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

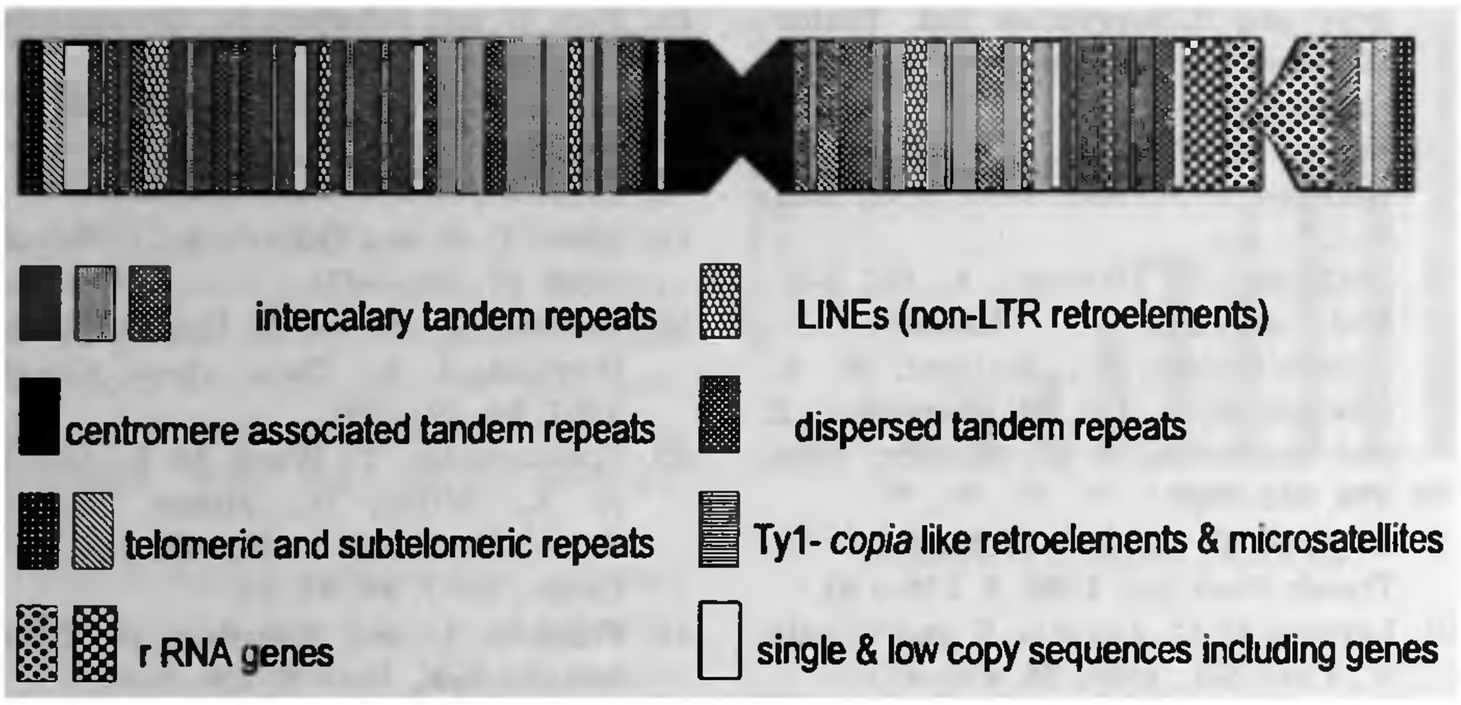

Figure 1. Large-scale organization of plant chromosome depicting characteristic distribution of different classes of repetitive DNA and low copy sequences (modified from the chromosome model by T. Schmidt and J. S. Heslop-Harrison[5]). Genes are clustered in discrete blocks and are characteristically located between the various repetitive DNA motifs. Although copy number of individual repeat motifs may vary between the chromosomes, all chromosomes within a plant genome have similar composition and organization of repeats denoting homogenization of the genome. The simple sequence repeats or microsatellites are dispersed between motifs and may have chromosome/genotype-specific distribution. The rRNA genes are clustered on one or more chromosomes within the complement, preferably along secondary constriction regions.

