* (Lakhotia and Jacob 1974b), which contradicted the then 'established fact' of heterochromatin being completely inert, (ii) absence of any replication of the α-heterochromatin during polytenization (Lakhotia 1974) and (iii) initiation of replication in chromatin regions that are not associated with the nuclear envelope (Lakhotia and Jacob 1974a). Based on these studies, my interest in heterochromatin and chromosomal replication in *Drosophila* also became more firmly rooted. After returning to Gujarat University, I decided to start studying the 93D puff, heterochromatin and replication in *Drosophila* chromosomes. While active research in the 93D puff has continued till date, I had to give up research relating to heterochromatin and chromosomal replication after several years of active involvement and many research papers, because of my increasing interest in the 93D puff and, emanating from it, the regulation and functions of other heat shock genes.

## 2. The 93D puff is a member of heat shock gene family but behaves differently from the other heat shock puffs

I initiated research on the 93D puff in 1974 at Gujarat University and one of the experiments, which appeared straightforward, was to see if following the BM treatment, a new protein was synthesized in the larval salivary glands, which could be correlated with the BM-induced enhanced transcription at the 93D puff. However, this rather simple experiment proved to be quite frustrating for several years because of the extremely limited laboratory and financial resources. During this period, two seminal papers (Tissieres *et al.* 1974; Lewis *et al.* 1975) on heat-shock-induced synthesis of new proteins in *Drosophila* tissues were published. The 93D puff was already identified to be one of the major heat-shock-induced puffs in salivary glands of *D. melanogaster* (Ashburner 1970). Thus, our task appeared narrowed down to identifying which of the heat-shock-induced proteins was also induced by BM treatment, presuming *prima facie* that the same gene was induced by the two treatments, which was subsequently indeed proven to be a correct surmise (sections 3 and 4). However, our protein synthesis studies were not making any headway because of the inability to get a vertical slab gel electrophoresis apparatus and a gel drier. We tried various 'homemade' alternatives for running thin polyacrylamide slab gels without success.

I moved to the Department of Zoology at Banaras Hindu University in September 1976 and, after settling down, continued with my research interests in the 93D puff, heterochromatin and chromosomal replication in *Drosophila.* Tapas Mukherjee became my PhD student in 1977, the first to work on the 93D puff. By this time, I had two research grants but still not enough funding to buy a vertical slab gel electrophoresis apparatus and gel drier. Because the protein synthesis study was still not feasible, Tapas started examining the heat-shock- and BM-induced puffing.

An early interesting observation was that the relative level of heat-shock-induced transcriptional activity of the 93D puff varied independently of the other major heat shock puffs, which were induced in concert like a 'battery' (Mukherjee and Lakhotia 1979). It was reported by Compton and McCarthy (1978) that incubation of isolated polytene nuclei of *D. melanogaster* in cytoplasm from heat-shocked Kc cells resulted in only a weak induction of the 93D puff while the other heat shock puffs were strongly induced as after heat shock to intact salivary glands. Following this report, we (Mukherjee and Lakhotia 1981) found that incubation of intact salivary glands, rather than isolated polytene nuclei, in a homogenate of heat-shocked salivary glands resulted in specific activation of only the 93D puff. These observations indicated that the 93D puff was regulated differently from the other heat shock puffs, a surmise that found support in subsequent reports that the binding of the heat shock factor to the 93D site following a typical heat shock was very transient compared to that on the other heat shock induced puff sites (Westwood *et al.* 1991). We also found

that although BM and heat shock both induced the 93D puff, a combination of the two inducers (one following the other or both applied together) results in inhibition of puffing and transcriptional activity of the 93D puff (Lakhotia and Mukherjee 1980). Likewise, recovery from anoxia at 24°C results in strong induction of 93D and other heat shock puffs; the same at 37°C fails to induce the 93D although the other heat shock puffs are activated (Mukherjee and Lakhotia 1982). Another intriguing observation was that the non-induction of the 93D puff following heat shock in combination with BM or recovery from anoxia at 37°C is accompanied by differential activation (Lakhotia and Mukherjee 1980; Mukherjee and Lakhotia 1982) of the otherwise nearly equally induced 87A and 87C puffs, which were by then known to be duplicated loci encoding Hsp70 (Livak *et al.* 1978). The observation relating to the unequal puffing of the 87A and 87C puffs when the 93D puff is not induced by heat shock has been a recurring theme in later studies (see section 5), and this aroused my interest in the other heat shock genes and thus in stress biology.

Some studies from Mary-Lou Pardue's laboratory in late 1970s had also suggested unusual properties of the 93D puff. In a remarkable study, Spradling *et al.* (1977) used electrophoretically separated fractions of radioactively labelled newly synthesized RNA from heat-shocked cells of *Drosophila* to hybridize *in situ* with polytene chromosomes (for the present day molecular or genomic biologists, the approach used in this study would be analogous to identification of *newly synthesized RNA* species using a whole genome microarray analysis). They found that none of the newly synthesized polyA+ RNA fractions from heat-shocked cells hybridized to the 93D puff. On the other hand, Bonner and Pardue (1976) had earlier shown that, unlike the other heat shock sites, the 93D puff was activated to very different levels depending upon the conditions of heat shock. In yet another study, Lengyel *et al.* (1980) examined metabolism of RNA at the 93D puff in heat-shocked cells and showed that this RNA may contain repeated sequences and only a part of the sequences transcribed from 93D were exported from the nucleus to cytoplasm, while a greater part stayed in the nucleus itself.

I first met Mary-Lou Pardue in the late 1970s during a meeting at Tata Institute of Fundamental Research, Mumbai. Following this meeting, we started regular and long-lasting exchange of our results and ideas. She visited my laboratory at Varanasi in the 1980s several times, not only because of our common interest in the 93D puff but also because Varanasi was on her way to Kathmandu for trekking in the Himalayas. I also visited her laboratory in 1985 for 3 months. These exchanges resulted in two joint publications in 1989 (section 7).

## 3. A homolog of the 93D puff is present in different species of *Drosophila* but does not seem to encode a protein

It was clear from the publications in the late 1970s that the 93D locus differed from the other heat shock loci in features that warranted additional study, especially when the heat shock response was becoming 'hotter' for molecular biological studies on gene expression and regulation. Given this background, it was important for us to identify the protein product that was expected to be encoded by the 93D gene. In the absence of a commercial vertical slab gel apparatus and a gel drier, Tapas decided to fabricate these apparatuses himself and, remarkably, he succeeded. However, when we finally performed this experiment with the 'home-made' gel system, we found no detectable difference in protein synthesis in BM-treated and control salivary glands and we suggested that the heat shock and BM-inducible 93D puff may not encode a protein (Lakhotia and Mukherjee 1982; figure 2). The above-noted findings of Lengyel *et al.* (1980) on metabolism of this puff's transcripts also seemed to indicate such a possibility.

The apparent non-coding nature of this gene was generally disappointing in view of the emerging concept of 'selfish' or 'junk' DNA (Doolittle and Sapienza 1980; Orgel and Crick 1980). However, our interest in this gene grew because we had just found that a BM- and heat-shock-inducible puff was present in every species of *Drosophila* that we examined (Lakhotia and Singh 1982; figure 3). Having a strong faith in the power of natural selection, I continued to believe in functional requirement of the 93D puff and its homologs in *Drosophila*, even if these loci were not coding for a protein.

Identification of the 2-48B (also sometimes named as 2-48C) puff of *Drosophila hydei* as a homolog of the 93D puff of *D. melanogaster* (Lakhotia and Singh 1982) was especially interesting since the 2-48B puff was shown by the group at Nijmegen (the Netherlands), led by Dr HD Berendes, to be singularly induced by vit-B6 or pyridoxine (Leenders *et al.* 1973). While on a brief visit to the Nijmegen laboratory in 1973, I asked Dr HJ Leenders if his group had ever examined inducibility of the 2-48B puff of *D. hydei* with BM. Dr Leenders offered to check this inducibility and, shortly thereafter, informed me in the negative. However, it soon became apparent that he tested benzidine rather than BM; because the Nijmegen laboratory did not have BM, the issue was not followed further. Only 9 years later we identified the *D. hydei* 2-48B puff as BM inducible (Lakhotia and Singh 1982). Interestingly, the Nijmegen group had attempted to find homology of different heat shock puffs in *Drosophila* species through *in situ* hybridization of transcripts

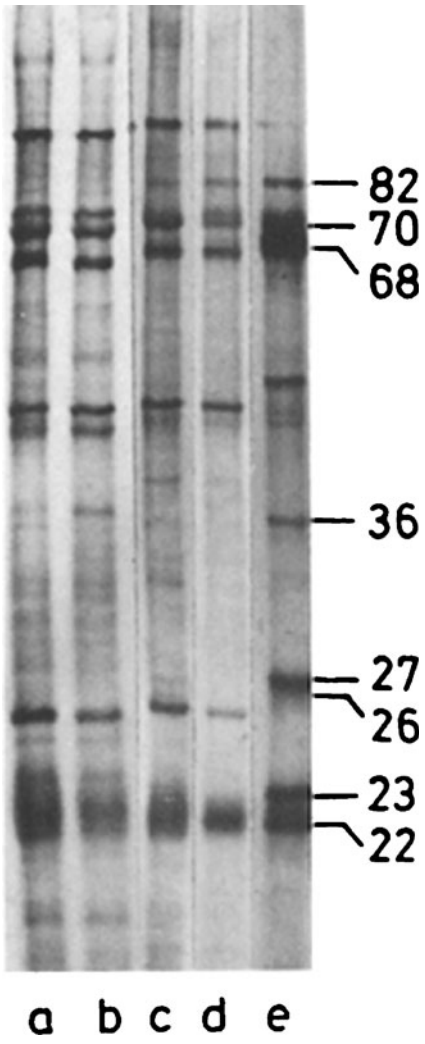



**Figure 2.** Specific induction of the 93D puff by BM is not accompanied by synthesis of a new polypeptide. Autoradiogram of $^{35}$S-methionine-labelled polypeptides separated by SDS-polyacrylamide gel electrophoresis reveals similar patterns of protein synthesis in control (lanes a and c) and in corresponding sister glands treated with benzamide at 24°C for 20 min (lanes b and d). Lane e shows protein synthesis pattern in heat-shocked glands. Note the absence of labelling of any novel polypeptide in glands treated with BM (lanes b and d). Image reproduced from Lakhotia and Mukherjee (1982)

induced in response to heat shock and vit-B6. Peters *et al.* (1980) reported that the nuclear transcripts from the 2-48B puff of *D. hydei* did not hybridize with any of the chromosome sites in *D. melanogaster* or in *D. virilis* and concluded that while the 20CD puff of *D. virilis* was an homolog of the 2-48B puff because of their common vit-B6 inducibility, a homolog of 2-48B did not exist in *D. melanogaster* because vit-B6 did not induce any puff in this species. Although the inference of Peters *et al.* (1980) that a 2-48B homolog was absent in *D. melanogaster* turned out to incorrect, their study did portend the subsequent discovery (section 5) of the rather rapid DNA base sequence divergence of the 93D and its homologs in other species. A later study in my lab (Singh and Lakhotia 1983) found that while the homogenate of heat shocked salivary glands of *D. melanogaster* retained its ability to specifically induce the 93D puff even after being heated to 100°C, that of *D. hydei* lost its 2-48B inducing ability when heated. The reason for the 93D puff of *D. melanogaster* remaining refractory to vit-B6 treatment as well as the nature and identity of 93D-inducing factors in the homogenate of heat-shocked cells remain unknown.

In view of a correlated increase in synthesis of a 40 kDa tyrosine aminotransferase following heat-shock- or pyridoxine-induced activation of the 2-48B puff of *D. hydei*, Belew and Brady (1981; also see Brady and Belew 1981) suggested that the cytoplasmic transcript of this gene (Peters *et al.* 1980) may encode the tyrosine aminotransferase. However, this has not been confirmed subsequently. Likewise, Scalenghe and Ritossa (1977) suggested that the 93D puff of *D. melanogaster* may encode subunit I of glutamine synthetase, but this paper was subsequently withdrawn.

With a view to identify a 93D-like locus in other dipterans, we examined heat-shock- and BM-inducible puffs in *Anopheles* and *Chironomus* (Nath and Lakhotia 1991). Although a BM- or colchicine-inducible puff was not detected in these two species, the nature of transcripts, base sequence evolution and binding of hnRNPs and Hsp90 at one of the heat-shock-inducible telomeric balbiani ring TBR-III in *Chironomus thummi* (Santa-Cruz *et al.* 1984; Carmona *et al.* 1985, Lakhotia 1989; Nath and Lakhotia 1991; Morcillo *et al.* 1994) suggested that this locus could be functionally similar to the 93D puff.

