Geology and Computer Sciences. S. C. Sarkar (Jadavpur University) threw light on different aspects of metallogeny throughout the Indian subcontinent. A. K. Gupta (Allahabad University) delivered lectures on different processes of magma generation and the genesis of Komatiites. G. S. Nurulla (DDG, GSI, NW region) talked on seismic hazards and their evaluation. K. R. Gupta (DST) presented a talk on the funding of research projects and the role of DST in it. A. Dey (Calcutta University) delivered his talk on the different aspects of basalt petrogenesis with special reference to the Deccan basalts. K. K. Sharma (Wadia Institute of Himalayan Geology, Dehradun) discussed the nature of crustal growth throughout geological past and correlations of different Indian shields. R. S. Sharma (emeritus professor) taught various processes responsible for melt generation during the process of metamorphism. R. N. Singh (CMMACS, Bangalore) dwelt at length on the thermal modelling of the lower continental crust. S. K. Tandon (Delhi University) talked on how to understand the past climates with the help of different proxy indicators. J. P. Srivastava (Delhi University) presented a series of lectures on the geochemistry of the Deccan basalts. C. S. Dubey (Delhi University) demonstrated how to extract geological informations of our own interest from the Internet.

Each participant was given a task to prepare a representative database related to his own research work, which all the participants completed with support from faculty members and the Central Library staff.

As this was the last school of the series, there were discussions on future planning for running such schools. A suggestion made was to organize a workshop next winter for the participants to present their progress. Such schools have been very helpful for research workers. They are provided a platform to discuss their problems and to arrive at solutions with the help of fellow-workers and scientists from the different parts of the country.

Vachaspati Pandey, Department of Earth and Planetary Sciences, Allahabad University, Allahabad 211 002, India; **Rajesh K. Singh,** Department of Geology Banaras Hindu University, Varanasi; **Divya Prakash,** Department of Geology, Banaras Hindu University, Varanasi, India

RESEARCH NEWS

Getting at the core of the colinearity of *Hox* genes

Rakesh K. Mishra

Discovery of homeotic genes was a major step towards understanding the genetic control of development. Mutations in homeotic genes convert one body part into another; for example, *bithorax* mutation of *Drosophila* transforms part of the haltare towards wing. Pioneering work of Ed Lewis showed that bithorax locus is complex of several distinct genetic elements and, most importantly, that bithorax mutations map in an order that corresponds to the order of the body parts that are affected by these mutations[1]. This colinearity – correspondence of order of genes on the chromosome with the order of body parts that are under the control of these genes – is conserved up to vertebrates.

All homeotic genes were later found to have a 60 a.a DNA binding domain called the 'homeobox'. In fact, most of the homeotic genes outside *Drosophila* have been identified on the basis of this homology. Although the 'colinearity' of homeotic genes was first discovered in *Drosophila*, the fly seems to be an exception to the rule! Unlike other organisms, homeotic genes of *Drosophila* exist in two clusters (Figure 1 *a*): the Antennapedia complex (ANT-C) that controls development of the head and half of the thorax and the bithorax complex (BX-C) that controls development of the rest of the thorax and abdomen. In vertebrates, the corresponding homeotic genes exist in one cluster and there are four such clusters per haploid genome (Figure 1 *b*). Moreover, genetic experiments showed that colinear arrangement within the BX-C is not essential; for example, both *Ubx* and *Abd-B* can function even when displaced to different locations in the genome. Another difference between fly and vertebrate *Hox* complexes is that the complexes in vertebrates are compact, 80 to 120 kb, while in *Drosophila* the BX-C alone spans more than 300 kb, 95% of which constitutes *cis*-regulatory elements. The colinearity in the BX-C extends to this vast *cis*-regulatory region also. Such a striking colinearity in homeotic gene clusters observed in all animals with anterior–posterior axial polarity could not be an accident. Unlike in *Drosophila* where determination of entire anterior–posterior axis takes place simultaneously, in vertebrates there is a temporal order for this – anterior parts are determined earlier and posterior parts later. The *Hox* genes are, accordingly, activated in this temporal order – the spatial colinearity is accompanied by a 'temporal colinearity'[2,3]. Why does this organization of homeotic genes persist? In other words, what is the mechanism that utilizes this organization in the transcriptional regulation of homeotic genes?

