Research News

Physical education – new technologies for mapping plant genomes

Robert Waugh, Paul H. Dear, Wayne Powell and Gordon C. Machray

Locating DNA sequences to specific chromosomal segments is essential for associating genes with phenotypes. It is routinely achieved by segregation analysis using meiotic mapping populations that have also been used to collect phenotypic information. However, meiotic mapping is struggling to cope with the shear volume of sequences emerging from high-throughput (HTP) gene-discovery programs. We describe two approaches, Radiation Hybrid and 'HAPPY' mapping, which, in conjunction with meiotic mapping, represent valuable HTP tools in the quest to link genes to phenotypes.

The discovery of extensive and easily monitored DNA sequence variability in plant genomes provided the impetus for the intensive genetic mapping studies that have been conducted since the mid-1980s. By correlating the pattern of inheritance of phenotypic trait scores with those of individually mapped genetic markers, many monogenic and polygenic characters have been located [1]. However, meiotic mapping has significant limitations in a high-throughput (HTP) environment (Box 1).

Physical mapping circumvents most of these. In particular, it precludes the requirement for markers to be polymorphic. Thus, STS (sequence tagged site)-content-analysis of libraries of large-insert clones [e.g. YACs or BACs (yeast artificial chromosomes or bacterial artificial chromosomes)] provides a high-resolution physical mapping approach that concurrently builds an archive of genomic information. Integrating physical and genetic maps facilitates map-based gene isolation, candidate gene identification and comparative genome analysis, and provides sequence-ready clones for genome sequencing and links genes with phenotypes. However, mapping genomes by STS-content-analysis is restricted by the process on which it relies: unclonable regions result in gaps, and rearranged or chimaeric fragments, or repeated regions larger than the size of the clones, lead to distortions. Target genome size, redundancy (required to achieve complete genome representation) and average clone insert size all affect the practicalities of this approach. This has led to the proposal that physical maps are most effective when built over an independently constructed STS scaffold, to ensure both long-range accuracy and map integrity. The following technologies might satisfy this goal.

Radiation hybrid mapping

Radiation hybrid (RH) mapping of the human genome culminated in the development of a physical map based on >40 000 STS [2], and fuelled RH technology development in many vertebrate and invertebrate species. Radiation hybrids [3] are generated by lethally irradiating cells of a donor species to induce random chromosomal breaks, producing fragments that can be subsequently rescued by fusion with a suitable recipient cell line from a different species [4]. Each resulting 'hybrid' contains a fraction of the donor genome retained as a collection of large genomic fragments integrated into the recipient genome (Fig. 1). A useful panel of hybrid cell lines should include representation of the entire donor genome. The mapping concept is as follows: the smaller the physical distance between two genes (or markers), the higher the probability that they will be retained on the same chromosomal fragment after fragmentation by irradiation, and thus incorporated and maintained in the same cell lines. By establishing the retention of target DNA sequences in individual panel members (by STS-PCR), co-retention frequencies can be translated into physical distance by statistical methods. Marker retention is scored on a plus or minus basis with no requirement for polymorphism. Consequently, RH panels constitute one of the most expedient methods for high-resolution physical mapping.

RH mapping could greatly advance studies of plant genome organization. However, attempts to adopt RH technology have been virtually non-existent. There are two exceptions. The first, unique to plants, exploits the capacity of polyploid host species to

Box 1. Meiotic mapping

Meiotic mapping monitors the inheritance of genes or markers in a population generated by sexual hybridization. During meiosis, recombination breaks and rejoins parental chromosomes forming a unique mosaic in each of the progeny plants. A limited number of recombination events occur. When parents have different alleles at a marker locus, their inheritance in the progeny can be followed. A high frequency of co-inheritance of gene or marker alleles indicates that they are genetically linked. Limitations of meiotic mapping include:

- The inability to map without polymorphism between the parents. Consequently, markers or genes in chromosomal regions identical by descent are impossible to order.
- The relationship between mapping resolution and population size. Ordering marker pairs that become ever closer is increasingly difficult because of the decreasing probability of recombination occurring between them. The only solution is to increase the population size.
- The poor relationship between meiotic and physical distance. Markers with high co-inheritance frequencies are depicted close together on the meiotic map. The true physical distance between them is unknown – they might be separated by more than half the total length of the chromosome [a].
- Failure to cope with multi-gene families. This leaves no alternative but to use gene-specific approaches based on comprehensive sequence data [b].