## 4. Cloning and sequencing of the 93D puff and its homologs in other species confirm its non-coding nature as well as rapid sequence divergence

Cloning of some sequences from the 2-48B puff of *D. hydei* (Peters *et al.* 1980) indicated certain unusual features of this puff, which were further confirmed by a more detailed characterization of several clones from this region (Peters *et al.* 1984). In the same year, cloning of the 93D puff sequences using micro-dissection of polytene chromosomes of *D. melanogaster* was reported (Walldorf *et al.* 1984). These studies also indicated that the nuclear transcripts of these puffs were not translated since they appeared to be mostly made up of repeated sequences with

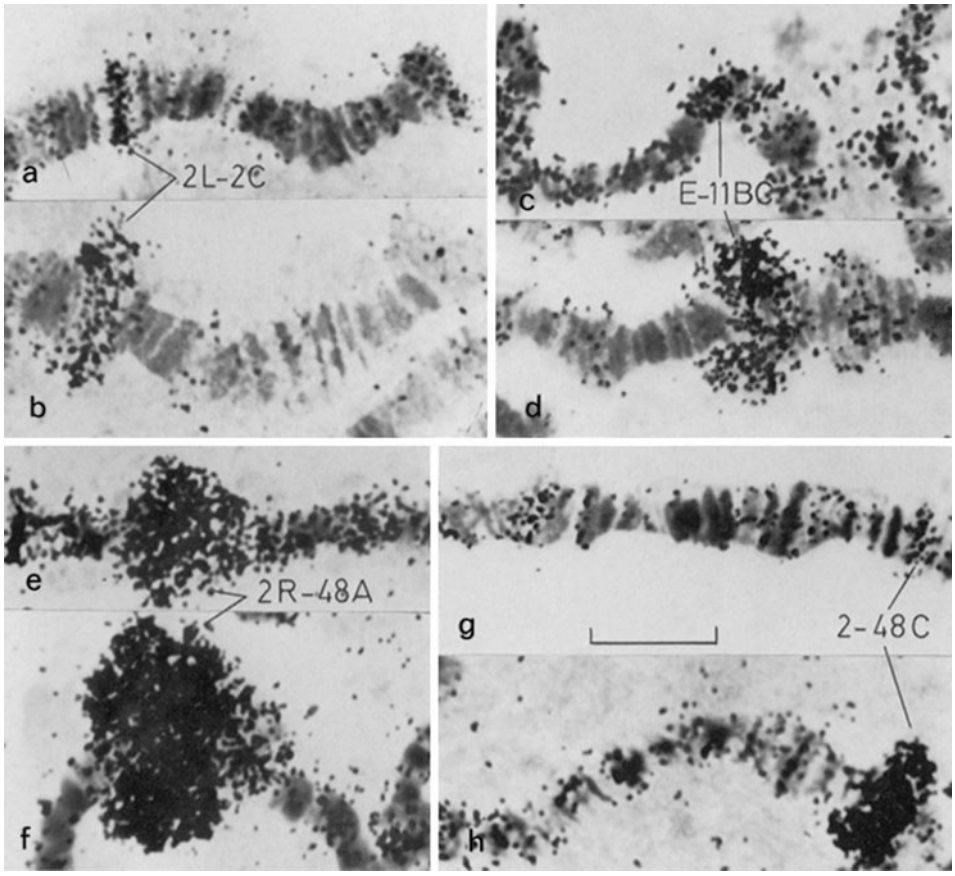



**Figure 3.** 93D homologs in other species of *Drosophila* identified through $^{3}$H-uridine labelling of the benzamide-inducible heat shock puff site: control (**a**, **c**, **e** and **g**) and BM-treated (**b**, **d**, **f** and **h**) salivary glands of (**a**–**b**) *D. ananassae,* (**c**–**d**) *D. kikkawai,* (**e**–**f**) *D. nasuta* and (**g**–**h**) *D. hydei*. Reproduced from Lakhotia and Singh (1982).

very limited coding potential. Subsequent cloning and sequencing studies in the laboratories of Pardue (Garbe *et al.* 1986; Garbe and Pardue 1986; Fini *et al.* 1989) and B Hovemann (Ryseck *et al.* 1985, 1987; Hovemann *et al.* 1986) provided data for comparison of the base sequences of the 93D and its homologs in *D. hydei* and *D. pseudoobscura*; a 93D-like puff was identified by Burma and Lakhotia (1984) in the latter species also. These data not only confirmed the earlier predictions (Lengyel *et al.* 1980; Peters *et al.* 1980; Lakhotia and Mukherjee 1982) that this locus may not encode a protein but also revealed production of multiple transcripts by this locus and its unexpected rapid sequence divergence in spite of the rather conserved architecture in *Drosophila* species (reviewed in Lakhotia 1987; Bendena *et al.* 1989a; Pardue *et al.* 1990). On the basis of comparison of the base sequence data for three species, *D. melanogaster*, *D. hydei* and *D. pseudoobscura*, the following unconventional features of the organization and transcripts of this locus emerged (figure 4). It may be mentioned that a bioinformatic analysis (Eshita Mutt and Lakhotia, unpublished) of the genomic sequences of different *Drosophila* species available on the Flybase (*http://flybase.org*) has confirmed the below noted architectural features of the 93D-like loci in all the species:

a. The 93D locus and its homologs in all other species have ~10- to 15-kb-long transcription unit with the proximal part (~2.6 kb) comprising of unique sequence and distal region (>5 kb) comprising entirely of tandem repeats of a short sequence, unique to this locus.
b. Two primary and independent transcripts, both limited to nucleus, are produced by this locus: the longer transcript (usually >10 kb) represents the

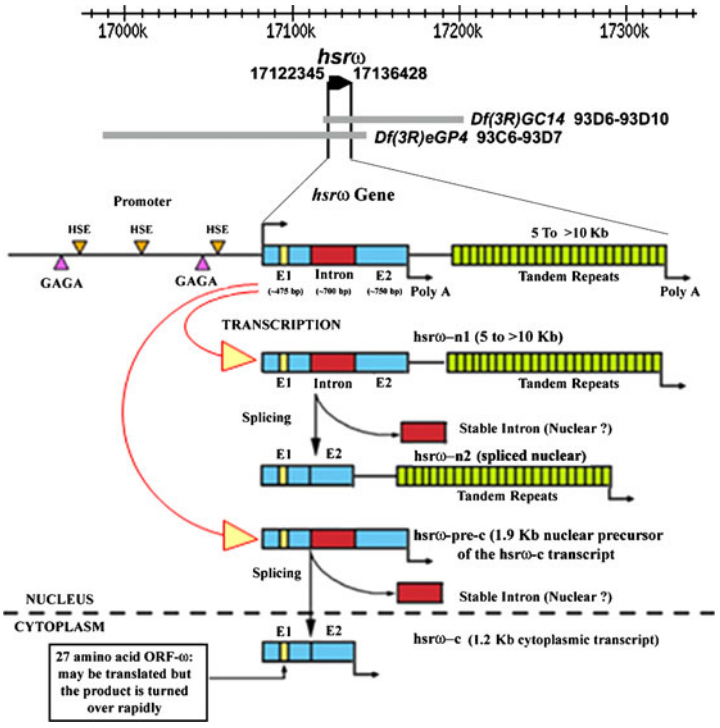



**Figure 4.** Architecture of the *hsrω* locus in *D. melanogaster*. On the top are shown the genomic coordinates of the *hsrω* gene with the numbers on a scale reflecting the base numbers on chromosome 3; the two parallel vertical lines indicate the span of the *hsrω* gene with the arrowhead showing the direction of transcription. The base numbers of transcription start and termination sites, with reference to the genomic sequence of chromosome 3 are also indicated (data from *http://www.flybase.org*). Spans of two deletions, *Df(3R)GC14* and *Df(3R)e$^{Gp4}$*, used in several studies described in the text are indicated by the gray horizontal bars. Some of the regulatory elements in the proximal promoter regions, as identified by Mutsuddi and Lakhotia (1995) and the transcribed region comprising of two exons (E1 and E2), one intron, and the long stretch (5 to >10 kb) of tandem repeats (280 bp repeat unit) are shown in the middle. Two primary transcripts, hsrω-n1 and hsrω-pre-c are produced, each of which is spliced to remove the ~700 b long intron. The hsrω-n1 and hsrω-n2 transcripts remain within the nucleus while the 1.2 kb hsrω-c transcript moves to cytoplasm where its ORF-ω may be translated. The hsrω homologs in other species of *Drosophila* also show a comparable architecture and transcript pattern.

entire transcription unit, while the shorter one is about 1.9 kb and corresponds to part of the proximal unique region till the first poly-A signal, which is about 700 bp prior to the beginning of the stretch of repeats.

c. The shorter 1.9 kb nuclear transcript is spliced to remove the single intron (~700 bases long) and the 1.2 kb transcript is transported to cytoplasm.

d. Neither the unique part nor the repeats show any significant degree of homology between the species except for a remarkably conserved stretch of about 15 and 59 bases at the 5′ and 3′ exon–intron junctions, respectively.

e. Amidst the high sequence divergence of the repeat units at the 3′ end of the gene in the two species (*melanogaster* and *hydei*), a nonamer (ATAGGTAGG) appears to be conserved so that it is present at nearly equal intervals in the repeat part of the large nuclear transcript.

f. The ~1.2 kb cytoplasmic transcript is associated with mono- or di-ribosomes and carries a potentially translatable open reading frame (ORF-ω) that may encode 23–27 amino acids in different species (Garbe *et al.* 1986) but whose translational product is not detectable (Fini *et al.* 1989). Intriguingly, the amino acid sequence potentially encoded by the ORF-ω in different species shows little conservation (Garbe *et al.* 1986).

Since the 93D locus was heat shock inducible but yielded only RNA as the final product, the gene was named *hsrω* or *heat shock RNA omega* (Bendena *et al.* 1989a), and the

three transcripts were named (Hogan *et al.* 1994) as hsrω-n (>10 kb, nuclear), hsrω-pre-c (1.9 kb, nuclear and precursor to the cytoplasmic transcript) and hsrω-c (1.2 kb, cytoplasmic). These three transcripts have also been named as ω-1, ω-2 and ω-3 (Garbe *et al.* 1989) or Hsrω-RB, Hsrω-RC and Hsrω-RA (*http://flybase.org*), respectively. It may be noted that until recently it was believed that the >10 kb hsrω-n transcript is not spliced. However, we (Mallik and Lakhotia 2011) recently found that this transcript is also spliced and that the unspliced as well as spliced forms, named, respectively, as hsrω-n1 and hsrω-n2, exist in the nucleus.

The realization that the 93D or the *hsrω* locus does not produce any detectable protein product was a disappointment in the 1980s, when biology was undergoing revolutionary changes ushered in by the increasingly efficient recombinant DNA methods and at a time when any DNA sequence not involved in production of proteins was generally considered 'junk' or 'selfish' and, therefore, not worthy of funding. The rapid sequence divergence of *hsrω* locus in *Drosophila* species also appeared to agree with its 'selfish' nature, a view further supported by the failure to isolate any point mutation in the *hsrω* locus (Mohler and Pardue 1982, 1984). As a consequence, both the Hovemann as well as the Pardue laboratories gradually stopped an active pursuit of this locus in the 1990s.

## 5. Diverse inducers or modifiers of its induction exist but the mechanism is not understood

We could not initiate typical molecular biological studies of the locus for several years because of the limited resources and facilities. In the absence of any mutant allele of the *hsrω* gene, a direct functional analysis of this non-coding gene was also not possible. Consequently, 93D puffing kept us engaged for some years and most of our studies during the 1980s and early 1990s essentially addressed the phenomena rather than the mechanisms in the hope that studies of the phenomena would give us some insight into functional significance of the 93D puff and the underlying mechanisms.

### 5.1 *Amides as specific inducer of 93D puff*

A review on heat-shock-induced gene activity by Ashburner and Bonner (1979) cited a publication by Gubenko and Baricheva (1979) in which colchicine was reported as a specific inducer of the 20CD puff of *D. virilis*, which on the basis of the above-noted studies of Peters *et al.* (1980) was a homolog of the 2-48B puff of *D. hydei*. Since the 2-48B puff was identified as a homolog of the 93D puff, we wanted to test if colchicine would also be a singular inducer of the 93D puff. We expected that colchicine, being a well-known disruptor of microtubules (Dustin 1978), may provide a better understanding of the significance of the 93D puff than BM, whose cellular effects were little understood. Colchicine as well as colcemid indeed mimicked BM in selective induction of the 93D puff (Lakhotia and Mukherjee 1984). To confirm the involvement of microtubules in 93D puff induction, Singh and Lakhotia (1984) examined cold shock and several other known microtubule poisons such as chloral hydrate, diamide, podophyllotoxin, vinblastin, nocodazole and griseofulvin to see if they induced the 93D puff or its homologs in other species. Surprisingly, none of these, except the cold shock, induced the 93D or other heat shock puffs (Singh and Lakhotia 1984). Thus, the disruption of microtubule organization did not appear to be the cause for selective induction of the 93D puff and its homologs in other species. Subsequent studies suggested that the induction of 93D puff by colchicine was due to its being an amide (*see* below).

Behnel (1982) reported that thiamphenicol specifically induced only the 93D puff, while Srivastava and Bangia (1985) noted that paracetamol also specifically induced the 93D puff. A structural analysis of these 93D-inducing chemicals (benzamide, colchicine, colcemid, thiamphenicol and paracetamol) with very different known effects on cellular metabolism revealed that all of them were amides (Tapadia and Lakhotia 1997). Therefore, we tested several other amides (nicotinamide, acetamide, formamide and 3-aminobenzamide) for their 93D-inducing activity, and indeed, all of them had the same effect as BM (Tapadia and Lakhotia 1997). Thus specific induction of the 93D puff and its homologs in other species by these diverse chemicals seem to be dependent upon their being amides. By this time, 3-aminobenzamide, a derivative of BM, was widely used as an inhibitor of poly-ADP ribose polymerase, or PARP (Sims *et al.* 1983), especially in DNA repair studies. In our paper (Tapadia and Lakhotia 1997) we discounted the possibility that the specific induction of the 93D puff may be related to the well-known inhibition of PARP by several of these amides because while that enzyme was inhibited with as little concentration of the amides as 3 mM, the 93D-inducing activity required 10 mM concentration. However, this conclusion needs re-examination in view of the recent report (Ji and Tulin 2009) of interactions between PARP and different hnRNPs, which in turn interact with the hsrω-n transcripts (section 9).

### 5.2 *Developmental conditions affect 93D inducibility*

Lakhotia and Singh (1985) reported that the 93D puff was refractory to induction by heat shock (37°C) or BM

or colchicine in salivary glands of *D. melanogaster* late third instar larvae that were reared at 10°C since hatching. However, transferring these glands from 10°C to 24°C by itself induced the 93D puff, without induction of the other heat shock puffs. Another intriguing observation was that ingredients in the fly food may affect the BM or colchicine inducibility of the 93D puff (Lakhotia 1989). Significance and mechanism of the effect of developmental conditions on 93D puff inducibility have remained unexplored.

### 5.3 *β-Alanine as a modifier of heat shock inducibility of 93D*

Based on the consistency and reproducibility of effects of non-induction of the 93D puff under certain conditions of heat shock on the 87A and 87C puffing after heat shock, we felt that the state of active transcription at the 93D puff has some role to play in the relative puffing of the two Hsp70-coding loci (Lakhotia 1987). However, Hochstrasser (1987) questioned this because he found unequal puffing at 87A and 87C even when the 93D puff appeared to be typically induced by heat shock. In this context, we used the *T(1;3)eH*$^{2}$*, red e* chromosome which displayed position effect variegation at the *ebony* locus (Henikoff 1980; Lakhotia *et al.* 1990). We expected the 93D locus, being close to the *ebony* locus, to also show variegated expression in this chromosome, and therefore, we hoped to examine the effect of such variable induction of the 93D puff on the relative levels of the 87A and 87C puffs in different cells of the same gland. Although a variegated induction of the 93D puff was not observed, these studies revealed an unexpected and intriguing effect of mutant alleles of the *ebony* or the *black* locus on the heat-shock-induced activity of the 93D puff (Lakhotia *et al.* 1990). Loss of function alleles of the *ebony* enhance while those of the *black* reduce levels of β-alanine in larval hemolymph (Wright 1987). Both these mutant conditions were found to inhibit induction of 93D puff by heat shock (Lakhotia *et al.* 1990). Feeding wild-type larvae with excess β-alanine also caused the 93D puff to be refractory to heat shock, but feeding the *black* homozygous mutant larvae on β-alanine restored heat-shock-induced puffing of the 93D locus (Lakhotia *et al.* 1990). Interestingly, the changes in levels of β-alanine due to mutation or feeding had no effect on induction of the 93D puff by BM. The mechanism of the inhibition of 93D puffing after heat shock by altered levels of β-alanine remains unexplored. In the context of the effect of 93D puffing during heat shock on 87A and 87C activity, the absence of 93D puff during heat shock following changes in β-alanine levels was also found to be associated with unequal puffing of the 87A and 87C sites (Lakhotia *et al.* 1990).