Several observations suggest that the mechanisms involved in regulation of homeotic genes are conserved. Transgenic animals, both flies and mice, carrying a reporter gene driven by a *cis*-regulatory element from the *Hox* cluster often show an expression pattern which is consistent with the property of the regulatory element in question[4–8]. Also, when a reporter gene is inserted with the complex, its expression is driven by the *cis*-regulatory elements of complex in

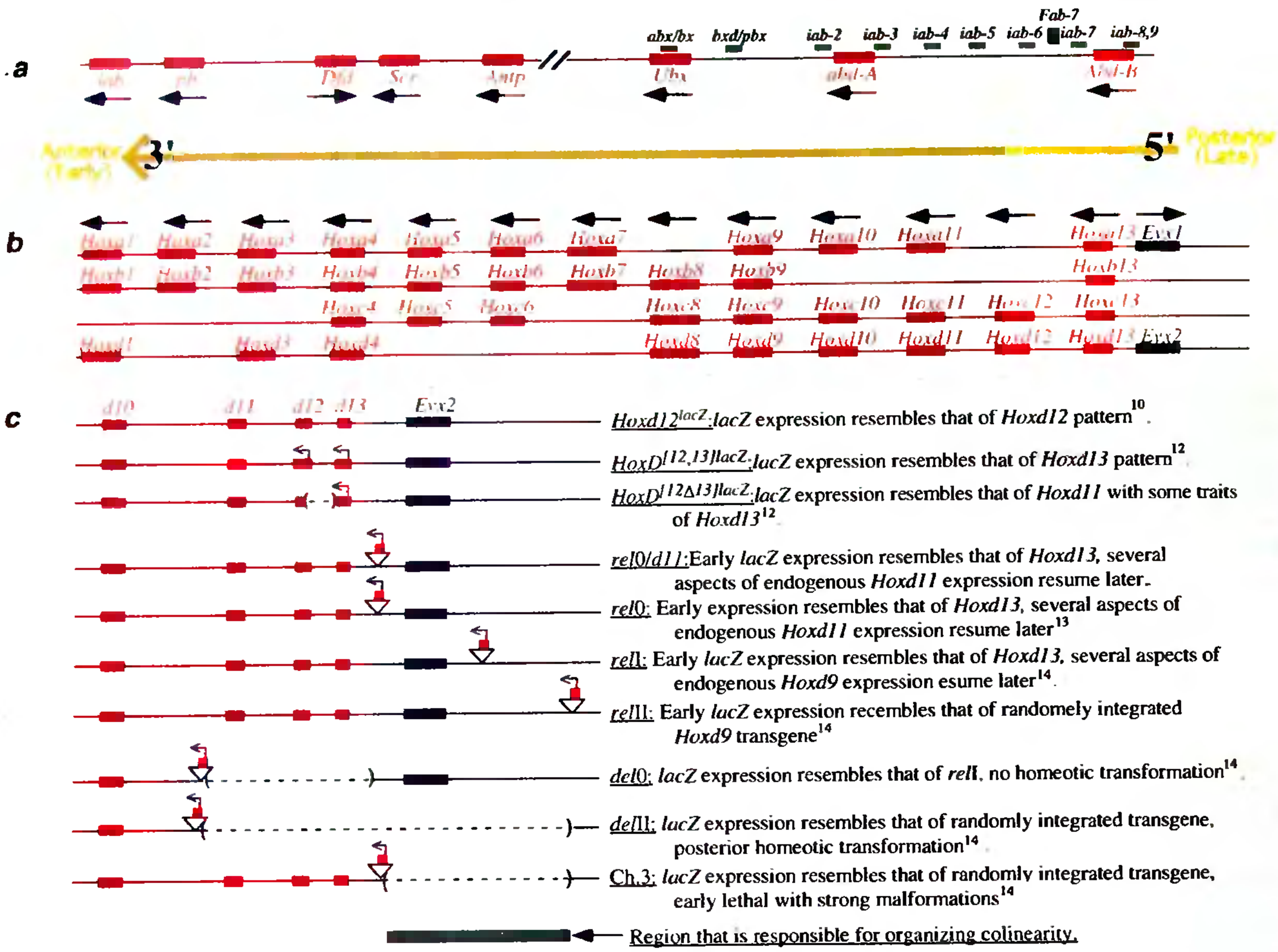

Figure 1. Organization of homeotic gene complexes (not drawn to scale). As indicated by the big yellow arrow, 3′ to 5′ arrangement of genes also reflects the order in which they are expressed along the anterior–posterior body axis. In addition to this spatial colinearity, in case of vertebrates there is a temporal colinearity – the 3′ to 5′ arrangement of genes also reflect the temporal order in which they are activated during development. Vertebrate genes at the 3′ end are also more sensitive to stimulation by retenoic acid (not shown). ***a***, Homeotic gene complex of *Drosophila melanogaster*. Arrows below the homeotic genes (red rectangles) indicate the direction of transcription. The complex is spilt into the Antennapedia complex (left) with five homeotic genes spread over 350 kb and the bithorax complex (right) with three homeotic genes spread over 330 kb. There are few non-homeotic homeobox genes and non-homeobox genes within these complexes (not shown) that are not subject to the colinearity rule. Some of the *cis*-regulatory elements (*abx/bx* to *iab-8,9*) are drawn in green rectangles. The spatial colinearity described for the homeotic genes also extends to these regulatory elements. ***b***, Four homeotic gene clusters of mouse, *HoxA*, *HoxB*, *HoxC* and *HoxD* containing a total of 39 homeotic genes. Unlike *Drosophila* these complexes are more compact, devoid of non-homeotic genes within a complex and all the genes are transcribed in the same direction (shown by black arrows). ***c***, Mutations and transgenes within *HoxD* complex discussed in the text. Homeotic genes *Hoxd-10* to *Hoxd-13* and the *Evx-2* are shown. Bent arrows indicate *lacZ* fusions and its direction of transcription. Deletions are shown as dashed lines within brackets. *lacZ* fusions with *Hoxd-11* for *rel0/d11* and *Hoxd-9* for the rest were used to generate relocation and deletion alleles. The DNA region that is responsible for organizing the spatial and temporal linearity, as inferred from the analysis of these relocations and deletions, is shown at the bottom.

the vicinity of the site of insertion. P element insertions in the BX-C show a precise expression pattern of the reporter gene that reflects the site of insertion in the complex[9]. Similarly, expression of *lacZ* in *Hoxd-12*^{lacZ} mice (Figure 1 *c*) resembles that of *Hoxd-12* gene[10]. This suggests that the domains of spatial information are organized in a colinear fashion in the *Hox* complexes. In *Drosophila* mutations have been recovered that delete the boundary elements separating such domains[11]. A