References

a Kuenzel, G. *et al.* (2000) Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. *Genetics* 154, 397–412

b Choi, D.W. *et al.* (1999) The barley (*Hordeum vulgare* L.) dehydrin multigene family: sequences, allele types, chromosome assignments, and expression characteristics of 11 Dhn genes of cv Dicktoo. *Theor. Appl. Genet.* 98, 1234–1247

 1360-1385/02/$ – see front matter © 2002 Elsevier Science Ltd. All rights reserved. PII: S1360-1385(02)02373-7

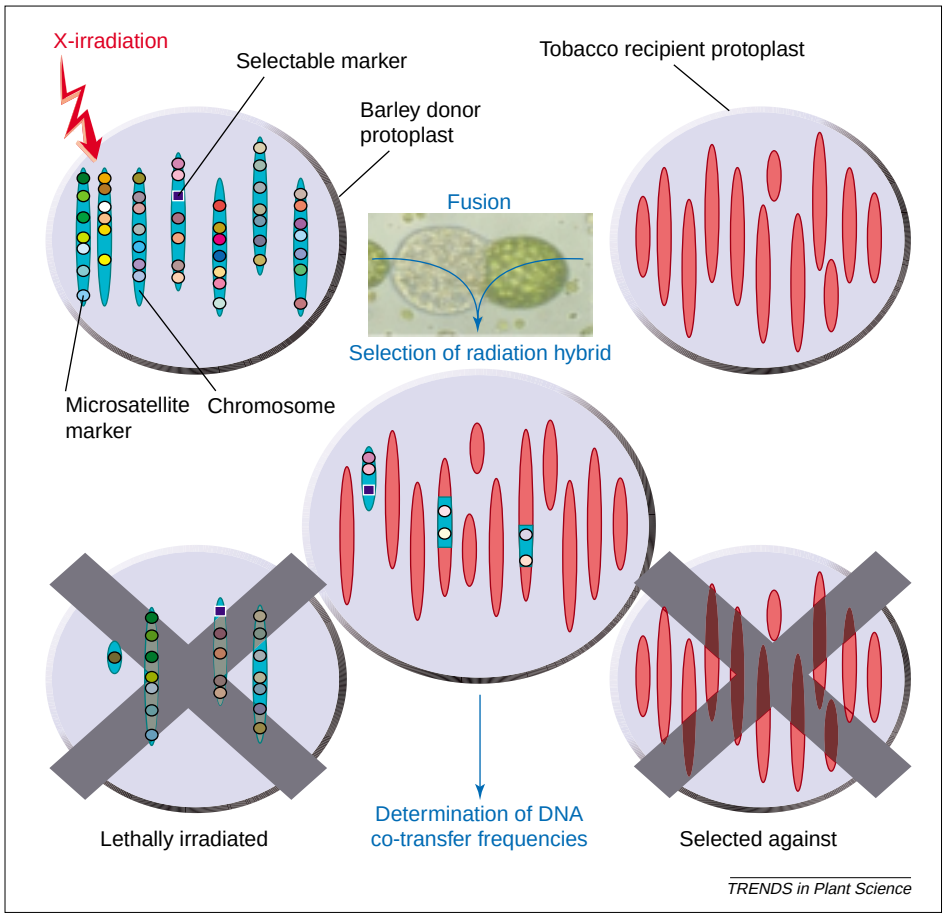


Fig. 1. Production of a whole genome radiation hybrid (RH) mapping panel. Protoplasts prepared from barley embryogenic cell suspensions are lethally irradiated and fused to protoplasts obtained from tobacco leaf mesophyll. Hybrid products are cultured and selected by the presence of the selectable marker present in the barley genome. They are then surveyed for additional random fragments of barley DNA through the use of barley-specific PCR-based molecular markers. Statistical analysis of marker co-retention frequencies across a panel of hybrids enables a physical map of the barley genome to be constructed.

genetically buffer changes induced by mutation. Oscar Riera-Lizarazu and colleagues [5] described the development of RH hybrids from a monosomic oat–maize chromosome-9-addition line. After exposure to gamma irradiation, the allohexaploid oat genome 'genetically buffered' potentially lethal chromosomal mutations and other aberrations and allowed regeneration of fertile 'RH' plants. Recovered plants were analysed by chromosome-9-specific marker analysis and FISH (fluorescence *in situ* hybridization). The results confirmed maize chromosome-9 fragmentation and the subsequent rescue of segments in regenerated plants. Between one and ten radiation-induced breaks were identified per line. The authors conclude that with an average of three breaks per chromosome and a chromosome size of 191 Mbp, a panel of 100 such lines assayed with ~400 evenly spaced markers would allow physical mapping to a resolution of 0.5–1.0 Mb. If used in combination with a complete set of oat–maize chromosome addition lines, panels of all ten chromosomes would provide a unique resource for mapping maize ESTs at high physical resolution. Similar approaches could be envisaged for other species that have analogous aneuploid stocks (e.g. the *Triticeae*).