### 5.4 *Dosage compensation for developmental expression of 93D*

During their screen for mutations at the 93D locus, Mohler and Pardue (1982, 1984) generated two small deficiencies (*Df(3R)e*$^{Gp4}$ and *Df(3R)GC14*) spanning the 93D region such that their overlap homozygously deleted the 93D heat-shock- and BM-inducible locus (Burma and Lakhotia 1986) while the flanking genes were believed to be present at least in one copy in the *Df(3R)e*$^{Gp4}$*/Df(3R)GC14* trans-heterozygotes (figure 4). The *93D* or *hsrω*-null *Df(3R)e*$^{Gp4}$*/Df(3R)GC14* trans-heterozygotes have poor viability, with nearly 80% dying as embryos or early larvae (Mohler and Pardue 1984; Lakhotia and Ray 1996) while the survivors emerge as very weak short-lived flies, suggesting an essential developmental requirement of this non-coding gene. Studies on the heat shock and benzamide inducibility of heat shock puffs in these deficiency chromosomes revealed another interesting property of the 93D puff. Burma and Lakhotia (1986) found that the developmental or BM-induced rates of transcription at the 93D locus in *Df(3R)e*$^{Gp4}$*/+* or *Df(3R)GC14/+* salivary glands, with only one copy of the locus, were equal to those in corresponding wild-type glands, which suggested that the developmental or BM-induced activities of the 93D puff were dosage compensated. On the other hand, heat shock caused regression of this puff in either of these genotypes, which was accompanied by unequal puffing of the 87A and 87C sites. Surprisingly, in *hsrω*-null *Df(3R)e*$^{Gp4}$*/Df(3R)GC14* trans-heterozygotes, the 87A and 87C puffs in about 50% of the glands were equally active, while in the remainder, the 87C puff was less active than the 87A (Burma and Lakhotia 1986).

## 6. Non-induction of the 93D puff during heat shock affects the Hsp70-encoding twin puffs at 87A and 87C

We were intrigued by the recurring observation that whenever the 93D puff was not induced during heat shock, the puffing activity at the 87A and 87C puffs, the twin loci harboring two and three copies, respectively, of *Hsp70* genes (Livak *et al.* 1978), was differential instead of being nearly equal as seen after a simple heat shock to larval salivary glands (reviewed in Lakhotia 1987, 1989). The 87C site also carries heat-inducible alpha-beta (αβ) repeats (Livak *et al.* 1978). To assess the actual levels of the hsp70 and αβ transcripts at the two sites following the different experimental conditions, we (Sharma and Lakhotia 1995) quantified their levels at the 87A and 87C heat shock loci of *D. melanogaster* by *in situ* hybridization of $^{35}$S-labelled antisense RNA probes with polytene chromosome squashes of larval salivary glands treated with heat shock,

benzamide, colchicine, heat shock followed by benzamide or heat shock in the presence of colchicine. These results (figure 5) revealed that the larger or smaller size of 87A and 87C puffs under different conditions of heat shock correlated with altered levels of the hsp70 transcripts at the two sites and αβ-transcripts at the 87C site. Interestingly, colchicine, but not benzamide, treatment itself also enhanced αβ transcripts at the 87C site, without any induction of the hsp70 transcripts at either of the puff sites. Parallel studies (Lakhotia and Sharma 1995, see section 7) showed that the hsrω transcript profile at the 93D puff in these different conditions also varied. These results seem to establish a causal relationship between hsrω transcripts synthesized at the 93D locus and RNA metabolism at the 87A and 87C sites, although the underlying mechanism remains unknown. In this context, it is interesting to note that another study (Kar Chowdhury and Lakhotia 1986) compared the 87A and 87C puffs of *D. simulans* and *D. melanogaster* sibling species; it was found that the 87A and 87C puffs of *D. simulans,* which does not carry αβ-repeats at the 87C site, were equally induced by heat shock even when the 93D puff was not induced due to combined treatment with BM or colchicine. Interestingly, in the interspecific hybrid polytene cells, the 87A and 87C loci on the two species' homologs behaved in their parental species-specific manner. This difference between the two species appeared to correlate with the absence of the αβ repeats at the 87C locus in *D. simulans* (Kar Chowdhury and Lakhotia 1986). Significance of the heat-shock-inducible αβ repeats at the 87C site of *D. melanogaster* and their interaction with the hsrω transcripts remains unknown.

Our recent unpublished studies show that conditions under which the optimal levels of hsrω transcripts are not available in the cell during or after heat shock, association of some of the hnRNPs with the 87A and 87C puff sites is affected. Further studies will help us understand this relationship.

## 7. Inducer-specific transcript profile of the *hsrω* gene

I visited the Pardue laboratory at MIT (USA) for 3 months in 1985 on a Fulbright fellowship and initiated studies to examine the nature of transcripts induced by BM and colchicine treatments. This collaboration resulted, some years later, in a research paper (Bendena *et al.* 1989b) and a review (Bendena *et al.* 1989a). It was found that BM and colchicine indeed enhance levels of different hsrω tran-

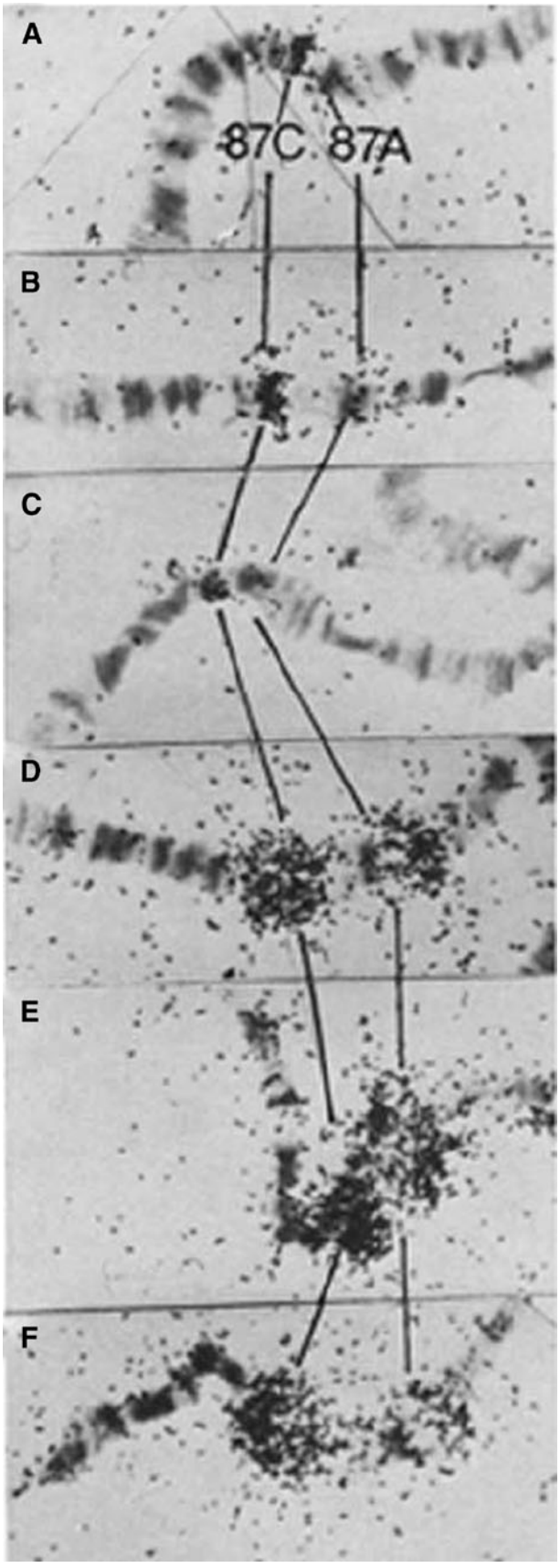



**Figure 5.** *In situ* hybridization of $^{35}$S-labelled hsp70 antisense RNA probe to RNA on the 87A and 87C sites of polytene chromosomes in (**A**) control salivary glands or after treatment with (**B**) benzamide, (**C**) colchicine, (**D**) heat shock, (**E**) heat shock followed by benzamide and (**F**) heat shock in the presence of colchicine. Reproduced from Sharma and Lakhotia (1995).

scripts in diploid Schneider cells as well, but the different hsrω transcripts are affected in an inducer-specific manner. While heat shock causes elevation of all the hsrω transcripts, BM and colchicine elevate only the large nuclear hsrω-n transcript. These studies also revealed interesting differences in kinetics of accumulation and degradation of the different hsrω transcripts in an inducer-specific manner. The developmentally expressed hsrω-n transcripts had a slow turnover, but those induced by heat shock, BM or colchicine turned over rapidly; however, this turnover appeared to be dependent upon continued transcription since in the presence of actinomycin D, the hsrω-n transcript levels remained high even after withdrawal of BM (Bendena *et al.* 1989b). In contrast, the hsrω-pre-c and hsrω-c transcripts had very short half-lives but were stabilized by inhibition of translation. The differential effects of heat shock, BM, colchicine and protein synthesis inhibitors (Bendena *et al.* 1989b) suggested that while the large hsrω-n transcript had roles in nuclear activities, the cytoplasmic hsrω-c transcript was related to translational activities. Because only the first four amino acids of the short but translatable ORF-ω in the cytoplasmic transcript (Fini *et al.* 1989) are strongly conserved, it appeared that it was the act of translation of ORF-ω, rather than its product, that was important to the cell (Bendena *et al.* 1989a). Accordingly, while reviewing the 93D puff story at the first conference of the Asia-Pacific Organization of Cell Biology (APOCB) at Shanghai in 1991, I suggested that the 1.2 kb hsrω-c transcript monitors the cells' translational activities. In a later more detailed review, we (Lakhotia and Sharma 1996) suggested that the hsrω locus 'has important house-keeping functions in transport and turnover of some transcripts and in monitoring the "health" of the translational machinery of the cell'.

Using quantitative RNA:RNA *in situ* hybridization with squash preparations of polytene chromosomes, we (Lakhotia and Sharma 1995) examined relative levels of the different hsrω transcripts *in situ* at the 93D puff following various treatments. Heat shock enhanced levels of all the three transcripts, viz., hsrω-n, hsrω-pre-c and hsrω-c at the 93D puff site in a coordinated manner; BM treatment caused a significant increase in the levels of hsrω-n and hsrω-pre-c, whereas colchicine resulted in increased levels of the hsrω-n and hsrω-c RNA species at the 93D site. This study also suggested that the hsrω-pre-c RNA is spliced at the site of synthesis with the spliced-out 'free' intron (hsr-omega-fi) accumulating at the puff site in a treatment specific manner. It was also found that in certain conditions under which the 93D puff was not induced in polytene cells and showed little $^{3}$H-uridine incorporation, the amount of hsrω transcripts present at the 93D region of polytene chromosomes remained nearly as high as when it was typically induced (Lakhotia and Sharma 1995). This suggests additional regulation at the level of turnover of these transcripts.

## 8. Complexity of regulation of developmental and induced expression of *hsrω*

Although the 93D puff has been known mostly for its unique inducible properties, this locus also forms a developmental puff in salivary glands of wild-type late third instar larvae (Ashburner 1967; Mukherjee and Lakhotia 1979). Northern analyses of total cellular RNA from unstressed salivary glands/embryos/whole organisms or from cells in culture always showed this gene's different transcripts in all preparations (Garbe *et al.* 1986). The very poor viability of *hsrω*-nulls (*Df(3R)e*$^{Gp4}$*/Df(3R)GC14*) also suggested this gene to have developmental expression and essential functions. Bendena *et al.* (1991), using a radioactively labelled RNA probe for *in situ* hybridization with hsrω transcripts in histological sections of different tissues from the various developmental stages, showed that the three omega transcripts are produced in almost all tissues in a regulated manner, with ecdysone probably playing a role in the tissue- and stage-specific changes in hsrω transcript profile. Only the preblastoderm embryos and the primary spermatocytes were reported (Bendena *et al.* 1991) to be devoid of hsrω transcripts.

We (Mutsuddi and Lakhotia 1995) re-examined the developmental expression of the *hsrω* gene by whole organ colorimetric RNA:RNA *in situ* hybridization and by following expression of the *lacZ* reporter gene placed downstream of different stretches of the *hsrω* promoter in transgenic lines generated by us. Our observations were generally in agreement with those of Bendena *et al.* (1991) but additionally revealed that the male as well as female gonial cells and oocytes in ovary do not contain hsrω transcripts. Based on these and other results, it was suggested that the variety of non-protein-coding transcripts of the *hsrω* gene have vital 'house-keeping' functions (Mutsuddi and Lakhotia 1995; Lakhotia and Sharma 1996), although the nature of such 'functions' remains elusive.

The β-galactosidase reporter expression studies in germline transformants carrying the *lacZ* reporter under different regions of the *hsrω* promoter (Mutsuddi and Lakhotia 1995) also helped in analysis of the promoter of the *hsrω* locus. The β-galactosidase reporter expression suggested that the region between −346 bp to −844 bp upstream contained major regulatory element/s for most of the developmental expression of this gene although this region was not sufficient for its expression in male and female reproductive systems. It was further shown that heat shock but not the BM or colchicine response elements are present within the proximal 844 bp of the *hsrω* promoter (Lakhotia and Mutsuddi 1996). Based on results of another unique study utilizing small chromosomal deficiencies, Lakhotia and Tapadia (1998) suggested that the amide response element

may be located more than 21 kb upstream of the *hsrω* gene and that intervening sequences may not interfere with action of the putative amide response element.

In a later more detailed study (Lakhotia *et al.* 2001), an enhancer-trap line carrying *P-lacZ* transposon insertion at −130 bp position of the *hsrω* promoter ($hsr\omega^{05241}$) was used to monitor the developmental and induced expression of the *hsrω* gene in a greater variety of embryonic, larval and adult tissues. The developmental, heat-shock- and BM-induced β-galactosidase reporter gene expression in this enhancer-trap line was compared with expression of the *hsrω* gene itself in different tissues through RNA:RNA *in situ* hybridization in the transposon insertion line and in wild type. Further, the two earlier generated transgenic lines (Mutsuddi and Lakhotia 1995) with the β-galactosidase reporter activity driven by different regions of the *hsrω* promoter were also used for comparison. This detailed study of the developmental expression revealed that the embryonic pole cells and the hub cells in testes and ovaries also do not express the *hsrω*, neither normally nor after heat shock. Significantly, in spite of insertion of a big transposon in the promoter, expression of the $hsr\omega^{05241}$allele in the enhancer-trap line, as revealed by *in situ* hybridization to cellular hsrω transcripts, was comparable to that in unstressed and heat-shocked wild-type embryonic, larval and adult somatic tissues examined. The BM inducibility of the 93D puff also appeared to be unaffected by the *P*-transposon insertion in the $hsr\omega^{05241}$ chromosome. On the other hand, expression of the *lacZ* reporter in this enhancer-trap line paralleled the *hsrω* RNA in all diploid cell types, but in the polytene cells, the β-galactosidase reporter activity in unstressed as well as heat-shocked tissues was completely absent (Lakhotia *et al.* 2001).