relay element that might work in a similar way as the boundary elements of bithorax complex has been identified in mouse *HoxD* cluster[12]. This is shown by expression patterns of *Hoxd-12*^{lacZ}, *Hoxd*$^{[12,13]lacZ}$ and *Hoxd*$^{[12\Delta13]lacZ}$ alleles, Figure 1 *c*. While in the first two alleles expression patterns resemble those of *Hoxd-12* and *Hoxd-13*, respectively, *lacZ* expression in the *Hoxd*$^{[12\Delta13]lacZ}$ resembles that of *Hoxd-11*. This anteriorization of the reporter gene expression suggests that the deleted region demarcates two domains in different states of activity.

What happens when *cis*-regulatory region or *Hox* gene from one part of the complex is translocated to another part of the same complex? While such experiments are technically very difficult to do in fly, the answer comes from the work of Duboule and colleagues on *HoxD* complex of mouse. When relocated within the *HoxD* complex, expression of *Hoxd-9* and *Hoxd-11* genes fused with *lacZ* reporter is under the control of the site of relocation, *rel0/d11* and *rel*0 alleles in Figure 1 *c* (ref. 13). This suggested that although the immediate neighbouring regulatory regions control a *Hox* gene expression in randomly located transgenes (as they often mimic wild type expression pattern), within the complex there is a higher order regulatory mechanism that is responsible for the 'colinearity' and can override other regulatory elements.

In a recent report, Kondo and Duboule extend this observation to ask how far the colinear regulation of the *HoxD* cluster extends[14]. Furthermore, they take a significant step towards understanding the basis of 'colinearity' by mapping a DNA region 5′ to the *Hoxd-13* that is responsible for it and by showing that when this region is deleted the colinearity of the entire *HoxD* cluster is lost. As shown in Figure 1 *c*, while the expression of relocated *Hoxd9*^{lacZ} fusion gene in *rel*0 and *rel*I alleles is controlled by the positional information coming from the *HoxD* complex, that of *rel*II is not. This means that the fusion gene at this position escapes the higher order regulatory mechanism responsible for the colinearity and behaves as it would anywhere else in the genome. Also, this indicates that colinearity originating at 3′ of the *rel*II insertion exerts its effect further downstream. As the next step, authors recover three deletions, *del*0, *del*II and Ch. 3, Figure 1 *c*. While *del*0 is essentially wild type with respect to both the expression of the reporter gene and the phenotype, *del*II shows early derepression of the reporter gene and a posterior homeotic transformation in heterozygous animals. Since *Hoxd-4* and *Hoxd-10* are prematurely expressed in *del*II (and not in *del*0) this transformation is likely to be a direct consequence of the misregulation of *Hox* genes during their initial phase of activation. These results map the DNA between *rel*0 and *rel*II sites as the region that executes the colinearity of *HoxD* complex regulation. By analysing the Ch. 3 animals, authors further confirm this conclusion.