In a study more accurately paralleling the highly successful mammalian approach, Julie Wardrop and colleagues [6] constructed a whole genome RH panel for barley using *Nicotiana tabacum* cv. Xanthi as the host. The authors fused lethally irradiated bialaphos-resistant (i.e. *bar* gene transgenic) barley cell suspension protoplasts with tobacco mesophyll protoplasts by electro-fusion. Hybrid cells were rescued by culture on bialaphos-containing medium optimized for the regeneration of tobacco protoplasts. Hybridity of the resulting calli was verified by assaying (by PCR) for the presence of the *bar* gene, which indicated that ~20% of the putative hybrids contained barley chromatin. These lines were further analysed to estimate co-retention frequency and genome representation. The first of these parameters was estimated by assaying a barley genome marker located 3.7 recombination units away from the transgene insertion site, and a pair of markers on a different chromosome that were known to be only 16 kb apart. The linked marker was co-inherited with *bar* in 77% of the hybrids and the marker pair 16 kb apart was always co-inherited. Further analysis using a genome-wide panel of barley-specific SSR markers showed that the overall donor DNA retention frequency (the percentage of markers retained within individual hybrids) averaged 26% and all regions of the genome tested were represented. These figures are within the effective range used in mammalian panels.

The objective here was not to regenerate plants. Calli were harvested after culture for a limited time on selective medium to prevent the loss of acentric chromosome segments. However, sacrificing the calli provided only a limited DNA resource. Thus, although demonstrating the potential of whole genome RH mapping in crop plants, current work is directed towards the development of an effective approach for whole hybrid-genome amplification to provide sufficient material for the thousands of assays required to develop a RH map.

HAPPY mapping

HAPPY mapping [7] is an *in vitro* mapping technique. Its utility has been demonstrated by mapping human chromosome 14 [8], and the genomes of *Cryptosporidium parvum* [9] and *Dictyostelium discoideum* [10,11]. More recently, it has been applied to plants. HAPPY mapping involves breaking intact naked genomic DNA at random, segregating the fragments into aliquots (the 'mapping panel'), each aliquot containing less than one genome's worth of DNA fragments (the *in vitro* analogue of an RH cell), and measuring the frequency of co-segregation of markers among the aliquots. Because closely linked markers are rarely separated by an intervening break, they tend to co-segregate, making it analogous to RH mapping. Just as each RH cell line

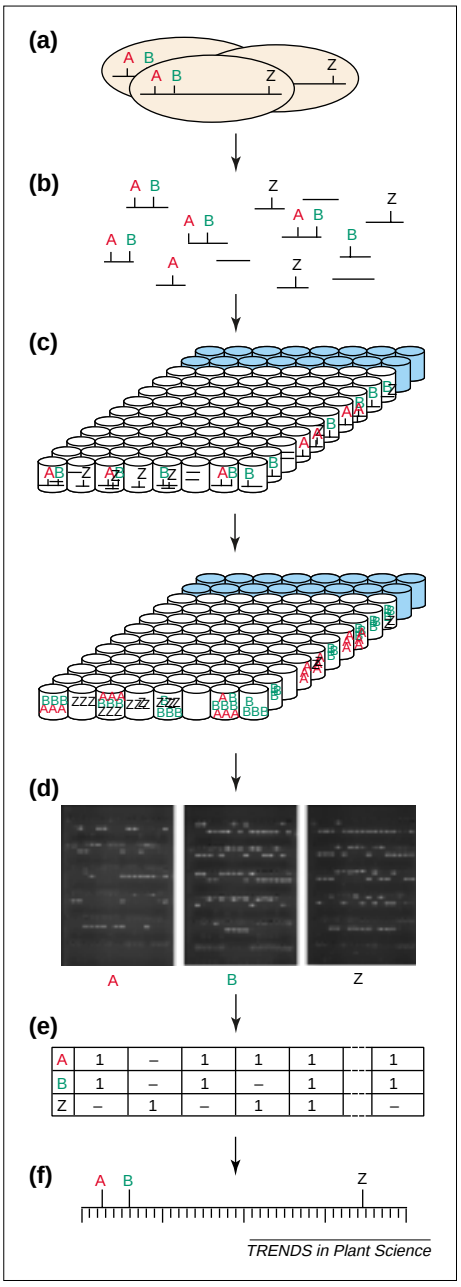


Fig. 2. Principle of HAPPY mapping and marker typing. High molecular weight DNA carrying STS (sequence tagged site) markers (A, B, Z) separated by different physical distances along a chromosome is extracted from cells (a) and broken randomly to give a pool of fragments of 'mean size' (b). After screening empirically to identify a dilution that contains 0.5–0.7 genomes in a given aliquot, the chosen dilution is used to construct a mapping panel comprising 80 sample individuals (white) and 16 'no DNA' controls (light blue) in a microtiter plate. These sub-genomic amounts of DNA per aliquot are first bulked by whole-genome PCR (c) to 'immortalize' the panel and facilitate typing of thousands of markers. The panel is screened by marker-specific PCR and gel analysis (d) to determine the marker content of each aliquot. The resulting data are used to construct a data table (e), which is subjected to analysis. Closely linked markers (A,B) show almost complete co-segregation; remote markers (B,Z) do not. Co-segregation frequencies reflect marker-to-marker distances, allowing a map (f) to be computed. A known physical distance between any two markers is used to calibrate the panel and to convert co-segregation frequencies into physical distance estimates.