Small differences in the site of insertion of P-transposon in the *hsrω* gene promoter have been found to result in distinctly different phenotypic consequences. Insertion of the EP (enhancer promoter)-element in the promoter region of *hsrω* gene by itself causes small differences in its developmental expression, which are associated with specific phenotypes in adult flies. The EP-transposon insertion in the *EP93D* and *EP3037* lines is at −130 and −144 bp, respectively (Mallik and Lakhotia 2009a). However, even with this small difference in the insertion sites, the *EP3037* homozygous flies are generally much weaker and less fecund than the *EP93D* homozygotes (Mallik and Lakhotia 2011). Further, when driven by GAL4 drivers, these two EP alleles result in distinctly different phenotypic consequences in certain instances (Mallik and Lakhotia 2009a, b, 2010a, 2011). Apparently, the local chromatin structure/regulatory elements in the *hsrω* promoter affect the EP allele expression.

The recently generated Exelixis and Drosdel collections of chromosomal deletions with molecularly defined breakpoints have provided chromosomes in which defined region of the *hsrω* promoter or only its transcribed region is deleted. Using some of these molecularly defined short deletions, Akanksha in my laboratory has recently (unpublished) found that even the proximal 124 bp of the *hsrω* promoter can drive some developmental expression so that genotypes with only one copy of the *hsrω* gene carrying just the proximal 124 bp promoter region can survive under normal conditions although they are sensitive to heat shock.

Together, these results clearly indicate that the *hsrω* gene's promoter is complex, as may be expected from the diversity of tissues in which this gene is expressed in a regulated manner and the variety of environmental factors that influence its transcriptional activity and the stability/turnover of its different transcripts. Several non-exclusive possibilities about its transcriptional regulation need further studies: (i) most of the developmental and heat-shock-induced regulation is at least partially controlled by the proximal 124 bp of the upstream region, (ii) some of the regulatory sites may be present within the transcribed region and/or (iii) the far upstream regulatory elements can exert their effects across the transposon insertion or other intervening genes.

## 9. Analysis of *hsrω*'s 'functions' through genetic interaction studies

Identification of functions of a non-coding gene is a difficult task when no point mutation is available and, especially more so, when the base sequence of the homologs changes as rapidly between related species as recorded for the *hsrω* locus. The two small deficiencies, $Df(3R)e^{Gp4}$ and *Df(3R) GC14* (Mohler and Pardue 1982, 1984), spanning the *hsrω* locus provided the only opportunity until recently for a genetic approach to examine functions of this enigmatic gene. When these two deletions are brought in trans, they completely delete both copies of the *hsrω* gene (figure 4). However, since their proximal and distal breakpoints are not molecularly mapped (Mohler and Pardue 1984), the extent of flanking regions that would be absent on both homologs or would be present only on one of the deficiency chromosomes remains uncertain. Nevertheless, in spite of this limitation, these two deficiencies did provide some insights into the functions of the *hsrω* gene. The high early lethality of the $Df(3R)e^{Gp4}$/*Df(3R)GC14* trans-heterozygous 93D-null indicated essential developmental functions of this locus. On the other hand, their increased thermo-sensitivity suggested a role of these transcripts in the heat shock response, notwithstanding the fact that the synthesis of heat shock proteins was not affected in the *hsrω*-null individuals (Mohler and Pardue 1984; Lakhotia 1987, 1989). A role of the *hsrω* gene in thermal adaptation has also been noted in a series of population genetic studies by S. McKechnie's group using its natural indel and other alleles in Australian

populations (McKechnie *et al.* 1998; McColl and McKechnie 1999; Anderson *et al.* 2003; Collinge *et al.* 2008, also see section 11.5 for my laboratory's recent observations on the role of this gene in surviving heat shock).

To investigate the possible functions of these non-coding transcripts, we initially made use of the above-noted two deletions in genetic interaction studies. Following a report by Morcillo *et al.* (1993) that heat shock enhanced binding of the Hsp83 with the 93D puff, we found that *hsp83* mutant alleles dominantly enhanced the high lethality of *Df(3R)* $e^{Gp4}$*/Df(3R)GC14* trans-heterozygous progeny (Lakhotia and Ray 1996). Likewise *Ras* mutant alleles, which were shown to interact with Hsp83 (Cutforth and Rubin 1994), also dominantly enhanced the *Df(3R)*$e^{Gp4}$*/Df(3R)GC14* trans-heterozygous phenotype (Ray and Lakhotia 1998). Currently, Mukulika Ray in my laboratory is examining interactions of the Ras pathway and Hsp83 with the hsrω transcripts using the *hsrω-RNAi* transgenic lines recently generated in our laboratory (section 11.2).

In an attempt to generate mutant alleles of the *hsrω* gene through transposon mutagenesis, TK Rajendra mobilized the P-transposon insertion in the $hsr\omega^{05241}$ allele. Several of the lines in which the P-element was completely lost shared a common phenotype that the homozygotes displayed a prolonged larval life with small-sized imaginal discs, brain ganglia and salivary glands, culminating in death as very early pupae. This was initially very exciting as we hoped that we may possibly have isolated some mutant alleles of the *hsrω* that had distinct phenotypes! However, further genetic and molecular analyses revealed that in all these lines, the *hsrω* gene and its promoter were normal, having been fully repaired following excision of the P-transposon from their genome and thus the recessive lethal phenotype appeared to be due to a new mutation in some other gene. Sonali Sengupta (Sengupta 2005) mapped the second site mutation generated during the P-mobilization to a 39 kb interval in the 93E13-94 F1 region on the right arm of chromosome 3 by recombination as well as deletion mapping, and renamed the mutation as *l(3)pl* (*pupal lethal*). Since all the then publicly available mutant alleles of genes predicted to be present in the 39 kb interval complemented this phenotype, the identity of the *l(3)pl* remained enigmatic (Sengupta 2005). Of interest in the context of *hsrω* story, however, was the observation that the omega speckles (section 10) in larval tissues of the $l(3)pl^{10R}$ homozygotes were clustered as seen in wild-type cells after a mild heat shock, and thus we believe that the $l(3)pl^{10}$ allele interacts with *hsrω* transcripts (Sengupta 2005). Current studies by Akanksha in my laboratory suggest that the $l(3)pl^{10R}$ mutation nay affect the *DNA-pol epsilon* (*CG6768*) gene located in this region, but for which no mutant allele is publicly available so far. This is being examined further.

## 10. Association of different hnRNPs and other proteins with *hsrω* locus during heat shock and the essential role of hsrω-n transcripts in organizing the nucleoplasmic omega speckles

Initial ultrastructural studies revealed unique characteristic large RNP particles in the nucleoplasm and on or in the vicinity of the 2-48B puff of *D. hydei* (Derksen *et al.* 1973; Derksen 1975) and the 93D puff of *D. melanogaster* (Dangli *et al.* 1983). These studies further suggested that the proteinaceous core of these particles was surrounded by RNA (Derksen *et al.* 1973).

A remarkable collection of monoclonal antibodies against chromosomal proteins was generated by Saumweber *et al.* (1980). These antibodies, which continue to be generously provided by Dr H Saumweber to users, have significantly helped in understanding some functions of this non-coding gene in normal development and under conditions of cellular stress. Dangli *et al.* (1983; also *see* Dangli and Bautz 1983) reported that in unstressed polytene cells, the P11 and Q18 antibodies, now known to recognize the Hrb87F and Hrb57A hnRNPs, respectively (Haynes *et al.* 1991; Buchenau *et al.* 1997), decorated many active and potentially active chromosome sites. However, their localization changed dramatically following heat shock so that most of the chromosome-bound antibodies were detectable only at the 93D puff site (Dangli *et al.* 1983). Subsequently, nearly all the tested hnRNPs and several other RNA-binding proteins have been found to show exclusive and/or high presence at the 93D puff in heat-shocked cells (reviewed by Lakhotia *et al.* 1999; Lakhotia 2003; Jolly and Lakhotia 2006).

Antibodies present in sera from certain class of patients suffering from autoimmune ankylosing spondilytis show specific binding with the 93D puff in heat-shocked salivary glands (Lakomek *et al.* 1991). It is possible that these auto-antibodies recognize some nuclear proteins like hnRNPs and thus specifically decorate the heat-shock-induced 93D puff. This interesting issue, although with the potential of becoming a diagnostic tool, has not been followed further.

The significance of an early observation of Spruill *et al.* (1978) that cyclic-GMP is present at many active chromosome sites in control cells but preferentially accumulates at the 93D puff following heat shock also remains unknown. In the context of the accumulation hnRNPs, etc., on the 93D puff in heat-shocked cells, it will be interesting to examine if cyclic-GMP is involved in some modifications of the hnRNPs while they are sequestered at the 93D puff.

An interesting finding of Samuels *et al.* (1994) was that the Sxl protein, one of the key players in sex determination and dosage compensation in *Drosophila* in female cells, moved from its normal presence at many chromosomal sites in unstressed polytene cells and became exclusively

localized at the 93D puff following prolonged heat shock. It remains to be seen if the association of Sxl with the hsrω transcripts in omega speckles and the 93D locus has some bearing on the sex-specific lethality seen under certain conditions of under- or overexpression of the *hsrω* gene (Mallik and Lakhotia 2011).

Dangli *et al.* (1983) as well as Samuels *et al.* (1994) suggested that the 93D puff may serve as a storage site for the RNA processing machinery under stress conditions. However, the significance of such unique association of a variety of proteins at a non-coding site during heat shock essentially remained a strange phenomenon until we examined (Lakhotia *et al.* 1999; Prasanth *et al.* 2000) the distribution of the hsrω transcripts in normal and heat-shocked cells in intact, rather than squashed, cells by fluorescence *in situ* hybridization. These studies revealed that in unstressed cells the large nuclear hsrω-n transcripts are present, besides at the 93D site, as distinct speckles in the nucleoplasm but closer to chromatin regions (figure 6). However, in heat-shocked cells this RNA was exclusively localized in large quantities at the 93D site only. This observation was contrary to that of Hogan *et al*. (1994), who, using the spatially less resolved *in situ* hybridization with radioactive-labelled probe, reported that the nuclear distribution of hsrω-n transcripts in control and heat-shocked Schneider cells was similar.

Fluorescence co-immunostaining and RNA:RNA *in situ* hybridization showed that all the so far tested hnRNPs (Lakhotia *et al.* 1999; Prasanth *et al.* 2000 and other unpublished observations) are present in unstressed and *hsrω*-expressing cells at (i) the active chromosome sites, (ii) the 93D puff site and (iii) in the hsrω-n-RNA-containing speckles. Following heat shock, the hnRNPs disappear from nearly all the chromosome sites, and along with the disappearance of the nucleoplasmic hsrω-n RNA speckles, they accumulate at the 93D site (figure 6). Because (i) the formation of these nucleoplasmic speckles containing the diverse hnRNPs and related proteins is dependent upon the presence of hsrω-n RNA and (ii) these speckles are distinct from the interchromatin granule clusters (Spector *et al.* 1991), we (Prasanth *et al.* 2000) named these novel nuclear structures as omega speckles (figure 6). The primary role of hsrω-n transcripts in organizing the omega speckles has been reconfirmed by the observation that ablation of these transcripts through RNAi results in loss of omega speckles so that the diverse hnRNPs get diffusely dispersed in nucleoplasm while overexpression results in larger clusters of omega speckles (Mallik and Lakhotia 2009a, 2011). Prasanth *et al.* (2000) concluded that the hsrω-n transcripts 'play essential structural and functional roles in organizing and establishing the hnRNP containing omega speckles and thus regulate the trafficking and availability of hnRNPs and other related RNA binding proteins in the cell nucleus'. It is now well known that hnRNPs constitute a family of diverse but conserved RNA-binding proteins with multiple roles in nucleic acid metabolism, including the packaging and processing of different species of nuclear RNAs and their translational regulation (Han *et al.* 2010). The omega speckles thus are expected to influence a diverse array of the nuclear RNA processing activities in view of their association with a variety of hnRNPs.

Recent studies in my laboratory indicate that the omega speckles may also be heterogeneous in their composition in different cell types or even in the same nucleus since some of the omega speckle associated proteins, for example, NonA, PEP and S5, are not present in all the speckles in a given nucleus. In addition, the recent finding that the hsrω-n RNA also exists in spliced and unspliced hsrω-n1 and hsrω-n2 forms and that both may be associated with omega speckles (Mallik and Lakhotia 2011), may provide a further means of functional differentiation of the omega speckles. Curiously, while the omega speckles are very distinct in the

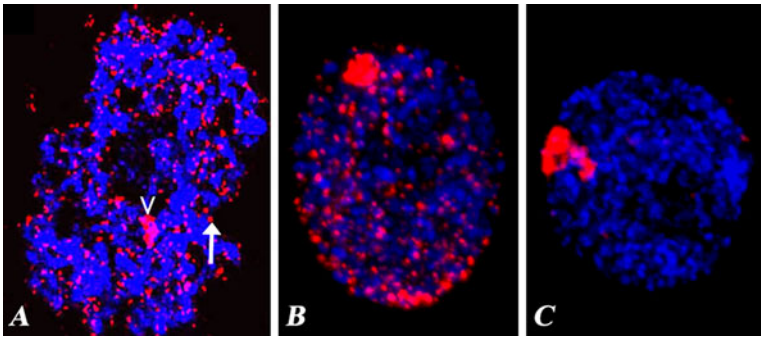



**Figure 6.** (**A**) Nucleoplasmic omega speckles (red, arrow) seen after *in situ* hybridization of hsrω 280b repeat riboprobe with RNA in Malpighian tubule polytene nucleus (chromatin blue). The arrowhead indicates the *hsrω* locus on the chromosome. Immunostaining of control (**B**) and heat shocked (**C**) Malpighian tubule nuclei with Hrb87F antibody. The Hrb87F (red) follows the same speckled pattern as the hsrω-n to RNA in control cells, but after heat shock, it shows near exclusive localization at the 93D site. DNA, stained with DAPI, is shown in blue in all the images. The image in **A** is from Sengupta (2005), while those in **B** and **C** were provided by Anand K Singh.

larval Malpighian tubule polytene cells, they are less prominent in the more polytenized larval salivary glands due to a more diffuse presence of hnRNPs in the nucleoplasm. The larval salivary glands also appear to have more of the spliced hsrω-n2 form compared to the hsrω-n1; whereas in whole larval or adult, the reverse situation obtains (unpublished observations). The significance of such differences remains to be examined.

