Genetics

John W. S. Brown, John E. Bradshaw, Gordon C. Machray & Robbie Waugh

The year 2000 was the first full year of the reorganised Genetics Division. One of the main aims of the reorganisation was to bring the different Genetics groups closer together to share and benefit more fully from the breadth of knowledge of plant genetics and molecular biology, and to make implementation of common approaches and technologies more efficient.

Success in forging closer associations is illustrated by the establishment of a new molecular marker lab in Applied Genetics, the start of a programme on the molecular characterisation of *Rubus* and *Ribes* germplasm, and the formulation of a new programme to characterise the Commonwealth Potato Collection genetically and phenotypically. All three areas are supported by increased interactions and collaborations among scientists from the different Units within Genetics. Similarly, gene expression during grain development and germination, based on cloning and sequencing of Expressed Sequence Tags (ESTs) and eventually on microarrays, run in parallel in the Units of Gene Expression and Genomics. The importance of Single Nucleotide Polymorphisms (SNPs) as simple markers for following traits in crop improvement, for mapping ESTs, and for association mapping of traits, will increase in parallel with gene discovery programmes. SNP discovery and implementation are being co-ordinated across all three research Units. Interactions were further enhanced by strategy meetings, particularly on barley and potato research, involving a number of scientists across the Division. Collaborations with the Pathology, Biochemistry and Cell Biology, and Plant, Soils and Environment Divisions in areas such as pathogen genomics, malting barley, and root structure, continue to increase as programmes develop and new programmes are initiated. The forthcoming year will also see the transfer and integration of two plant research groups from the University of Dundee into the Genetics Division and we look forward to this opportunity for closer collaboration.

Apart from a number of scientific successes described in the Unit sections below, scientists in the Genetics Division are responsible for a number of 'firsts' in plant research. These include the successful establishment of physical mapping techniques such as HAPPY and radiation hybrid mapping for the first time in plants, new exciting results in the area of homologous recombination, the characterisation of plant intron signals, and the theory and practice of linkage and QTL analysis in autotetraploids (with BioSS).

In terms of external funding, the IGF programme with SCRI, IACR, and JIC on grain development in barley (jointly funded by SEERAD and BBSRC) was a major success for the Division and Institute, and research began in mid-2000. Further significant funds were obtained for chromosome manipulation,

biodiversity, blackcurrant breeding and potato breeding. Significantly, SCRI now has breeding contracts for improved processing and quality traits from the major potato and blackcurrant processors in the UK. In addition, the international reputation of many scientists in the Genetics Division has been demonstrated by the organisation of, and outstanding contributions to, international meetings. Early in 2000, a barley genomics meeting at SCRI ('The Dundee Workshop') saw the gathering of scientists involved in cereal genomics from around the globe which catalysed the reorganisation of the International Triticeae Mapping Initiative (ITMI) and the location of the ITMI Head Office at SCRI. The 8th International *Rubus* and *Ribes* Symposium was held in Dundee in July 2001 and the XVIth EUCARPIA Congress in Edinburgh in September 2001. In parallel with the EUCARPIA Congress, there was an ITMI Workshop on Association Genetics.

The year 2001 will see large increases in the generation of DNA sequencing and genotyping data at SCRI, due to the increasing number of programmes using genomics approaches. This has placed more emphasis on automating processes to achieve higher throughput. SCRI is investing in more robotics and this investment will have to continue to meet the demands of genomics, proteomics and metabolomics. The Institute's DNA Sequencing and Genotyping Facility, run by the Unit of Gene Expression, underpins all of the DNA work of the Institute, and its importance is reflected in the increasing number of sequencing and genotyping samples, exceeding 40,000 in 2000 and likely to exceed 80,000 in 2001. Again, to meet the necessity for increased capacity, the Institute has recently purchased a new capillary DNA sequencing machine. More importantly, however, is the need for Bioinformatics services and support across the Institute. This is required to handle and analyse a variety of sources of data generated at SCRI and to mine and utilise the vast compendium of data being generated world-wide. A key challenge in the immediate future is to resource a Bioinformatics infrastructure which will meet the needs of the Institute for the next decade.

Unit of Applied Genetics During the year, we produced a document outlining our vision for plant breeding at SCRI, based on applying the new sciences of genomics, proteomics and metabolics to crop improvement, whilst recognising the different end-user communities and their needs, and hence the most appropriate commercialisation routes. Our aim is to connect genome science to plant breeding in a way that allows the novel utilisation of our unique germplasm collections and the development of innovative approaches to the production of crop plants with improved, novel and value-added properties. We want to extend our breeding targets beyond traditional products such as barley malt, and potato crisps and French fries, to improved nutritional value and anti-carcinogenic properties that can have a positive impact on human health, and to altered biosynthetic pathways that result in new products for a range of processing industries. This is already happening in our commercially-funded soft fruit breeding programmes and their underpinning research, where there is increased emphasis on fruit quality and nutritional traits. We are also conscious of the need to train a new generation of young scientists in the plant breeding methods of the future.