can be grown to provide material for many assays, the members of a HAPPY panel can be pre-amplified to allow thousands of marker-typings. Constructing and screening a HAPPY mapping panel is straightforward, and only small quantities of DNA are needed. The basic principle is outlined in Fig. 2.

The potential of HAPPY mapping in plants has been demonstrated by Madan Thangavelu and co-workers [12] who constructed a high-resolution HAPPY map of a section of *Arabidopsis thaliana* chromosome 4 with an average marker spacing of 16 kbp. They prepared a HAPPY mapping panel comprising sheared genomic DNA that was subsequently analysed for the presence of PCR-amplified STS markers spanning the region of interest. Analysis of marker co-segregation frequencies enabled their relative order to be estimated and the known distance between two markers (from the sequence) allowed the map to be calibrated. When compared to the DNA sequence, a remarkably good correspondence was observed. The order of >85% of the markers was identical to the genome sequence and the errors in the positions of all but one of the markers were well within the margins of error expected from the resolution of the panel. The authors are currently focussing their attention on crop plants. These experiments indicate that HAPPY mapping is a uniquely versatile tool for physical mapping, even in species with large genomes. It also provides a valuable 'comparative physical mapping' tool, avoiding the need to construct large insert DNA libraries for different accessions followed by comparative structural analysis.

Future prospects

RH and HAPPY mapping represent technologies that bridge the resolution gap between meiotic and BAC-based STS-content mapping. By varying radiation dose or applied shear, the range and resolution of the mapping panels are manipulable. The different methodologies used to induce chromosome fragmentation are likely to provide complementarity; either might have the ability to map regions refractory to the other. Both are suited to constructing physical maps and preclude the need to detect polymorphism. As such, they should greatly aid the analysis of large or unfashionable crop plant genomes for which complete genomic sequencing projects are not yet established.

Acknowledgements

We acknowledge Julie Wardrop, Alan James, Glenn Bryan, Luke Ramsay and Madan Thangavelu for their diligence and skill in developing RH and HAPPY technologies in plants, and BBSRC/SEERAD and the Leverhulme Trust for financial support.

References

1 Tanksley, S.D. (1993) Mapping polygenes. *Annu. Rev. Genet.* 27, 205–233
2 Olivier, M. *et al.* (2001) A high-resolution radiation hybrid map of the human genome draft sequence. *Science* 291, 1298–1302
3 Goss, S.J. and Harris, H. (1975) New method for mapping genes in human chromosomes. *Nature* 255, 680–684
4 Barrett, J.H. (1992) Genetic mapping based on radiation hybrid data. *Genomics* 13, 95–103
5 Riera-Lizarazu, O. *et al.* (2000) Production and characterization of maize chromosome 9 radiation hybrids derived from an oat-maize addition line. *Genetics* 156, 327–339
6 Wardrop, J. *et al.* (2002) Constructing plant radiation hybrid panels. *Plant J.* 31, 223–228
7 Dear, P.H. and Cook, P.R. (1993) Happy mapping – linkage mapping using a physical analog of meiosis. *Nucleic Acids Res.* 21, 13–20
8 Dear, P.H. *et al.* (1998) A high-resolution metric happy map of human chromosome 14. *Genomics* 48, 232–241
9 Piper, M.B. *et al.* (1998) A happy map of *Cryptosporidium parvum*. *Genome Res.* 8, 1299–1307
10 Konfortov, B.A. *et al.* (2000) A high-resolution happy map of *Dictyostelium discoideum* chromosome 6. *Genome Res.* 10, 1737–1742
11 Glöckner, G. *et al.* (2002) Sequence and analysis of chromosome 2 of *Dictyostelium discoideum*. *Nature* 418, 79–85
12 Thangavelu, M. *et al.* HAPPY mapping in a plant genome: reconstruction and analysis of a high-resolution physical map of a 1.9 Mbp region of *Arabidopsis thaliana* chromosome 4. *Plant Biotechnol. J.* (in press)

Robert Waugh*
Wayne Powell
Gordon C. Machray
Scottish Crop Research Institute, Invergowrie, Dundee, UK DD2 5DA.
*e-mail: rwaugh@scri.sari.ac.uk

Paul H. Dear
MRC Laboratory of Molecular Biology, Protein and Nucleic Acid Chemistry Division, Cambridge, UK CB2 2QH.