Several reports (Jolly and Morimoto 1999; Denegri *et al.* 2001; Jolly *et al.* 2004) on the formation of nuclear stress bodies in human cells following heat shock revealed that the non-coding sat III repetitive sequences, associated with centromeric heterochromatin on chromosome 9, were transcriptionally induced by heat shock and were essential for accumulation of a variety of proteins at the site of transcription in the form of stress bodies. Intrigued by the striking parallels between the behaviour of hsrω transcripts and the human sat III transcripts in heat-shocked cells, I initiated email exchanges with Caroline Jolly, which finally developed into a review wherein we suggested that the hsrω-n and sat III transcripts are functional analogues and work through a common paradigm to dynamically regulate RNA-processing proteins through sequestration (Jolly and Lakhotia 2006).

Zimowska and Paddy (2002) reported that the Tpr protein, normally present at nuclear pores and as granules/speckles associated with the nuclear matrix, accumulates at the 93D puff following heat shock in parallel with proteins like NonA, PEP, etc. This paper, however, does not state if proteins like NonA or PEP associate with the Tpr granules in nucleoplasm although data indicate that the Tpr protein is not present at any of the developmental or other heat shock puffs. Thus, Tpr behaves differently from the other proteins that accumulate at the 93D puff after heat shock. Observations in our laboratory, using the Bx34 antibody against Tpr, obtained from Dr H Saumweber, reveal that, unlike the NonA or PEP or other hnRNPs, the Tpr granules do not colocalize with the omega speckles, although Tpr granules and the omega speckles are on the fibrillar nuclear matrix network that also contains Tpr, Surprisingly, however, contrary to the results of Zimowska and Paddy (2002), we (Anand K Singh and SC Lakhotia, unpublished) do not find any binding of the Bx34 antibody at the 93D puff in heat-shocked cells. These different results need further examination.

Around the time when we were excited about the discovery of omega speckles, we also noted (Rajendra *et al.* 2001) that males homozygous for the $hsr\omega^{05241}$ P-transposon insertion allele were completely sterile and they also displayed overexpression of hsrω-n transcripts in the cyst cells, a pair of which surrounds the bundle of 64 elongating sperm until their individualization. This correlation appeared very attractive and led us to surmise that the larger clusters of omega speckles seen in cyst cells in testes of $hsr\omega^{05241}$ homozygotes sequestered a greater proportion of hnRNPs, etc., and this affected the cyst cell function, which in turn prevented individualization of spermatozoa, resulting in male sterility (Rajendra *et al.* 2001). Subsequent studies in the laboratory, however, revealed that the $hsr\omega^{05241}$ mutation was actually not responsible for the recessive male sterility. It turned out that our earlier study (Rajendra *et al.* 2001) with the $hsr\omega^{05241}$ P-transposon insertion chromosome erroneously failed to detect a second site recessive mutation. Akanksha *et al.* (2008) showed that the male sterility was actually due to a second site recessive mutation rather than to the P-transposon insertion at −130 bp position in the *hsrω* promoter. The second site mutation ($ms^{21}$) has subsequently been mapped by Roshan Fatima (unpublished) in my laboratory to a novel axonemal dynein intermediate chain gene (*CG7051*) at 61B1 region on left arm of chromosome 3. The cause and consequence of the clustering of omega speckles in cyst cells in dynein intermediate chain mutant background remain unknown.

Although Lakhotia *et al.* (2001) did not notice a difference in the levels of hsrω transcripts in larval tissues of $hsr\omega^{05241}$ homozygotes, either through colorimetric RNA:RNA *in situ* hybridization or the *lacZ* reporter assay, a later study (Sengupta and Lakhotia 2006) using more sensitive fluorescence *in situ* hybridization showed that the omega speckles in eye disc cells of $hsr\omega^{05241}$ homozygous larvae were also clustered and larger, suggesting this allele to be overexpressing. The *hsrω* chromosome used by Sengupta and Lakhotia (2006) still carried the then unsuspected $ms^{21}$ mutation. Subsequently, however, it has been seen that the male fertile $hsr\omega^{05241}$ line also shows slightly enhanced expression of the hsr*ω* transcripts in the larval eye discs and a mild roughening of adult eyes.

Binding of polyADP-ribose (pADPr) to many proteins, mediated by PARP or otherwise, significantly modulates their activities; the removal of pADPr from the ribosylated proteins by poly(ADP-ribose) glycohydrolase (PARG) is also a regulated process. Among the many proteins that undergo such modifications, the ribosylation of hnRNPs is very interesting in the context of *hsrω* activity. Ji and Tulin (2009) have shown that heat shock enhances ribosylation of hnRNPs like Squid and Hrb98DE, so that these proteins lose their affinity for association with RNA at most of the developmentally active loci. PARG activity relieves the released hnRNPs of their pADPr moieties; these de-ribosylated hnRNPs then accumulate at the 93D puff in heat-shocked cells. Ji and Tulin (2009) reported that there is enhanced ribosylation in PARG mutants and, consequently, reduced binding of the two hnRNPs with the heat-shock-induced 93D puff. Further, the compromised PARG activity affected splicing of hsrω and Ddc transcripts, which they

correlated with a reduced binding of proteins like Squid and Hrb87F with RNA in the spliceosome machinery. While this model needs further studies and validation, the observed relationship between PARP and PARG activities and the accumulation of hnRNPs at the *hsrω* locus during heat shock is very interesting. The mechanism and processes underlying the congregation of hsrω-n transcripts and the various proteins at the 93D locus following heat shock remain to be examined. Nuclear matrix components are expected to play important roles in the dynamic nuclear relocation of hnRNPs and hsrω-n transcripts. In this context, our recent finding (Mallik and Lakhotia 2011) that hsrω transcripts interact with nuclear lamins is of interest. Further, as noted in sections 11.4 and 11.5, the *hsrω* activity also relates to ISWI and HP1 proteins. Further studies to explore the interactions between ribosylation, ISWI activity, hnRNPs, HP1, lamins, nuclear matrix, etc., with each other and with hsrω transcripts should provide novel insights into the overall chromatin regulation and nuclear activity.

A number of different speckle compartments are now known in eukaryotic nuclei, all of which contain variety of RNA species together with different RNA-processing and other proteins (Jolly and Lakhotia 2006; Prasanth and Spector 2007). It is likely that all classes of nuclear speckles carry non-coding RNA species. Based on our understanding of the omega speckles, it appears that the non-coding RNA species in these nuclear compartments help them function as dynamic 'storage depots' for a regulated on demand release of different classes of nuclear proteins and other factors involved in functions like organization and modification of chromatin, processing and transport of nuclear (including nucleolar) RNA, etc.

The discovery of omega speckles at the turn of the century was a milestone and made it possible to experimentally address the functional significance of this intriguing non-coding gene. Fortunately, this also happened at a time when non-coding RNAs were becoming increasingly attractive. Although much of the excitement about non-coding RNAs was with reference to the small miRNAs (and their various namesakes), by the beginning of this century, several large non-coding RNAs were also known as valid and functional molecules (Lakhotia 1996), and thus the hsrω transcripts' role in organizing the omega speckles was 'acceptable'. Our paper (Lakhotia *et al.* 1999) in which we raised the possibility that the hsrω transcripts may regulate the dynamics of hnRNPs, etc., was part of a collection of articles on large non-coding RNAs that I edited for *Current Science*. In my editorial (Lakhotia 1999), I argued for a due recognition of the significance of non-coding RNA and stated, 'Among a variety of factors that are already known to affect the higher order chromatin organization and consequently gene expression and "ribotype" of a cell, RNA is one as exemplified by the inactive X-chromosome in female mammals. Since the fine-tuned "ribotype" of a cell results in individual cell phenotype, the "ribotype" actually is subjected to natural selection. Additionally, since the "ribotype" can also generate new components of the genotype through reverse transcription, RNA molecules in a cell remain the prime players'. Since the prejudice against non-coding RNAs as functional molecules essentially stemmed from the popular interpretation of 'central dogma of molecular biology' that biological information that leads to production of proteins only is relevant, I concluded my editorial saying, 'Dogmas are helpful in providing directions for searches in a defined framework but they need continued revisions and modifications so that newer directions are found. Followers of the central dogma of molecular biology need to become less dogmatic since the living world is full of diversity and surprises' (Lakhotia 1999). It may be noted that although the basic point of the 'central dogma' that the flow of information from DNA to protein via RNA is unidirectional remains valid, the common (mis-)interpretation that has generally prevailed is that the genetic information not involved in production of proteins may be 'junk' or 'selfish'. Fortunately, with the large non-coding RNAs making their presence felt, this misconception is losing its ground.

## 11. Conditional RNAi or overexpression unravels pleiotropic roles of hsrω transcripts

Since classical mutant alleles of the *hsrω* locus have not been available, we (Mallik and Lakhotia 2009a) generated transgenic lines for the GAL4-UAS based (Brand and Perrimon 1993) conditional RNAi, using the 280 bp *hsrω* repeat sequence. We believe that this RNAi construct acts within the nucleus (Mallik and Lakhotia 2011) and thus essentially down-regulates only the hsrω-n transcripts. We also used the EP-transposon (Rorth 1996) for targeted overexpression of hsrω transcripts. These two approaches have provided very exciting insights into the multiple functions that these transcripts perform in a cell during normal development as well as under conditions of cell stress.

### 11.1 *Developmental effects*

Since the *hsrω* gene is expressed in a regulated manner in somatic cells of all developmental stages, it is expected to perform some functions in all cells. In agreement, overexpression or ablation of the hsrω transcripts in different tissues and developmental stages with a variety of GAL4 drivers revealed that a balanced expression of these non-coding transcripts is critical for survival and normal

development; a change in cellular levels of these transcripts generally had detrimental consequences, with extreme cases resulting in organismal lethality (Mallik and Lakhotia 2011). Since altered levels of the hsrω nuclear transcripts immediately affect the omega speckles, we believe that the developmental effects following the targeted down- or up-regulation of hsrω transcripts disrupt the dynamic homeostasis of RNA-processing proteins in the given tissue, which in turn has cascading effects on downstream events. Interestingly, we also found that in a few cases, ablation of these transcripts suppressed the mutant phenotype resulting from mis-expression of other genes, such as inactivated *apterous* gene or UAS-driven overexpression of lamin C protein (see Mallik and Lakhotia 2011). In a few cases, the hsrω-RNAi or overexpression had sex-specific effects (Mallik and Lakhotia 2010a, 2011); it remains to be seen if such sex-specific effects relate to the earlier noted interactions of Sxl protein with the hsrω transcripts.

In agreement with the abundant presence of the hsrω transcripts in nurse cells (Mutsuddi and Lakhotia 1995; Lakhotia *et al.* 2001), absence or down- or up-regulation of hsrω transcripts in ovarian follicles, due to chromosomal deletion, P-transposon insertion, RNAi or EP allele expression, affects oogenesis (Lakhotia *et al.* 1999; Srikrishna 2008, and other unpublished results). In parallel with the limited transcription of *hsrω* in the male germline, even in the complete absence of the hsrω transcripts, fertile sperms are produced in *Df(3R)e*$^{gp4}$*/Df(3R)GC14* males (Ray 1997; Lakhotia *et al.* 1999).

The *hsrω*-null condition or global activation of *hsrω-RNAi* transgene with *Act5C-GAL4* driver causes extensive embryonic or larval death but a certain proportion of such individuals regularly survive, and surprisingly, *Act5C-GAL4* and *UAS-hsrω-RNAi* can even be maintained together in a stock (Mallik and Lakhotia 2011). We believe that the successful survival of these flies with global down-regulation of the hsrω transcripts is due to other pathways that can take care of the critical *hsrω* functions, at least to a limited extent. Existence of such 'backup' or alternative pathways appears to be supported by the observation that the continued presence of *Act5C-GAL4* driver and the *UAS-hsrω-RNAi* responder through several generations in a stock reduces, but does not abolish, the lethality compared to that when the driver and the responder are brought together for the first time (Mallik and Lakhotia 2011). This may suggest a rapid selection or activation of the 'backup' pathways through epigenetic or other means. The nature and population dynamics of the possible 'backup' for *hsrω* functions need further study.

Apparently supporting the role of hsrω-n transcripts in sequestering hnRNPs and thus controlling processing of many hnRNAs, Johnson *et al.* (2009) reported that adult flies of certain lines of *D. melanogaster* showed reduced levels of hsrω-n transcript and also displayed higher rates of protein synthesis in adult ovaries. The enhanced rate of protein synthesis was presumed to result from a faster processing of mRNAs by the enhanced availability of hnRNPs consequent to reduced hsrω transcripts in these lines. While this looks attractive, some limitations of the experimental procedures of this study need to be considered. Johnson *et al.* (2009) measured the rate of protein synthesis only in ovaries but the hsrω transcript levels were measured in total RNA from whole flies, rather than from their ovaries only. Thus, a direct correlation between levels of the hsrω transcripts in ovaries with the rate of protein synthesis remains uncertain. Further, their design for measurement of rate of protein synthesis did not eliminate the possibility that unincorporated $^{35}$S-methionine also contributed significantly to their estimated 'rate' of protein synthesis. An additional complication arises from the recent finding in our laboratory (Mallik and Lakhotia 2011) that the hsrω-n transcripts can exist in unspliced as well as spliced forms; consequently, the RT-PCR-based estimates of hsrω-c transcripts in earlier studies failed to differentiate between the spliced hsrω-n and hsrω-c transcripts, since the amplicons size in both cases are identical. Therefore, the inferences of Johnson *et al.* (2009) need re-examination.

### 11.2 *Modulation of induced apoptosis*

Following our observation that eye damage as well as the high frequency of apoptosis in eye discs in organisms carrying two copies of the GMR-GAL4 driver transgene are nearly completely suppressed by hsrω-RNAi, we examined the consequences of down- or up-regulation of hsrω transcripts on apoptosis induced either by directly expressing the proapoptotic Reaper, Grim or Hid proteins or caspases (precursor as well as activated) in the eye and other imaginal discs (Mallik and Lakhotia 2009b). We showed for the first time that down-regulation of hsrω transcripts through RNAi suppressed JNK singalling and stabilized the *Drosophila* inhibitor of apoptosis protein 1 (DIAP1), presumably through its increased association with Hrb57A, which is released following the disappearance of omega speckles (Mallik and Lakhotia 2009a).