This discovery argues against the enhancer sharing, promoter competition and other such models (reviewed in ref. 15) and strengthens the view that higher order chromatin organization is directly involved in the regulation of homeotic genes. The results lead to the suggestion that a DNA region at the 5′ end of the *Hox* complex establishes a repressive environment possibly through a particular chromatin conformation. This repressive conformation releases domains of *Hox* genes for transcription in a sequential manner. Boundary or relay elements serve as landmarks that define the limits of such domains. It is very satisfying and exciting to see that we are beginning to have some explanation for the mystery of spatial and temporal colinearity. These studies should provide a framework to investigate molecular aspects of this process. It remains to be seen, though, if other three *Hox* complexes of mouse are regulated by a similar mechanism. Comparison of homologous regions from other vertebrates should provide insight into the DNA sequence requirements to organize the repressive chromatin conformation at one end of the *Hox* complexes.

The misregulation of *Hox* genes – break in the colinearity – in *del*II mice is mainly in the early embryos indicating that *Hox* gene regulation has more complexities particularly at later stages and in different tissues. Split of homeotic gene complex at different sites in two different species of *Drosophila*[16] indicates that there could be a mechanism to add up two clusters into one spatial colinearity. This along with interaction of widely spread out *cis*-regulatory elements of the BX-C and unusual transvection phenomenon described for the *Abd-B* region[17–19] suggest that at least in *Drosophila* long range interactions involving even unlinked homeotic genes and their regulatory elements are part of this complex regulatory system. Polycomb and trithorax group (Pc-G and trx-G) genes are involved in the maintenance of the expression pattern of the homeotic genes that are set by early acting segmentation genes in *Drosophila*. Homologues of the Pc-G genes have been found in mammals, plants and nematodes. Interestingly, mutations in these genes in mouse lead to the kind of phenotype that is similar to that of fly mutants for Pc-G genes[20,21]. This highlights the conservation of mechanisms regulating homeotic genes beyond the initial phase. While DNA elements that are thought to target the Pc-G gene products, the Polycomb Response Elements or PREs, are known in *Drosophila*, are yet to be identified in mammals. Several PREs have been identified all over ANT-C and BX-C. Since PREs are known to interact with each other both in *cis* and in *trans*, this may help bring the spread out or split clusters together at least at some point of time during development. Characterization and functional analysis of such elements in other organisms should help us understand how an early expression pattern is maintained and diversified to implement later steps of the developmental programme.

1. Lewis, E. B., *Nature*, 1978, **276**, 565–570.
2. Dolle, P., Izpisua-Belmonte, J. C., Falkenstein, H., Renucci, A. and Duboule, D., *Nature*, 1989, **342**, 767–772.
3. Izpisua-Belmonte, J. C., Falkenstein, H., Dolle, P., Renucci, A. and Duboule, D., *EMBO J.*, 1991, **10**, 2279–2289.
4. Busturia, A. and Bienz, M., *EMBO J.*, 1993, **12**, 1415–1425.
5. Zhou, J., Ashe, H., Burks, C. and Levine, M., *Development*, 1999, **126**, 3057–3065.
6. Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P., W., Stott, D. and Allemann, R. K., *Genes Dev.*, 1991, **5**, 2048–2059.
7. Puschel, A. W., Balling, R. and Gruss, P., *Development*, 1990, **108**, 435–442.
8. Vogel, R., de Graaf, W. and Deschamps, J., *Development*, 1993, **118**, 71–82.
9. McCall, K., O'Connor, M. and Bender, W., *Genetics*, 1994, **138**, 387–399.
10. Kondo, T., Dolle, P., Zakany, J. and Duboule, D., *Development*, 1996, **122**, 2651–2659.
11. Gyurkovics, H., Gausz, J., Kummer, J. and Karch, F., *EMBO J.*, 1990, **9**, 2579–2585.
12. Kondo, T., Zakany, J. and Duboule, D., *Mol. Cell*, 1998, **1**, 289–300.
13. van der Hoeven, F., Zakany, J. and Duboule, D., *Cell*, 1996, **85**, 1025–1035.
14. Kondo, T. and Duboule, D., *Cell*, 1999, **97**, 407–417.