double-stranded portions of telomeres and facilitate telomere protective function and regulation of length of telomeric DNA tract[9,10]. The DNA binding domain to telomere binding proteins has been dubbed as 'telobox'[11]. SSRs are ubiquitous in plants that evolve rapidly, having implications for genome organization and chromosome evolution. Members of a particular subclass of SSRs, flanked by DNA sequences that are present, ideally, only once in a genome are valuable as molecular markers, and are described by the term 'microsatellite markers'[12]. Also, some satellite DNA repeats can be used as species-specific probes. The other group of repetitive sequences comprises elements with a dispersed organization. Dispersed repetitive DNA elements are scattered throughout the genome, interspersed with other sequences and distributed along the chromosomes, although regions of depletion or amplification can be found[7]. Blocks of nested copies of such elements have also been observed[8,13]. Dispersed DNA sequences include mobile elements like DNA transposable elements and retroelements, their remnants – long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs), and other dispersed repeats. Ty1-*copia* and other families of long terminal repeat (LTR) retroelements constitute the major component of plant nuclear genomes, representing up to 50% of the nuclear DNA[14]. The Ty1-*copia*-like elements show dispersed distribution along chromosome arms, and are under-represented or absent from NORs, centromeres and prominent heterochromatic regions[15,16].

Isolation and molecular characterization of multiple repetitive sequences, each representing a substantial fraction of the genome followed by their physical localization provides a novel top–down chromosomal approach to complement bottom–up DNA marker and clone-based genome analysis for viewing genomic organization, chromosome structure and landmarks for looking at genes, their clustering and orientation. The extensive data obtained on sugar beet and its wild relatives with respect to isolation and *in situ* localization of repetitive DNA elements and C-DNAs led to the proposal of a generalized chromosome model showing the various repetitive elements along the chromosome arms, with clusters of genes between some of the regions of tandemly repeated DNA[5]. A simple modified version of this model is depicted in Figure 1. It is seen from the figure that the large number of families of repetitive DNA, their high amplification and different characteristic locations fill most of the chromosome with repetitive DNA, and genes occur in clusters between blocks of one or more different repeat arrays. That the transcribing regions are present in clusters, and are usually located towards telomeric ends has been explicitly demonstrated in wheat from the analysis of cytogenetic ladder mapping and deletion mapping[17–19].

However, there are some major features that differentiate plant genomes from mammalian and other animal genomes at the level of chromosome organization. It is surmised that repetitive DNA behaves differently in animals and plants, with plants showing greater homogenization between all chromosomes in a species[20]. Whereas each chromosome shows a characteristic GC:AT nucleotide ratio in a mammalian species, all chromosomes within a complement show similar nucleotide composition (i.e. genome homogenization) in plants, except for rDNA sequences that are GC rich[21,22].

In principle, most of the plant species may follow the above model with respect to large-scale chromosome organization, although animal species with much larger genomes may not[5]. The information provided may have value for the following: (i) That the functional genes are organized in clusters, mainly in subtelomeric regions, is a very important information secured from large-scale chromosome organization. This is of great value in facilitating map-based cloning. This is commensurate with the observation that subtelomeric regions are the favoured sites for stable transgene integration, on account of their inherent gene-rich composition[23]. (ii) Although plant chromosomes evince characteristic linear distribution of dispersed repetitive DNA motifs along the chromosomes, their distribution is homogenous with respect to genomic complement *per se*, unlike animal systems[21,22]. Therefore, it may not be easy to develop dispersed region-specific probes to facilitate chromosome painting in plants. Nevertheless, some satellite DNA repeats that show characteristic GC:AT ratio could be used as species-specific probes for plants, as well. (iii) The tandemly re-

peated SSRs are ubiquitous in plants. The SSRs flanked by DNA sequences that are present only once in a genome, i.e. 'microsatellites', can be used as valuable molecular markers, and may be ideal for assessment of genetic diversity and facilitating core collections. (iv) The retroelements comprising repetitive DNA elements with dispersed organization constitute the major component of plant nuclear genome, and exhibit regions of depletion or amplification scattered throughout. Identification of changes associated with retroelements, particularly retrotransposons that account for most of the variation in genome size[13] would be of great value to understand means and ways to genome evolution.