### 11.3 *Modulation of neurodegeneration*

Fernandez-Funez *et al.* (2000) reported the *hsrω*$^{05241}$ allele to be a dominant enhancer of the expanded polyQ toxicity in a fly model of Spinocerebellar Ataxia type 1. Following this report, we examined the interaction in greater detail. In the first study, Sengupta and Lakhotia (2006) confirmed the dominant enhancing effect of the *hsrω*$^{05241}$ allele on 127Q-induced neurodegeneration but found that the *hsrω*$^{05241}$

allele was actually an overexpressing allele, rather than loss-of-function allele as presumed by Fernandez-Funez *et al.* (2000). This study also showed that the omega speckles do not colocalize with the polyQ nuclear inclusion bodies (IB). However, the hsrω transcripts may affect the polyQ pathogenesis via the association with hnRNPs in omega speckles since a monosomic condition for Hrb87F (also known as Hrp36) was, by itself, also found to dominantly enhance the polyQ phenotype (Sengupta and Lakhotia 2006). The availability of the *hsrω-RNAi* and *EP* lines for down- or up-regulation, respectively, of the hsrω transcripts permitted more detailed mechanistic analyses of the modifying effect of these transcripts on polyQ pathogenesis in several different fly models (Mallik and Lakhotia 2009a, 2010a). Reduction in hsrω-n transcripts through RNAi nearly completely suppressed neurodegeneration following expression of different expanded polyQ proteins, while elevation of these transcripts through EP allele expression aggravated neurodegeneration. Significantly, the down-regulation of hsrω-n transcripts through RNAi was associated with disappearance/reduction of the polyQ IBs (Mallik and Lakhotia 2009a). Levels of the chromatin and transcription modulator CREB-binding protein (CBP) mRNA as well as protein were elevated under this condition (Mallik and Lakhotia 2010a). The elevated levels of free hnRNPs, following hsrω-RNAi, resulted in their enhanced association with CBP. Down-regulation of hsrω transcripts also improved proteasomal activity. The hsrω-RNAi thus interferes with the polyQ pathogenesis pathways at multiple levels (Mallik and Lakhotia 2010b) and thereby facilitates clearance of the polyQ IBs. Together with the above-noted suppression of induced apoptosis (section 11.2), hsrω-RNAi results in near complete suppression of neurodegeneration (Mallik and Lakhotia 2010a, b). Interestingly, unlike the suppression of polyQ damage, hsrω-RNAi had little protective effect on damage in eye discs caused by overexpression of normal or mutated tau protein (Mallik and Lakhotia 2009a).

### 11.4 *Interaction with ISWI chromatin remodeller*

An email from Dr Davide Corona in October 2007 stating that his laboratory had found *hsrω* to be one of the modifiers of phenotypes resulting from mis-expression of the ISWI-ATP-dependent chromatin remodeller initiated a regular exchange of ideas and reagents (fly stocks, antibodies, DNA clones, etc.) between our labs. This collaboration provided very exciting insights into the interaction between ISWI and hsrω transcripts (Onorati *et al.* 2011). The phenotypes of ISWI-nulls, like poor condensation of the polytene chromosomes, especially the X chromosome in male salivary glands, larval lethality and eye degeneration in mosaic eyes, etc., (Corona *et al.* 2007) are suppressed by hsrω-RNAi. These studies show that ISWI is essential for organization of omega speckles since in ISWI-null cells, the nucleoplasmic hsrω-n transcript and associated hnRNPs are seen as elongated trail-like structures, rather than the typical speckles. Interestingly, while the hsrω-n transcripts and ISWI display limited colocalization, immunoprecipitation with ISWI antibody pulls down the hsrω-n transcripts as well; further, the 280b repeat unit of hsrω-n transcripts can stimulate the ATPase activity of the ISWI *in vitro* (Onorati *et al.* 2011). The mechanistic details of interaction between *hsrω* and ISWI are not yet understood, but the interactions of either of these with PARP/PARG, HP1, nuclear lamins, CBP, etc., appear to have significant roles in this context.

### 11.5 *Role in recovery from heat shock*

We have also used these RNAi and overexpressing *EP* lines to understand the role of hsrω transcripts in heat shock response. As noted earlier, the *Df(3R)e*$^{Gp4}$*/Df(3R)GC14* trans-heterozygotes, which are null for the *hsrω* gene, show thermosensitivity in spite of normal induction of synthesis of heat shock proteins. In agreement with these earlier observations, it has been recently observed in our laboratory that, unlike wild-type embryos and larvae, those expressing *hsrω-RNAi* or *EP* during heat shock or those being nullosomic for the *hsrω* gene show nearly 100% death several hours after the thermal stress. To understand this rather enigmatic situation, we (Lakhotia *et al.* 2011, in preparation) have compared the dynamics of hnRNPs, RNA pol II and HP1 after heat shock and during recovery in wild-type salivary glands that have relatively reduced (because of RNAi or *hsrω* mono- or nullo-somy) or elevated (through EP expression) levels of hsrω transcripts during heat shock. The rapid re-distribution of hnRNPs to the 93D puff and of the RNA pol II on the heat-shock-induced loci is comparable in all the cases. However, unlike the rapid return of the hnRNPs and the RNA pol II to different chromosomal regions within 1 hr of recovery from heat shock, glands with reduced or elevated levels of the hsrω transcripts fail to re-localize these proteins to their expected chromosome sites as well as to the omega speckles. In cells with reduced hsrω transcripts, the hnRNPs neither get back to omega speckles nor move efficiently to chromosome sites but get distributed in a rather diffuse manner in the nucleoplasm in addition to some chromosomal sites. The RNA pol II is also seen on fewer developmentally active chromosome sites during recovery in these glands. In the *hsrω* overexpressing glands most of the hnRNPs as well as the RNA pol II continue to remain, even after 1 or 2 hr of recovery, at sites they moved to following the heat shock (Lakhotia *et al.* 2011, in preparation). In glands with down- or up-regulated hsrω transcripts, the hnRNPs showed abnormal association with the 87A and 87C puffs, which may provide an explanation

for the earlier noted (section 6) unequal puffing of these twin puffs when the 93D is not induced during heat shock.

The HP1 protein also shows redistribution to heat shock loci, including the 93D puff, following heat shock (Piacentini *et al.* 2003). Our results (Lakhotia *et al.* 2011, in preparation) show that depletion or overexpression of hsrω transcripts during heat shock affects the redistribution of HP1 protein to pre-stress condition during recovery.

These observations suggest that a balanced level of the hsrω transcripts is essential for a regulated movement of hnRNPs, RNA pol II and other proteins during recovery from heat shock. Consequently, in the absence of normal resumption of cellular activities, the organisms die over a period time, even though they produce the usual set of heat shock proteins during the heat shock period.

In view of the above-noted interactions of CBP, ISWI, lamin C, HP1, etc., with hsrω transcripts, it is likely that components of nuclear matrix and chromatin remodelling complexes are involved in the dynamic movement of hnRNPs, etc., away from the 93D puff during recovery from heat shock. In this context, it is significant that, like the hnRNPs that are associated with the nuclear matrix (Mattern *et al.* 1999; Arao *et al.* 2000), the omega speckles are also mostly distributed along the nuclear matrix and show rapid, but constrained, movements in live cells (Anand K Singh and SC Lakhotia, unpublished; also see supplementary video S1 in Onorati *et al.* 2011).

## 12. Non-coding transcripts as hubs in cellular networks

The wide variety of apparently unrelated pathways, with which the *hsrω* gene appears to interact, may lead one to suspect that these may be non-specific interactions. However, given the primary role of the hsrω-n transcripts in organizing omega speckles and thus modulating, directly or indirectly, activities of a diverse array of regulatory proteins, such pleiotropy is indeed expected. As discussed above, the *hsrω* locus is also very sensitive to intra- and extra-cellular conditions. Therefore, we believe that the non-coding *hsrω* gene functions as a hub in cellular networks (Arya *et al.* 2007; Mallik and Lakhotia 2010a). Based on the known genetic, cell biological or biochemical interactions of the *hsrω* gene with other pathways, a simplified schematic of the role of hsrω-n transcripts as a potential hub in cellular networks is presented in figure 7. Since each interacting protein has its own network connections, it is likely that the different hsrω transcripts help maintain homeostasis through their influence on the diverse networks; the actual signal transduced to and between the different networks would depend upon the input signals from other intra- and extra-cellular events. This model predicts that an imbalance between these non-coding transcripts and their interacting proteins or other factors will have consequences of varying magnitudes in different cells depending upon the temporal and spatial factors. It also appears likely that for such critical roles, the evolutionary processes would have generated 'backup' systems, mentioned in section 11.1, to provide for at least a limited survival under conditions when the prime transcripts or their activities are compromised.

A significant advantage that a large non-coding RNA has in its role as hub is that it may not only provide platforms for multiple proteins in view of the usually short nucleotide sequences needed at their binding sites, but being single-stranded, the RNA molecules can have a much greater plasticity in their secondary and tertiary structures. The tandem repeats seem to provide additional versatility. The higher-order structures of such large non-coding RNAs may be specifically modulated by binding of one or the other proteins or other conditions. A strong functional conservation in spite of the poor conservation of the *hsrω* base sequence even between different related *Drosophila* species is an indication of the importance of secondary and tertiary structures of the hsrω transcripts rather than the base sequence itself. The functional analogy between the hsrω-n transcripts and human sat III transcripts without any sequence homology is also a pointer in the same direction. A search for functional equivalents of *hsrω* in other organisms will, therefore, have to look for large non-coding RNAs with comparable protein association properties and/or similar higher-order structural features. Improved bioinformatic approaches and knowledge of the non-coding RNomes in different species will hopefully permit identification of the *hsrω* equivalents in increasing number of eukaryotes in the near future.

## 13. Epilogue

The journey of the 93D locus and its transformation from a 'nice-looking' puff to a mysteriously conserved but apparently non-coding *hsrω* 'gene', to the organizer of omega speckles and, therefore, a potential hub regulating multiple interconnected networks, has indeed been fascinating. As this story progressed, my conviction in a major role played by the enormously large non-coding component of eukaryotic genome in generating the biological complexity has grown (Lakhotia 1987, 1989, 1996, 1999, 2003; Jolly and Lakhotia 2006; Mallik and Lakhotia 2007). My interest in the inactive heterochromatin arose during my doctoral studies primarily because of the chromosomal level differences in the inactive-X in female mammals and the hyperactive-X in male flies (Lakhotia 1970). Karyotyping of some mammals with prominent centromeric heterochromatin (Rao and Lakhotia 1972) and participation in writing a review on heterochromatin (Shah *et al.* 1973) enhanced

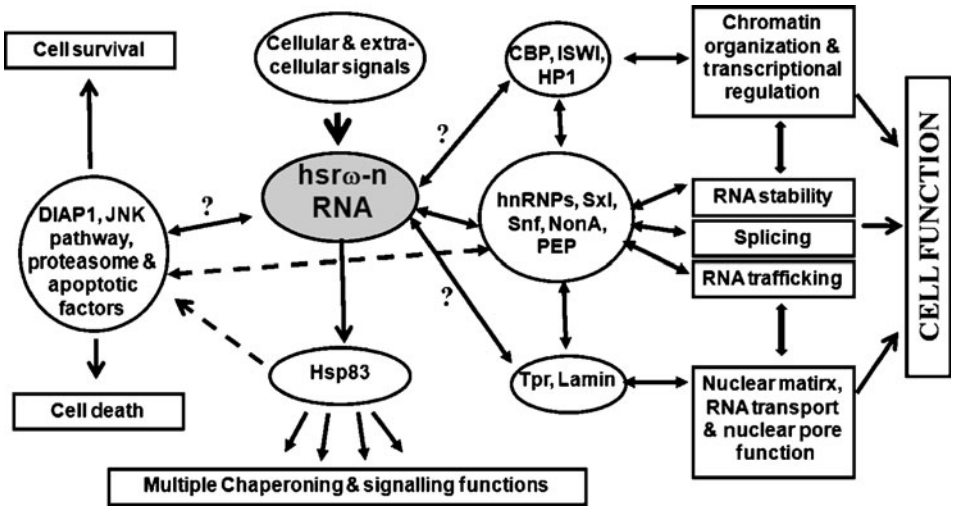



**Figure 7.** The hsrω-n RNA seems to function as a hub to coordinate several critical cellular networks (shown within individual ovals) that control different activities (noted in rectangles). The hsrω-n RNA is sensitive to a variety of cellular and extra-cellular signals (arrowhead) and relays the signal through its transcripts to multiple networks like (i) chromatin modulators, (ii) factors regulating nuclear RNA stability, processing and transport, (iii) nuclear pore and matrix components, (iv) Hsp83 and/or (v) cell survival or cell death pathways . A direct interaction of the hsrω-n transcripts with other networks is indicated through solid line arrows while indirect affects through other networks are shown with broken line arrows. The interrogation mark close to some solid line arrows indicates that a direct interaction with one or more components of the particular network is yet to be established.

my interest in the enigma of heterochromatin. That heterochromatin has definite functional significance was further strengthened when active transcription in the classical 'gene-deficient' β-heterochromatin was discovered for the first time (Lakhotia and Jacob 1974b). Thus, the increasingly clear evidence that the 93D locus does not code for a typical protein (Lakhotia and Mukherjee 1982; Peters *et al.* 1984; Ryseck *et al.* 1985; Garbe *et al.* 1986) was exciting rather than disappointing. During the 1995 session of the series of annual discussion meetings named 'TRendys', while reviewing the then known large non-coding RNAs, I predicted that they would become 'trendy' in coming years (Lakhotia 1996). It is indeed gratifying that this prediction has come true.

As is natural, many questions that arose during the course of our studies on the 93D puff have remained unanswered or even unaddressed. Several of these have been noted above and would be worth following up. More importantly, while our recent studies have mostly focused on the large nuclear hsrω-n transcripts, significance of the stable spliced-out intron and the 1.2 kb cytoplasmic hsrω-c transcript need more focused attention. The nearly 100% conservation of the intron–exon junction sequences, especially of the junction of the intron and second exon of *hsrω*, in all species of *Drosophila* whose genome sequences are available (Eshita Mutt and SC Lakhotia, unpublished) remains unexplained. This requires in-depth experimental analysis to understand if this conservation has something to do with regulation of the *hsrω* gene and/or with subtle differences in properties and activities of the hsrω-n1 (unspliced) and hsrω-n2 (spliced) transcripts. The mechanistic details of specific interactions of the various hsrω transcripts with the diverse variety of proteins can be more specifically addressed as novel experimental strategies become available.

As the years go by, it is expected that functions of the enormous diversity of large non-coding RNAs, which commonly exist in eukaryotes but currently remain largely unknown or unappreciated, will be understood soon so that the '93D puff' would no more be an exception but would be remembered as one of the 'pioneer' non-coding gene.

