Plant molecular and cell biology

Wayne Powell and Howard V. Davies

Reorganisation of the Institute's research activities brings enhanced prominence to the thematic area of plant molecular and cellular biology. As the primary sequences of most plant genes become identified, new opportunities will arise for their exploitation in fundamental and applied research. In the future, greater emphasis will need to be placed on basic genetics and biochemistry to fill the knowledge gap between genome structure and function. This will demand that we have a better understanding of how living cells operate, function and most importantly interact and communicate. These scientifically challenging goals will require a focused commitment from us all.

RNA processing and gene expression This year has seen the culmination of 3-4 years of research on the novel organisation of small nucleolar RNA (snoRNA) genes in plants. These genes encode small RNAs required for the production of ribosomal RNAs (rRNAs). The molecular characterisation of the unique organisation, transcription and processing of these snoRNAs in plants has been complemented by cell biological localisations of various rRNA and snoRNA components carried out in collaboration with Dr Peter Shaw's group at the John Innes Institute. This exciting and productive area has had impact in the field of rRNA processing and ribosome biogenesis in general.

Research on pre-mRNA splicing continues to maintain a high profile and has concentrated on demonstrating the importance of exon scanning mechanisms in plant splicing. The establishment of experimental systems to show exon scanning and the molecular characterisation of particular *Arabidopsis* mutants, has provided the majority of evidence for such mechanisms in plants. In addition, progress is being made in understanding the regulation of alternatively spliced systems, and, in particular, how a mini-exon of only 9 nt becomes included in potato invertase genes. A greater knowledge of these post-transcriptional processes will allow better understanding of how plant genes and transgenes are expressed and how expression may be regulated.

Finally, the genes which encode the various components of the pre-mRNA splicing or rRNA processing machinery, provide a rich source of promoters for transgene expression in plant biotechnology. As most genes are organised in multigene families with great variability in expression levels and patterns, a novel approach has been developed to allow the identification and isolation of promoters with the required expression characteristics. This approach will prove valuable for future promoter isolation and exploitation.

The DNA Sequencing Facility has been run from June 1997 by Clare McQuade. The facility is well-organised and runs efficiently to meet the current sequencing needs of users across the breadth of the Institute.

The structural characterisation of all potato apoplastic invertase gene promoters has been completed. This included determination of the full nucleotide sequence of the single promoter for which a genomic clone was available and the nucleotide sequence for three further invertase gene promoters which were obtained by genomic walking. The acquisition of single copy sequence information from closely-related members of a gene family by genomic walking in the large tetraploid genome of potato, represents a significant achievement in the application of this technology to polyploid plant genomes. Transcription start points were determined for each promoter by the use of a sequence-specific RT-PCR approach which was also employed to obtain expression information at the organ level. This expression analysis was extended and refined by the construction of promoter-reporter gene fusions for each of the invertase genes. A set of confirmed transgenic potato plants was prepared for each of these fusions allowing accurate histological assessment of gene expression at the level of specific tissues.

Intellectual property rights have been obtained to protect the information generated from these analyses which offer a detailed insight into the function of these enzymes in plant sucrose metabolism. Through this insight, novel projects examining gene expression and targeted mutation in pollen have become possible. Potential commercial applications of the manipulation of sucrose metabolism in cold-stored potato tubers have also been advanced with the completion of a series of constructs designed to manipulate levels of invertase and other enzymes of potato carbohydrate metabolism. An extensive interdepartmental effort has produced *c.* 10,000 transgenic potato plants which are undergoing field-trialling extending over several years. A further collaborative effort is under way to characterise the signals which regulate processing of the unique mini-exon found in plant invertase genes.

Our work on plant carbohydrate metabolism has recently been extended by the initiation of new projects which will yield essential knowledge and technology to underpin the manipulation of quality traits of the barley grain. The first of these involves the generation of a catalogue of gene expression in malting barley - the determination of an initial 1,000 expressed sequence tags (ESTs) is well under way. These, when extended into an expression profile and map, will provide both the raw material for genetic manipulation and an indication of the metabolic systems in which it might most usefully be employed. Such manipulation will require robust and efficient cereal transformation systems - work on both barley and wheat has commenced, with the engineering of virus-resistance in Chinese wheat, analogous to that engineered effectively into dicots, providing a real framework for the establishment of this key technology.

