

Molecular mapping of gene-rich chromosome-specific sequences – A complementary approach to plant genome initiative in India

U. C. Lavania, Seshu Lavania and Y. Vimala

Our understanding of organismal genetics is now extending from phenotypic to molecular basis. Molecular physical maps are important resource for facilitating positional cloning of trait genes, sequencing of genomic DNAs and analysis of chromosome and genome structure in detail. The physical maps, therefore, will become central to all types of genetic and molecular enquiry and manipulation, including genome analysis, gene cloning, and crop genetic improvement. It is, therefore, in our immediate interest that concerted efforts are initiated in India in the fast emerging area of genomics, and the short-term and long-term objectives are prioritized. A systematic programme on molecular mapping of gene-rich segments in major Indian plants with special emphasis on the identification of novel sequences is desirable.

A major recent discovery in plant genetics is occurrence of gene synteny, i.e. parallelism of gene order. The complement and chromosomal arrangement of genes in different plant genomes is very similar. This is particularly true for grasses, such as maize, sorghum, wheat and rice. Thus, maize and rice differ more in the version of genes (i.e. alleles) that they carry than they do in the actual type of genes. Interestingly, not only is the order of DNA markers well conserved for most of the genome, but also a small number of quantitative trait loci (QTLs) that determine some agronomic phenotypes correspond closely in sorghum, rice and maize¹⁻⁴.

This implies that the isolation and study of genes from one grass can provide information and sequences that can be used in the improvement of all other grass species. Consequently, genes isolated from one flowering plant can generally be used to isolate the corresponding genes from other plants by heterologous hybridization. It should be interesting to mention here that the three major cereal crops, rice, maize and

wheat, differ distinctly with respect to their genome size – estimated to be 430 Mb, 2000 Mb and 16,000 Mb, respectively, but comprise almost similar number of genes, i.e. 30,000 as predicted from comparative genetics⁵. Therefore, information obtained from rice map could be of great help in the identification of genes in other cereals like wheat having large genome. Furthermore, promoters for many genes retain normal patterns of expression when introduced into distantly-related plant species⁶.

Nevertheless, this potential synergy does not extend indefinitely, and it is clear that even a full understanding of the *Arabidopsis* genome will not provide sufficient information for a complete understanding of the biology of monocot crops⁷. Therefore, an obvious question is – what level of details are needed? There are two major viewpoints. One may like to fish out from corn or other model plants just the sequences of genomic DNA most likely to code for genes, i.e. typically a small part of any genome. The other view is that sequencing the entire genome is the only way to find all the genes and understand their relation to each other⁵. Curiously, the two views are constitutive and complementary in nature. Bennetzen *et al.*⁷ further opine that it may be essential to undertake intensive investigations into a small number of key species to generate useful data and tools with respect to their entire genome in view of their immediate usefulness to all major crops. The frequent collinearity and similar gene content of plants may further serve as potential resource for making such connections.

In respect to flowering plants, two major international efforts are underway for whole genome mapping: *Arabidopsis* (150 Mb) and rice (430 Mb). Lately, National Science Foundation, USA, has announced a Plant Genome Research Programme (PGRM) to expand genom-

ics – and have identified maize (2000 Mb) for developing the total genome maps⁷. The Department of Biotechnology, Government of India, has also taken laudable steps in this direction through establishment of National Plant Genome Research Centre (NPGRC) at the Jawaharlal Nehru University, New Delhi. Map information gained from whole genome maps could be of immense value in isolating corresponding gene for a given target species. Simultaneously, there may be some valuable genes that may be unique but may not be covered by the species where whole genome map would be available. Therefore, it would be desirable to search for such novel sequences in selected important plants as well.

Development of chromosome-specific libraries. The large size and complexity of most of the higher plant genomes make the map saturation a difficult and tedious task. However, division of the genome into smaller parts that can be studied separately would simplify and accelerate its analysis. This approach would also allow isolation of region-specific markers and integration of genetic and physical maps⁸.

One promising way to solve this problem is to focus on the DNA of individual chromosome/chromosome arm. This requires separating an individual chromosome type and constructing a genomic library from the isolated chromosomes⁹. Individual chromosome types could be purified by flow sorting and their DNA amplified by degenerate oligonucleotide primed (DOP) polymerase chain reaction⁸. The chromosomal DNA can be cloned into a suitable vector in order to generate a library that can serve as a source for linkage group-based molecular markers from the regions of interest. Utilization of reconstructed cytogenetic stocks can further facilitate such efforts.

Preferential mapping of gene-rich regions. The main difference in genome

structure between a large and small genome plant is in the size of gene-poor block separating the gene-rich blocks. The 22-fold difference in the size of gene-poor regions compares well with the 35-fold difference in the genome size of wheat and rice. It is, therefore, not surprising that the base pair/centimorgan estimates for the gene cluster regions are nearly comparable in the two species. Thus, it is mainly the presence of substantial amount of nontranscribing repetitive DNA blocks that brings about increase in the genome size of wheat¹⁰. Further, it has been demonstrated through cytogenetic ladder mapping that the transcribing regions in wheat are present in clusters, and are usually located towards the telomeric ends¹¹.