## Acknowledgements

I express my appreciation of the stimulating and enthusiastic work carried out by my former and present PhD students and of the various stimulating discussions with colleagues in the Cytogenetics Laboratory, all of which helped develop new ideas. The technical and other support by the different support staff in the laboratory has been invaluable. I thank Mary-Lou Pardue, Caroline Jolly and Davide F Corona for collaborating and sharing ideas/reagents. I thank H Saumweber for developing the remarkable collection of monoclonal antibodies against a large variety of nuclear proteins and making them available for our work and which were largely responsible for appreciation of the omega speckles. The generous and 'no-questions-asked' approach of the fly

community in sharing fly stocks and other reagents has been a great support. I also thank the various funding agencies in India (University Grants Commission, Board of Research in Nuclear Sciences, Indian National Science Academy, Council of Scientific and Industrial Research, Department of Science & Technology and Department of Biotechnology) for supporting my research on various topics at different times. I also thank the anonymous reviewers of this manuscript for their very useful comments/suggestions.

## References

Akanksha, Mallik M, Fatima R and Lakhotia SC 2008 The hsromega(05241) allele of the noncoding hsromega gene of *Drosophila melanogaster* is not responsible for male sterility as reported earlier. *J. Genet.* **87** 87–90

Anderson AR, Collinge JE, Hoffmann AA, Kellett M and Mckechnie SW 2003 Thermal tolerance trade-offs associated with the right arm of chromosome 3 and marked by the hsr-omega gene in *Drosophila melanogaster*. *Heredity* **90** 195–202

Arao Y, Kuriyama R, Kayama F and Kato S 2000 A nuclear matrix-associated factor, SAF-B, interacts with specific isoforms of AUF1/hnRNP D. *Arch. Biochem. Biophys.* **380** 228–236

Arya R, Mallik M and Lakhotia SC 2007 Heat shock genes – integrating cell survival and death. *J. Biosci.* **32** 595–610

Ashburner M 1967 Patterns of puffing activity in the salivary gland chromosomes of *Drosophila*. I. Autosomal puffing patterns in a laboratory stock of *Drosophila melanogaster*. *Chromosoma* **21** 398–428

Ashburner M 1970 Patterns of puffing activity in the salivary gland chromosomes of *Drosophila*. V. Responses to environmental treatments. *Chromosoma* **31** 356–376

Ashburner M and Bonner JJ 1979 The induction of gene activity in Drosophilia by heat shock. *Cell* **17** 241–254

Behnel HJ 1982 Comparative study of protein synthesis and heat shock puffing activity in *Drosophila* salivary glands treated with chloramphenicol. *Exp. Cell. Res.* **142**

Belew K and Brady T 1981 Induction of tyrosine aminotransferase by pyridoxine in *Drosophila hydei*. *Chromosoma* **82** 99–106

Bendena WG, Ayme-Southgate A, Garbe JC and Pardue ML 1991 Expression of heat-shock locus hsr-omega in nonstressed cells during development in *Drosophila melanogaster*. *Dev. Biol.* **144** 65–77

Bendena WG, Fini ME, Garbe JC, Kidder GM, Lakhotia SC and Pardue ML 1989a hsrω: A different sort of heat shock locus; in *Stress-induced proteins* (eds) ML Pardue, J Ferimisco and S Lindquist (Alan R Liss, Inc) pp 3–14

Bendena WG, Garbe JC, Traverse KL, Lakhotia SC and Pardue ML 1989b Multiple inducers of the *Drosophila* heat shock locus 93D (*hsrω*): inducer-specific pattern of the three transcripts. *J. Cell Biol*. **108** 2017–2028

Berendes HD and Beermann W 1969 Biochemical activity of interphase chromosomes (polytene chromosomes); in *Handbook of molecular cytogenetics* (ed) Lima-de-Faria (Amsterdam: London North Holland Publishing Company) pp 501–519

Bonner JJ and Pardue ML 1976 The effect of heat shock on RNA synthesis in *Drosophila* tissues. *Cell* **8** 43–50

Brady T and Belew K 1981 Pyridoxine induced puffing (II-48C) and synthesis of a 40 KD protein in *Drosophila hydei* salivary glands. *Chromosoma* **82** 89–98

Brand AH and Perrimon N 1993 Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118** 401–415

Buchenau P, Saumweber H and Arndt-Jovin DJ 1997 The dynamic nuclear redistribution of an hnRNP K-homologous protein during *Drosophila* embryo development and heat shock. Flexibility of transcription sites in vivo. *J. Cell Biol.* **137** 291–303

Burma PK and Lakhotia SC 1984 Cytological identity of 93D-like and 87C-like heat shock loci in *Drosophila pseudoobscura*. *Indian J. Exp. Biol.* **22** 577–580

Burma PK and Lakhotia SC 1986 Expression of 93D heat shock puff of *Drosophila melanogaster* in deficiency genotype and its influence on activity of the 87C puff. *Chromosoma* **94** 273–278

Carmona MJ, Morcillo G, Galler R, Martinez-Salas E, de la Campa AG, Diez JL and Edstrom JE 1985 Cloning and molecular characterization of a telomeric sequence from a temperature-induced Balbiani ring. *Chromosoma* **92** 108–115

Collinge JE, Anderson AR, Weeks AR, Johnson TK and McKechnie SW 2008 Latitudinal and cold-tolerance variation associate with DNA repeat-number variation in the hsr-omega RNA gene of *Drosophila melanogaster*. *Heredity* **101** 260–270

Compton JL and McCarthy BJ 1978 Induction of the *Drosophila* heat shock response in isolated polytene nuclei. *Cell* **14** 191–201

Corona DF, Siriaco G, Armstrong JA, Snarskaya N, McClymont SA, Scott MP and Tamkun JW 2007 ISWI regulates higher-order chromatin structure and histone H1 assembly in vivo. *PLoS Biol.* **5** e232

Cutforth T and Rubin GM 1994 Mutations in Hsp83 and cdc37 impair signaling by the sevenless receptor tyrosine kinase in *Drosophila. Cell* **7** 1027–1036

Dangli A and Bautz EK 1983 Differential distribution of nonhistone proteins from polytene chromosomes of *Drosophila melanogaster* after heat shock. *Chromosoma* **88** 201–207

Dangli A, Grond C, Kloetzel P and Bautz EK 1983 Heat-shock puff 93 D from *Drosophila melanogaster*: accumulation of a RNP-specific antigen associated with giant particles of possible storage function. *EMBO J.* **2** 1747–1751

Denegri M, Chiodi I, Corioni M, Cobianchi F, Riva S and Biamonti G 2001 Stress-induced nuclear bodies are sites of accumulation of pre-mRNA processing factors. *Mol. Biol. Cell* **12** 3502–3514

Derksen J 1975 The submicroscopic structure of synthetically active units in a puff of *Drosophila hydei* giant chromosomes. *Chromosoma* **50** 45–52

Derksen J, Berendes HD and Willart E 1973 Production and release of a locus-specific ribonucleoprotein product in polytene nuclei of *Drosophila hydei*. *J. Cell Biol.* **59** 661–668

Doolittle WF and Sapienza C 1980 Selfish genes, the phenotype paradigm and genome evolution. *Nature (London)* **284** 601–603

Dustin P 1978 *Microtubules* (Berlin, Heidelberg, New York: Springer-Verlag)

Edstrom JE and Beermann W 1962 The base composition of nucleic acids in chromosomes, puffs, nucleoli, and cytoplasm of *Chironomus* salivary gland cells. *J. Cell Biol.* **14** 371–379

Fernandez-Funez P, Nino-Rosales ML, de Gouyon B, She WC, Luchak JM, Martinez P, Turiegano E, Benito J, *et al.* 2000 Identification of genes that modify ataxin-1-induced neurodegeneration. *Nature (London)* **408** 101–106

Fini ME, Bendena WG and Pardue ML 1989 Unusual behavior of the cytoplasmic transcript of hsr omega: an abundant, stress-inducible RNA that is translated but yields no detectable protein product. *J. Cell Biol.* **108** 2045–2057

Garbe JC and Pardue ML 1986 Heat-shock locus 93D of *Drosophila melanogaster:* a spliced RNA most strongly conserved in the intron sequence. *Proc. Natl. Acad. Sci. USA* **83** 1812–1816

Garbe JC, Bendena WG and Pardue ML 1989 Sequence evolution of the *Drosophila* heat shock locus hsr omega. I. The nonrepeated portion of the gene. *Genetics* **122** 403–415

Garbe JC, Bendena WG, Alfano M and Pardue ML 1986 A *Drosophila* heat shock locus with a rapidly diverging sequence but a conserved structure. *J. Biol. Chem.* **261** 16889–16894

Grossbach U 1969 Chromosome activity and biochemical cell differentiation in the salivary glands of Camptochironomus. *Chromosoma* **28** 136–187

Gubenko IS and Baricheva EM 1979 *Drosophila virilis* puffs induced by temperature and other environmental factors. *Genetika (USSR)* **15** 1399–1414

Han SP, Tang YH and Smith R 2010 Functional diversity of the hnRNPs: past, present and perspectives. *Biochem. J.* **430** 379–392

Haynes SR, Johnson D, Raychaudhuri G and Beyer AL 1991 The *Drosophila* Hrb87F gene encodes a new member of the A and B hnRNP protein group. *Nucleic Acids Res.* **19** 25–31

Henikoff S 1980 A more conventional view of the 'ebony' gene. *Dros. Lnf. Serv.* **55** 61

Hochstrasser M 1987 Chromosome structure in four wild-type polytene tissues of *Drosophila melanogaster*. The 87A and 87C heat shock loci are induced unequally in the midgut in a manner dependent on growth temperature. *Chromosoma* **95** 197–208

Hogan NC, Traverse KL, Sullivan DE and Pardue ML 1994 The nucleus-limited Hsr-omega-n transcript is a polyadenylated RNA with a regulated intranuclear turnover. *J. Cell Biol.* **125** 21–30

Hovemann B, Walldorf U and Ryseck RP 1986 Heat-shock locus 93D of *Drosophila melanogaster*: An RNA with limited coding capacity accumulates precursor transcripts after heat shock. *Mol. Gen. Genet.* **204** 334–340

Ji Y and Tulin AV 2009 Poly(ADP-ribosyl)ation of heterogeneous nuclear ribonucleoproteins modulates splicing. *Nucleic Acids Res.* **37** 3501–3513

Johnson TK, Carrington LB, Hallas RJ and McKechnie SW 2009 Protein synthesis rates in *Drosophila* associate with levels of the hsr-omega nuclear transcript. *Cell Stress Chaperones* **14** 569–577

Jolly C and Lakhotia SC 2006 Human sat III and *Drosophila* hsr omega transcripts: a common paradigm for regulation of nuclear RNA processing in stressed cells; *Nucleic Acids Res.* **34** 5508–5514

Jolly C and Morimoto RI 1999 Stress and the cell nucleus: dynamics of gene expression and structural reorganization. *Gene Expr.* **7** 261–270

Jolly C, Metz A, Govin J, Vigneron M, Turner BM, Khochbin S and Vourc'h C 2004 Stress-induced transcription of satellite III repeats. *J. Cell Biol.* **164** 25–33

Kar Chowdhury D and Lakhotia SC 1986 Different effects of 93D on 87C heat shock puff activity in *Drosophila melanogaster* and *D. simulans*. *Chromosoma* **94**: 279–284

Lakhotia SC 1970 Gene physiological studies on dosage compensation in *Drosophila*. PhD Thesis, University of Calcutta, Kolkata

Lakhotia SC 1974 EM autoradiographic studies on polytene nuclei of *Drosophila melanogaster*. 3. Localisation of non-replicating chromatin in the chromocentre heterochromatin. *Chromosoma* **46** 145–159

Lakhotia SC 1987 The 93D heat shock locus in *Drosophila* - a review. *J. Genet.* **66** 139–157

Lakhotia SC 1989 The 93D heat shock locus of *Drosophila melanogaster*: modulation by genetic and developmental factors. *Genome* **31** 677–683

Lakhotia SC 1996 RNA polymerase II dependent genes that do not code for protein. *Indian J. Biochem. Biophys.* **33** 93–102

Lakhotia SC 1999 Non-coding RNAs: versatile roles in cell regulation. *Curr. Sci.* **77** 479–480

Lakhotia SC 2003 The non-coding, developmentally active and stress inducible hsrω gene of *Drosophila melanogaster* integrates post-transcriptional processing of other nuclear transcripts; in *Noncoding RNAs: molecular biology and molecular medicine* (eds) J Barciszweski and VA Erdmann (New York: Kluwer Academic/Plenum Publishers) pp 202–219

Lakhotia SC and Jacob J 1974a Electron microscopic autoradiographic studies on polytene nuclei of *Drosophila melanogaster*: Part I. Replication and its relationship with nuclear membrane. *Indian J. Exp. Biol.* **12** 389–394

Lakhotia SC and Jacob J 1974b EM autoradiographic studies on polytene nuclei of *Drosophila melanogaster*. II. Organization and transcriptive activity of the chromocentre. *Exp. Cell Res.* **86** 253–263

Lakhotia SC and Mukherjee AS 1969 Chromosomal basis of dosage compensation in *Drosophila*. I. Cellular autonomy of hyperactivity of the male X-chromosome in salivary glands and sex differentiation. *Genet. Res.* **14** 137–150

Lakhotia SC and Mukherjee AS 1970a Chromosomal basis of dosage compensation in *Drosophila*. 3. Early completion of replication by the polytene X-chromosome in male: further evidence and its implications. *J. Cell Biol.* **47** 18–33

Lakhotia SC and Mukherjee AS 1970b Activation of a specific puff by benzamide in *D. melanogaster*. *Dros. Inf. Serv.* **45** 108

Lakhotia SC and Mukherjee T 1980 Specific activation of puff 93D of *Drosophila melanogaster* by benzamide and the effect of benzamide treatment on the heat shock induced puffing activity. *Chromosoma* **81** 125–136

Lakhotia SC and Mukherjee T 1982 Absence of novel translation products in relation to induced activity of the 93D puff in *Drosophila melanogaster*. *Chromosoma* **85** 369–374

Lakhotia SC and Mukherjee T 1984 Specific induction of the 93D puff in polytene nuclei of *Drosophila melanogaster* by colchicine. *Indian J. Exp. Biol.* **22** 67–70

Lakhotia SC and Mutsuddi M 1996 Heat shock but not benzamide and colchicine response elements are present within the −844 bp upstream region of the *hsr omega* gene of *Drosophila melanogaster*. *J. Biosci.* **21** 235–246

Lakhotia SC and Ray P 1996 *hsp*83 mutation is a dominant enhancer of lethality associated with absence of the non-protein coding *hsr omega* locus in *Drosophila melanogaster*. *J. Biosci.* **21** 207–219

Lakhotia SC and Sharma A 1995 RNA metabolism *in situ* at the 93D heat shock locus in polytene nuclei of *Drosophila melanogaster* after various treatments *Chromosome Res.* **3** 151–161