15. Duboule, D., *Curr. Opin. Genet. Dev.*, 1998, **8**, 514–518.
16. von Allmen, G., Hogga, I., Spierer, A., Karch, F., Bender, W., Gyurkovics, H. and Lewis, E., *Nature*, 1996, **380**, 116.
17. Hendrickson, J. E. and Sakanju, S., *Genetics*, 1995, **139**, 835–848.
18. Hipmann, R., Duncan, D. and Duncan, I., *Genetics*, 1995, **139**, 815–833.
19. Sipos, L., Mihaly, J., Karch, F., Schedl, P., Gausz, J. and Gyurkovics, H., *Genetics*, 1998, **149**, 1031–1050.
20. van der Lugt, N. M., Domen, J., Linder, K., van Roon, M., Robanus-Maanadag, E., te Reile, H., van der Valk, M., Deschamps, J., Sofroniew, M., van Lohuizen, M. *et al.*, *Genes Dev.*, 1994, **8**, 757–769.
21. Yu, B. D., Hess, J L., Horning, S. E., Brown, G. A. and Kormeyer, S. J., *Nature*, 1995, **378**, 505–508.

Rakesh K. Mishra is in the Department of Zoology and Animal Biology, University of Geneva, 1211 Geneva 4, Switzerland.

SCIENTIFIC CORRESPONDENCE

Large-scale chromosome organization in plants underpin genome homogenization, characteristic distribution of repetitive DNAs and occurrence of genes in discrete clusters

Eukaryotic genomes vary greatly in size – which for diploid plants may range from approximately 150 Mbp in *Arabidopsis thaliana* (1C = 0.2 pg, $2n = 10$) to more than 85,000 Mbp in *Fritillaria davisii* (1C = 90 pg, $2n = 24$). So much so, there are vast differences in the nuclear DNA content of even closely related taxa; for example the genomes of *Oryza sativa*, rice (1C = 0.6 pg, 580 Mbp) and *Secale cereale*, rye (1C = 9.5 pg, 9300 Mbp) belonging to same family differ by a factor of sixteen[1]. Such large-scale differences are brought about by the repetitive DNAs, that show extensive differences in sequence motifs and abundance on account of different selective pressures from those acting on genes and evolutionarily successful multigene modules[2,3]. Nevertheless, the distribution of DNA is highly organized with respect to sequence composition, chromosome architecture, and the complex machinery associated with gene regulation, recombination and development.

The chromosome structure is almost as important as the DNA sequence to understand chromosome behaviour. The same gene can have different levels of activity at different chromosomal positions within the same cell, and certain chromosome segments may show linear differentiation although not necessarily defined by specific DNA sequences. Such features may reflect regional differences in condensation between different parts of the chromosomes in interphase cells. To ensure faithful segregation of genetic information during mitosis and meiosis, the DNA is systematically compacted through a fundamental process of chromosomal condensation – for each chromosome this means packing of about 4 cm of DNA into a rod 10 µm long 1 µm in diameter. Data generated in the recent years have provided meaningful insights in understanding the mechanism of chromosome condensation[4] and elucidation of large-scale chromosome organization that have value in gene cloning and evolutionary studies[5].

Plant and animal genomes consist largely of repetitive DNA – perhaps 30 sequence motifs ranging in size from dinucleotides to more than 10,000 bp. Copy numbers of individual repetitive DNA motifs can vary from several hundred to hundreds of thousands, and single motifs may represent 10 or even 50% of a genome. Families of repetitive DNA sequences are differentiated by their degree of sequence homology, distribution among species and/or genome and physical organization[5–7]. The organization of repetitive and single copy sequences along the chromosomes, and positioning of these sequences within the nucleus at interphase have important consequences for plant genetics. Evidence derived from comparative genome analysis of a range of taxa suggests for a strong conservation of gene order – conserved synteny or collinearity of genes, and indeed genes represented in all species can be often regarded as allelic variants. However, the DNA sequences of low-copy genes and regulatory sequences make up only a small proportion, as little as 5% of the total DNA. The interspersed repetitive DNAs found between these genes are very different, making the physical distance between similar loci highly variable[6,8].

Repetitive DNA elements can be divided into two major groups, distinguished by their genome organization and localization on the chromosomes; although intermediate forms of organization can also exist[7]. One group includes sequences showing an organization in tandem repeating units, where individual copies are arranged adjacent to each other forming tandem arrays of the monomeric unit. Such tandemly repeated DNAs are found preferentially at specific positions of the chromosomes, such as the pericentromeric, subtelomeric, telomeric or intercalary regions. DNA elements arranged in tandem arrays include simple sequence repeats (SSRs) and minisatellites, different types of satellite DNAs, the telomeric repeat and the rDNA. In addition to the characteristic telomeric repeat (5'-TTTAGGG-3')n, the telomeres also contain specific non-nucleosomal proteins that coat the