1997/98 saw the development of further plant RNA virus-based, foreign protein expression vector systems utilising the '2A-linker to coat protein' rationale behind the original Potato Virus X-OVERCOAT®. Specifically, Tobacco Mosaic Virus was found, counter intuitively, on X-ray diffraction derived structure data and from prior reports, to be able to accommodate large OVERCOAT® proteins up to the size of the green fluorescent jellyfish protein (27 kDa), fused to 5% approximately of the $2{,}400^{+}$ coat protein sub-units required to encapsidate the recombinant genome. In addition, a vector was created which increased the efficiency and frequency of co-translational release of the upstream, 'foreign' OVERCOAT® moiety by virtue of arranging a double 2A linker peptide sequence between the foreign 'gene' and the TMV coat protein gene (work done by Christophe Lacomme).

A substantial Scottish Office/CHABOS co-ordinated programme initiative (£938 K) was compiled and funded during this period (SCR/824/97) with the objective of studying novel delivery routes for therapeutic/prophylactic or growth promoting peptides and/or proteins with veterinary applications (collaboration with the Hannah, Moredun and Rowett Research Institutes).

Work also began on a six partner, EU-INCO-funded programme between the UK, Germany and China to produce virus-resistant elite germplasm in Chinese cultivars. Several 'biolistic' transformation vectors for wheat embryos have been constructed and are being tested at present.

Genomics Marker technology development has progressed well in potatoes, barley and various tree species. For potato, more than 100 nuclear- and 20 chloroplast-derived simple sequence repeat based-markers have been assembled. Approximately 60% of the nuclear SSRs have now been genetically mapped. As with barley, the development of this class of markers for potato has provided tools which are highly appropriate for a range of applications. In diploid potato populations, the SSRs have been used to study the inheritance of quantitative resistance to late blight and other characters such as vigour and earliness, and SSRs in regions of the potato genome which harbour genes affecting these characters (particularly on potato chromosome V) have been identified. Establishing the map location of SSRs in experimental diploid populations has been particularly valuable in attempts to unravel some of the complexities of genetic inheritance in cultivated, tetraploid potatoes. This is because the number of distinguishable SSR alleles at a single locus allows a large portion of the parental genomes to be accurately followed in segregating populations. Thus, the SSRs have been used for the first time to examine the inheritance of quantitative resistance to late blight, *S. tuberosum* subsp. *andigena*-derived polygenic resistance to PCN and sensitivity to the important processing characteristic of low temperature sweetening in a tetraploid population. Multiple allelism allows allelic-bridges to be formed between (the 12 sets of 8) independently segregating genetic linkage groups and these linkage groups to be assigned specific chromosomal designations. Information derived from these analyses can then be directly compared to that available from diploid populations. Informative SSR-markers which explain a large proportion of the phenotypic variation in the population for late blight and PCN resistance, have been identified on potato chromosome IV. A particularly informative application of nuclear and chloroplast SSRs is in assessing the amount and distribution of genetic diversity in collections of germplasm. These marker types are complementary, allowing bi-parental and maternal inheritance respectively to be studied with a very high degree of resolution. Both have been applied to genotype the complete list of potato cultivars on the current UK National List and a genotypic database has been established and installed at SASA as a reference for comparison when applications are made to register new potential varieties. Eight nuclear SSRs reveal a substantial amount of genotypic diversity and can uniquely fingerprint all of the test varieties (except two pairs of 'sports'). In contrast, analysis with cpSSRs reveals a substantial cytoplasmic bottleneck in the cultivated genepool with more than 85% of the analysed lines having an identical chloroplast haplotype. In this study, 26 chloroplast haplotypes were revealed using cpSSRs. This compares favourably with the five haplotypes identified previously using cpRFLPs and demonstrates the power of resolution of this approach.

European Union-sponsored research on conifer genomics is focused on comparative gene mapping in maritime pine and Norway spruce. Six research groups are utilising the tools of contemporary genetics to examine gene synteny, genome organisation, gene discovery based on ESTs and the development of software to support comparative mapping and QTL analysis.

The development of new technologies for functional genomics is of major significance. In this context, a new BBSRC-GAIT funded project to develop radiation-hybrid panels for plants is underway in the Unit of Barley Genomics and Breeding. The goals of this project are to develop new approaches to map gene sequences which are not dependent on the detection of polymorphism and meiotic recombination.