Lakhotia SC and Sharma A 1996 The 93D (*hsrω*) locus of *Drosophila*: non-coding gene with house-keeping functions. *Genetica* **97** 339–348

Lakhotia SC and Singh AK 1982 Conservation of the 93D puff of *Drosophila melanogaster* in different species of *Drosophila*. *Chromosoma* **86** 265–278

Lakhotia SC and Singh AK 1985 Non-inducibility of the 93D heat shock puff in cold-reared larvae of *Drosophila melanogaster*. *Chromosoma* **92** 48–54

Lakhotia SC and Tapadia MG 1998 Genetic mapping of the amide response element/s of the *hsrω* locus of *Drosophila melanogaster*. *Chromosoma* **107** 127–135

Lakhotia SC, Kar Chowdhuri D and Burma PK 1990 Mutations affecting β-alanine metabolism influence inducibility of the 93D puff by heat shock in *Drosophila melanogaster*. *Chromosoma* **99** 296–305

Lakhotia SC, Rajendra TK and Prasanth KV 2001 Developmental regulation and complex organization of the promoter of the non-coding hsrω gene of *Drosophila melanogaster*. *J. Biosci.* **26** 25–38

Lakhotia SC, Ray P, Rajendra TK and Prasanth KV 1999 The non-coding transcripts of hsr-omega gene in *Drosophila*: do they regulate trafficking and availability of nuclear RNA-processing factors? *Curr. Sci.* **77** 553–563

Lakomek HJ, Plomann M, Specker C and Schwochau M 1991 Ankylosing spondylitis: an autoimmune disease? *Ann. Rheum. Dis.* **50** 776–781

Leenders HJ, Derksen J, Mass PMJM and Berendes HD 1973 Selective induction of a giant puff in *Drosophila hydei* by vitamin $B_6$ and derivatives. *Chromosoma* **41** 447–460

Lengyel JA, Ransom LJ, Graham ML and Pardue ML 1980 Transcription and metabolism of RNA from the *Drosophila melanogaster* heat shock puff site 93D. *Chromosoma* **80** 237–252

Lewis M, Helmsing PJ and Ashburner M 1975 Parallel changes in puffing activity and patterns of protein synthesis in salivary glands of *Drosophila*. *Proc. Natl. Acad. Sci. USA* **72** 3604–3608

Lindell TJ, Weinberg F, Morris PW, Roeder RG and Rutter WJ 1970 Specific inhibition of nuclear RNA polymerase II by alpha-amanitin. *Science* **170** 447–449

Livak KJ, Freund R, Schweber M, Wensink PC and Meselson M 1978 Sequence organization and transcription at two heat shock loci in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **75** 5613–5617

Lyon MF 1961 Gene action in the X-chromosome of the mouse (*Mus musculus L.*). *Nature (London)* **190** 372–373

Mallik M and Lakhotia SC 2007 Noncoding DNA is not "junk" but a necessity for origin and evolution of biological complexity. *Proc. Natl. Acad. Sci. India Spl. Iss.* **77(B)** 43–50

Mallik M and Lakhotia SC 2009a RNAi for the large non-coding hsromega transcripts suppresses polyglutamine pathogenesis in *Drosophila* models. *RNA Biol.* **6** 464–478

Mallik M and Lakhotia SC 2009b The developmentally active and stress-inducible non-coding hsrω gene is a novel regulator of apoptosis in *Drosophila*. *Genetics* **183** 831–852

Mallik M and Lakhotia SC 2010a Improved activities of CBP, hnRNPs and proteasome following down regulation of non-coding hsromega transcripts help suppress polyQ pathogenesis in fly models. *Genetics* **184** 927–945

Mallik M and Lakhotia SC 2010b Modifiers and mechanisms of multi-system polyglutamine neurodegenerative disorders: lessons from fly models. *J. Genet.* **89** 497–526

Mallik M and Lakhotia SC 2011 Misexpression of the developmentally active and stress-inducible non-coding hsrω gene in *Drosophila* has pleiotropic consequences. *J. Biosci.* **36** 265–280

Mattern KA, van der Kraan I, Schul W, de Jong L and van Driel R 1999 Spatial organization of four hnRNP proteins in relation to sites of transcription, to nuclear speckles, and to each other in interphase nuclei and nuclear matrices of HeLa cells. *Exp. Cell Res.* **246** 461–470

McColl G and McKechnie SW 1999 The *Drosophila* heat shock hsr-omega gene: an allele frequency cline detected by quantitative PCR. *Mol. Biol. Evol.* **16** 1568–1574

McKechnie SW, Halford MM, McColl G and Hoffmann AA 1998 Both allelic variation and expression of nuclear and cytoplasmic transcripts of Hsr-omega are closely associated with thermal phenotype in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **95** 2423–2428

Mohler J and Pardue ML 1982 Deficiency mapping of' the 93D heat shock locus in *Drosophila*. *Chromosoma* **86** 457–467

Mohler J and Pardue ML 1984 Mutational analysis of the region surrounding the 93D heat shock locus of *Drosophila melanogaster*. *Genetics* **106** 249–265

Morcillo G, Diez JL, Carbajal M E and Tanguay R M 1993 Hsp90 associates with specific heat shock puffs (*hsrω*) in polytene chromosomes of *Drosophila* and *Chironomus*. *Chromosoma* **102** 648–659

Morcillo G, Diez JL and Botella LM 1994 Heat shock activation of telomeric sequences in different tissues of *Chironomus thummi*. *Exp. Cell Res.* **211** 163–167

Mukherjee AS and Beermann W 1965 Synthesis of ribonucleic acid by the X-chromosomes of *Drosophila melanogaster* and the problem of dosage compensation. *Nature (London)* **207** 785–786

Mukherjee T and Lakhotia SC 1979 $^3$H-uridine incorporation in the puff 93D and in chromocentric heterochromatin of heat shocked salivary glands of *Drosophila melanogaster*. *Chromosoma* **74** 75–82

Mukherjee T and Lakhotia SC 1981 Specific induction of the 93D puff in *Drosophila melanogaster* by a homogenate of heat shocked larval salivary glands. *Indian J. Exp. Biol.* **19** 1–4

Mukherjee T and Lakhotia SC 1982 Heat shock puff activity in salivary glands of *Drosophila melanogaster* larvae during

recovery from anoxia at two different temperatures. *Indian J. Exp. Biol.* **20** 437–439

Muller HJ and Kaplan WD 1966 The dosage compensation of *Drosophila* and mammals as showing the accuracy of the normal type. *Genet. Res.* **8** 41–59

Mutsuddi M and Lakhotia SC 1995 Spatial expression of the hsr-omega (93D) gene in different tissues of *Drosophila melanogaster* and identification of promoter elements controlling its developmental expression. *Dev. Genet.* **17** 303–311

Nath BB and Lakhotia SC 1991 Search for a *Drosophila*-93D-like locus in *Chironomus* and *Anopheles*. *Cytobios* **65** 7–13

Onorati MC, Lazzaro S, Mallik M, Ingrassia AMR, Singh AK, Chaturvedi DP, Lakhotia SC and Corona DFV 2011 The ISWI chromatin remodeler organizes the hsrω ncRNA-containing omega speckle nuclear compartments. *PLoS Genet.* **7** e1002096. doi:10.1371/journal.pgen.1002096

Orgel LE and Crick FH 1980 Selfish DNA: the ultimate parasite. *Nature (London)* **284** 604–607

Pardue ML, Bendena WG, Fini ME, Garbe JC, Hogan NC and Traverse KL 1990 *Hsr-omega*, A novel gene encoded by a *Drosophila* heat shock puff. *Biol. Bull.* **179** 77–86

Perry RP and Kelley DE 1968 Persistent synthesis of 5S RNA when production of 28S and 18S ribosomal RNA is inhibited by low doses of actinomycin D. *J. Cell Physiol.* **72** 235–246

Peters FP, Lubsen NH and Sondermeijer PJ 1980 Rapid sequence divergence in a heat shock locus of *Drosophila*. *Chromosoma* **81** 271–280

Peters FP, Lubsen NH, Walldorf U, Moormann RJ and Hovemann B 1984 The unusual structure of heat shock locus 2-48B in *Drosophila hydei*. *Mol. Gen. Genet.* **197** 392–398

Piacentini L, Fanti L, Berloco M, Perrini B and Pimpinelli S 2003 Heterochromatin protein 1 (HP1) is associated with induced gene expression in *Drosophila* euchromatin. *J. Cell Biol.* **161** 707–714

Prasanth KV and Spector DL 2007 Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. *Gene. Dev.* **21** 11–42

Prasanth KV, Rajendra TK, Lal AK and Lakhotia SC 2000 Omega speckles - a novel class of nuclear speckles containing hnRNPs associated with noncoding hsr-omega RNA in *Drosophila*. *J. Cell Sci.* **113** 3485–3497

Rajendra TK, Prasanth KV and Lakhotia SC 2001 Male sterility associated with over-expression of the non-coding *hsrω* gene in cyst cells of testis of *Drosophila melanogaster*. *J. Genet.* **80** 97–110

Rao SRV and Lakhotia SC 1972 Chromosomes of Rattus (Rattus) blanfordi (Thomas), Muridae, Rodentia. *J. Heredity* **63** 44–47

Ray P 1997 Studies on interaction of the 93D (hsr-omega) locus with other genes during development of *D. melanogaster.* PhD Thesis, Banaras Hindu University, Varanasi

Ray P and Lakhotia SC 1998 Interaction of the non-protein-coding developmental and stress-inducible *hsrω* gene with *Ras* genes of *Drosophila melanogaster*. *J. Biosci.* **23** 377–386

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

Ryseck RP, Walldorf U and Hovemann B 1985 Two major RNA products are transcribed from heat-shock locus 93D of *Drosophila melanogaster*. *Chromosoma* **93** 17–20

Ryseck RP, Walldorf U, Hoffmann T and Hovemann B 1987 Heat shock loci 93D of *Drosophila melanogaster* and 48B of *Drosophila hydei* exhibit a common structural and transcriptional pattern. *Nucleic Acids Res.* **15** 3317–3333

Samuels ME, Bopp D, Colvin RA, Roscigno RF, Garcia-Blanco MA and Schedl P 1994 RNA binding by Sxl proteins in vitro and in vivo. *Mol. Cell Biol.* **14** 4975–4990

Santa-Cruz MC, Morcillo G and Diez JL 1984 Ultrastructure of a temperature-induced Balbiani ring in Chironomus thummi. *Biol. Cell* **52** 205–211

Saumweber H, Symmons P, Kabisch R, Will H and Bonhoeffer F 1980 Monoclonal antibodies against chromosomal proteins of *Drosophila melanogaster*: establishment of antibody producing cell lines and partial characterization of corresponding antigens. *Chromosoma* **80** 253–275

Scalenghe F and Ritossa F 1977 The puff inducible in region 93D is responsible for the synthesis of the major 'heat-shock' polypeptide in *Drosophila melanogaster*. *Chromosoma* **63** 317–327

Sengupta S 2005 Studies on a novel gene interacting with hsrω and their roles as modifiers in polyglutamine induced neurodegeneration in *Drosophila melanogaster*. PhD Thesis, Banaras Hindu University, Varanasi

Sengupta S and Lakhotia SC 2006 Altered expression of the noncoding hsrω gene enhances poly-Q-induced neurotoxicity in *Drosophila*. *RNA Biol.* **3** 28–35

Shah VC, Lakhotia SC and Rao SRV 1973 Nature of heterochromatin. *J. Sci. Industrial Res.* **32** 467–480

Sharma A and Lakhotia SC 1995 *In situ* quantification of hsp70 and alpha-beta transcripts at 87A and 87C loci in relation to hsr-omega gene activity in polytene cells of *Drosophila melanogaster*. *Chromosome Res.* **3** 386–393

Sims JL, Berger SJ and Berger NA 1983 Poly(ADP-ribose) Polymerase inhibitors preserve nicotinamide adenine dinucleotide and adenosine 5'-triphosphate pools in DNA-damaged cells: mechanism of stimulation of unscheduled DNA synthesis. *Biochemistry* **22** 5188–5194

Singh AK and Lakhotia SC 1983 Further observations on inducibility of 93D puff of *Drosophila melanogaster* by homogenate of heat shocked cells. *Indian J. Exp. Biol.* **21** 363–366

Singh AK and Lakhotia SC 1984 Lack of effect of microtubules positions on the 93D or 93D-like heat shock puffs in *Drosophila*. *Indian J. Exp. Biol.* **22** 569–576

Sirlin JL and Jacob J 1964 Sequential and reversible inhibition of synthesis of ribonucleic acid in the nucleolus and chromosomes: effect of benzamide and substituted benzimidazoles on dipteran salivary glands. *Nature (London)* **204** 545–547

Spector DL, Fu XD and Maniatis T 1991 Associations between distinct pre-mRNA splicing components and the cell nucleus. *EMBO J.* **10** 3467–3481.

Spradling A, Pardue ML and Penman S 1977 Messanger RNA in heat-shocked *Drosophila* cells. *J. Mol. Biol.* **109** 559–587

Spruill WA, Hurwitz DR, Lucchesi JC and Steiner AL 1978 Association of cyclic GMP with gene expression of polytene chromosomes of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **75** 1480–1484

Srikrishna S 2008 Studies on expression of hsrω non-coding RNA and its interacting proteins during oogenesis in *Drosophila melanogaster*. PhD Thesis, Banaras Hindu University, Varanasi

Srivastava JP and Bangia KK 1985 Specific induction of puff 93D in the polytene chromosomes of salivary glands of *Drosophila melanogaster* by paracetamol. *J. Curr. Biosci.* **2** 48–50

Tapadia MG and Lakhotia SC 1997 Specific induction of the hsr omega locus of *Drosophila melanogaster* by amides. *Chromosome Res.* **5** 359–362

Tissieres A, Mitchell HK and Tracy UM 1974 Protein synthesis in salivary glands of *Drosophila melanogaster*: relation to chromosome puffs. *J. Mol. Biol.* **84** 389–398

Walldorf U, Richter S, Ryseck RP, Steller H, Edstrom JE, Bautz EK and Hovemann B 1984 Cloning of heat-shock locus 93D from *Drosophila melanogaster*. *EMBO J.* **3** 2499–2504

Westwood JT, Clos J and Wu C 1991 Stress-induced oligomerization and chromosomal relocalization of heat-shock factor. *Nature (London)* **353** 822–827

Wright TR 1987 The genetics of biogenic amine metabolism, sclerotization, and melanization in *Drosophila melanogaster*. *Adv. Genet.* **24** 127–222

Zimowska G and Paddy MR 2002 Structures and dynamics of *Drosophila* Tpr inconsistent with a static, filamentous structure. *Exp. Cell Res.* **276** 223–232

*e*Publication: 08 July 2011