1. Bennett, M. D. and Leitch, I. J., *Ann. Bot.*, 1995, **76**, 113–176.
2. Charlesworth, B., Sniegowski, P. and Stephan, W., *Nature*, 1994, **371**, 215–220.
3. Vershinin, A. V. and Heslop-Harrison, J. S., *Plant Mol. Biol.*, 1998, **36**, 149–161.
4. Murray, A. W., *Science*, 1998, **282**, 425–427.
5. Schmidt, T. and Heslop-Harrison, J. S., *Trends Plant Sci.*, 1998, **3**, 195–199.
6. Heslop-Harrison, J. S., in *Unifying Plant Genomes: Comparison, Collinearity and Conservation* (ed. Heslop-Harrison, J. S.), Company of Biologists, Cambridge, 1996, pp. 17–23.
7. Kubis, S., Schmidt, T. and Heslop-Harrison, J. S., *Ann. Bot.*, 1998, **A82**, 45–55.
8. SanMiguel, P., Tikhonov, A., Jin, Y-K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P. S., Edwards, K. J., Lee, M., Avramova, Z. and Bennetzen, J. L., *Science*, 1996, **274**, 765–768.
9. Shippen, D. E. and McNight, T. D., *Trends Plant Sci.*, 1998, **3**, 126–130.
10. Lavania, U. C., Lavania, S. and Vimala, Y., *Curr. Sci.*, 1999, **76**, 470–473.
11. Bilaud, T., Koering, C. E., Binet-Brassellet, E., Ancelin, K., Pollice, A., Grasser, S. M. and Gilson, E., *Nucleic Acids Res.*, 1996, **24**, 1294–1303.
12. Cuadrado, A. and Schwarzacher, T., *Chromosoma*, 1998, **107**, 587–594.
13. Higashiyama, T., Noutoshi, Y., Fujie, M. and Yamada, T., *EMBO J.*, 1997, **16**, 3715–3723.
14. SanMiguel, P. and Bennetzen, J. L., *Ann. Bot.*, 1998, **A82**, 37–44.
15. Heslop-Harrison, J. S., Brandes, A., Taketa, S., Schmidt, T., Vershinin, A. V., Alkhimova, E. G., Kamm, A., Doudrick, R. L., Schwarzacher, T., Katsiotis, A., Kubis, S., Kumar, A., Pearse, S. R., Flavell, A. J. and Harrison, G. E., *Genetica*, 1998, **100**, 197–204.
16. Pich, U. and Schubert, S., *Chromosome Res.*, 1998, **6**, 315–321.
17. Gill, K. S., Gill, B. S., Endo, T. R. and Taylor, T., *Genetics*, 1996, **144**, 1883–1891.
18. Endo, T. R. and Gill, B. S., *J. Hered.*, 1996, **87**, 295–307.
19. Castilho, A., Miller, T. E. and Heslop-Harrison, J. S., *Theor. Appl. Genet.*, 1997, **94**, 293–297.
20. Schwarzacher, T., Wang, M. L., Leitch, A. R., Miller, N., Moore, G. and Heslop-Harrison, J. S., *Theor. Appl. Genet.*, 1997, **94**, 91–97.
21. Fronicke, L. and Scherthan, H., *Chromosome Res.*, 1997, **5**, 254–261.
22. Fuchs, J., Houben, A., Brandes, A. and Schubert, I., *Chromosoma*, 1996, **104**, 315–320.
23. Lavania, U. C., Lavania, S. and Vimala, Y., *Curr. Sci.*, 1998, **75**, 538–539.

ACKNOWLEDGEMENT. I am grateful to Prof A. K. Sharma for his valuable suggestions and DBT, New Delhi for funds.

U. C. LAVANIA

Central Institute of Medicinal and Aromatic Plants,
Lucknow 226 015, India

Increased bilirubin binding to erythrocytes of tobacco chewers than non-chewers

The binding of unconjugated bilirubin, a catabolic product of haemoglobin, to various types of cells including erythrocytes and its toxic effects are well known[1]. Under physiological conditions, human erythrocytes bind a fixed amount of bilirubin from a given bilirubin load[2,3]. This bilirubin binding property of erythrocytes, however, changes with the variation in the physico-chemical properties of erythrocyte membranes[4]. Recently, we have demonstrated that the erythrocytes from healthy smokers bind more bilirubin than erythrocytes from healthy non-smokers[5]. This increase in bilirubin binding to the smokers' erythrocytes can be correlated to the smoke-induced changes in the physico-chemical properties of erythrocyte membrane[6,7]. The nicotine in tobacco can also be absorbed simply by contact with the mucus membranes of the mouth[8]. But, unlike the effects of cigarette smoke, incubation of nicotine with erythrocytes is reported to show decreased deformability of the cells[9]. Whether the effect of tobacco leaf extract on the bilirubin uptake by erythrocytes is same as that observed with cigarette smoke is not known. Therefore, it is interesting to study the binding of bilirubin to erythrocytes from healthy tobacco chewers. Here, we present our data on the binding of bilirubin to erythrocytes from healthy tobacco chewers and healthy non-chewers. Our data suggest that erythrocytes from healthy tobacco chewers bind more bilirubin than that from healthy non-chewers.

Blood samples of both healthy tobacco chewers consuming more or less similar amount of nicotine per day and non-chewers (females of the age group between 25 and 40 years) were collected in 4% sodium citrate after interviewing the persons concerned thoroughly. Erythrocytes were isolated from the blood samples by centrifugation at 1000 *g* for 20 min. Then, 1.0 ml of albumin (2.0%) in 0.07 M sodium phosphate buffer, pH 7.4, containing 0.08 M sodium chloride was added and the mixture was incubated for 20 min in