Such findings are commensurate with general features of linear differentiation of plant chromosomes that are richly endowed with heterochromatin mainly in telomeric and centromeric regions¹². Further, the matrix attachment regions (MARs) that serve as link between nuclear membrane attachment sites and expressed sequence tags (ESTs), and facilitate gene expression, remain mainly associated with the subtelomeric regions¹³. Also, it has been shown that even the transformed sequences when targeted to subtelomeric sites show

enhanced expression and genetic stability^{14,15}. It may, therefore, be considered that the subtelomeric regions are rich in functional genes. Therefore, emphasis may be laid to develop maps for this region as a first step, at least in the selected plants where India has major interest. It is not a very difficult approach, since the techniques for microdissection and microamplification of specific chromosome sequences are now available¹⁶⁻¹⁸.

1. Ahn, S. and Tanksley, S. D., *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 7980-7984.
2. Moore, G., Devos, K. M., Wang, Z. and Gale, M. D., *Curr. Biol.*, 1995, **5**, 737-739.
3. Paterson, A. H., Lin, Y-R., Li, Z., Schertz, K. F., Doeblay, J. F., Pinson, S. R. M., Liu, S-C., Stansel, J. W. and Irvine, J. E., *Science*, 1995, **269**, 1714-1718.
4. Bennetzen, J. L. and Freeling, M., *Genome Res.*, 1997, **7**, 301-306.
5. Cohen, J., *Science*, 1997, **276**, 1960-1962.
6. Gibson, S. and Somerville, C., *Trends Biotechnol.*, 1993, **11**, 306-313.
7. Bennetzen, J. L., Kellog, E. A., Lee, M. and Messing, J., *Plant Cell*, 1998, **10**, 488-490.
8. Macas, J., Gualberti, G., Nouzova, M., Samec, P., Lucretti, S. and Dolezel, J., *Chromosome Res.*, 1996, **4**, 531-539.

9. Lee, J-H., Arumuganathan, K., Yen, Y., Kaeppler, S., Kaeppler, H. and Baenziger, P. S., *Genome*, 1997, **40**, 633-638.
10. Gill, K. S., Gill, B. S., Endo, T. R. and Boyko, E., *Genetics*, 1996, **143**, 1001-1012.
11. Gill, K. S., Gill, B. S., Endo, T. R. and Taylor, T., *Genetics*, 1996, **144**, 1883-1891.
12. Lavania, U. C. and Sharma, A. K., *Proc. Indian Acad. Sci.*, 1983, **92**, 51-79.
13. Holmes-Davis, R. and Comai, L., *Trends Plant Sci.*, 1998, **3**, 91-97.
14. Iglesias, V. A., Moscone, E. A., Papp, I., Neuhuber, F., Michalowski, S., Phelan, T., Spiker, S., Matzke, M. and Matzke, A. J. M., *Plant Cell*, 1997, **9**, 1251-1264.
15. Pedersen, C., Zimny, J., Becker, D., Jahne-Gartner, A. and Lorz, H., *Theor. Appl. Genet.*, 1997, **94**, 749-757.
16. Ponelies, N., Stein, N. and Weber, G., *Nucleic Acids Res.*, 1997, **25**, 3555-3557.
17. Saitoh, Y. and Ikeda, J-E., *Chromosome Res.*, 1997, **5**, 77-80.
18. Hernould, M., Glimellus, K., Veuskens, J., Bergman, P. and Mouras, A., *Plant J.*, 1997, **12**, 703-709.

U. C. Lavania is in the Central Institute of Medicinal and Aromatic Plants, Lucknow 226 015, India; Seshu Lavania is in Lucknow University, Lucknow 226 007, India and Y. Vimala is in Jiwaji University, Gwalior 474 011, India.

COMMENTARY

This vision of structural biology and biomedical research is almost 14 years old. Reexamining 'old visions' in the light of new knowledge may be instructive.

- Editors

Looking back: A vision of molecular science and medicine*

V. Ramalingaswami

There is, today, a new and beautiful vision of life revealed to us by modern biology. The vision is linked to a hope

*Based on the inaugural address delivered at the International Symposium on 'Bio-molecular Structure and Interactions' at the Indian Institute of Science, Bangalore, on 17 December 1984.

of being able to cope with the problems of *population growth, agriculture and parasitic diseases*. Understanding molecular structure and relating it to biological function is today the most exciting and intense pursuit of science. So far as medicine is concerned, at first it was correlating altered morphology of organs in disease states with clinical

syndromes. This could be called the Rokitsansky era led by the Vienna school. With the discovery of the light microscope, this pursuit of relating structure to function reached its culmination in the concept of *cellular pathology* propounded by Rudolf Virchow. Virchow called the cell as the *theatre of life*. The pioneering work of Pasteur and