SCRI now interacts with the majority of major seed and processing companies in the UK in its mandate crops, and the termination in March 2000, by mutual agreement, of the Consortium Agreement (for barley and potatoes) of 1989 will allow us, through Mylnefield Research Services Ltd (MRS), to seek new marketing opportunities for wholly externally-funded plant breeding activities. The global perspective of our industrial partners in potato and blackcurrant breeding for processing will open up wider opportunities for SCRI germplasm.

During the year, after careful consideration, seed potato production was transferred from rented land at Blythbank Farm (one of the Roslin Institute farms) near West Linton, to rented land at Balruddery Farm, near Invergowrie, from 1 April 2001. Blythbank had proved a good site when potato breeding was based at Pentlandfield from 1955 to 1989. However, the move to Balruddery and modification of the potato store on site at SCRI will make our operations more efficient and convenient.

SCRI has been a leader in genetic marker development and application for a number of years and this is the area of immediate impact in our plant breeding programmes. AFLP markers linked to the *Ce* gall mite resistance gene in *Ribes nigrum* have now been confirmed, and deployment strategies are under development. The source of the major gene resistance to *Rhynchosporium* in the barley cultivar Livet has been identified through the use of molecular markers as being derived from Digger. An allele of one of SCRI's library of SSRs is associated with the resistance and is

also found in the newly recommended cultivar Pewter. This allele, therefore, appears to have value as a diagnostic marker for the resistance gene. The locus is independent of the *Rh* complex on chromosome 3H and, hence, there is the opportunity to pyramid the resistance genes. We have also established in barley that a β-amylase SSR is diagnostic for variation in enzyme thermostability. In potato, a major QTL for *Solanum vernei*-derived resistance to the white potato cyst nematode (*Globodera pallida*) was located on chromosome 5 using AFLP markers, as part of a SERAD flexible funded project between the Genomics, Applied Genetics, and Mycology, Bacteriology and Nematology Units. These AFLP markers were converted to an easy-to-use PCR marker which is now being used to explore the efficacy and efficiency of molecular marker assisted selection in the glasshouse-grown seedling generation of our potato breeding research programme.

In the longer term, plant breeding at SCRI will benefit from the expansion in genomics and functional genomics research at SCRI. For example, the barley IGF programme being led by Dr Robbie Waugh, will underpin our priorities for crop improvement for the malting and distilling industries, combined with in-built disease resistance for benign methods of control, which are highly relevant to sustainable agriculture and the environment end-users.

Unit of Gene Expression The Gene Expression Unit covers a wide range of research activities involved in understanding how plant genes and gene families are expressed at the transcriptional and post-transcriptional levels, plant transformation, and the development of new technologies for gene function analysis. Major areas of research in the Unit of Gene Expression are RNA processing, gene expression of isolated gene families (e.g. splicing factors and invertase genes), gene expression during barley germination and malting, and the development of transformation methods for targeted gene manipulation to examine gene function.

RNA processing research has concentrated on utilising the potato mini-exon splicing system to analyse *cis*- and *trans*-acting factors in plant pre-mRNA splicing. Characterisation of the mini-exon system, discovered in the Unit of Gene Expression, was published in *RNA*[1] early in 2000. The sensitivity of the system to mutation of key sequence signals has allowed the first detailed characterisation of plant intron branchpoint and polypyrimidine tracts. In collaboration with Profs. Witek Filipowicz (Basel) and Andrea Barta (Vienna), this system is now being utilised to examine the role of plant RNA binding proteins and splicing factors in intron recognition and removal. Further studies on intron splicing in collaboration with Prof. Artur Jarmolowski (Poznan), focus on AT-AC introns. These are rare introns which are removed by a secondary splicing machinery and may be important in the regulation of expression of genes containing them.

A second area of research is the analysis of small nucleolar RNA genes in *Arabidopsis*. The RNA processing lab has led plant snoRNA research for a number of years and, with the availability of the *Arabidopsis* genome sequence, it has now been possible to identify many new snoRNA genes. Of key interest is the finding that most snoRNA genes are organised in polycistronic gene clusters which are distributed throughout the *Arabidopsis* genome. Many gene clusters are duplicated and sequence changes in related genes provide information on the evolution of this gene family. The characterisation of these genes has involved mapping of modification sites on *Arabidopsis* ribosomal RNAs, and has led to the identification of a number of novel plant snoRNAs.

The analysis of the components and mechanisms of RNA processing represents basic science underpinning our understanding of plant gene expression. To exploit our knowledge of these essential processes in gene expression, an applied project to target and knockout splicing functions was established in 1998, with funding from SEERAD and BBSRC through the DTI-LINK Cell Engineering programme, and in collaboration with a UK Biotechnology company. The programme aimed to generate tolerance or resistance to nematode pests in potatoes by disrupting nematode feeding sites. One of the selected gene targets has now been shown to confer tolerance to these important pests, paving the way for further detailed studies and identification of new targets.

Besides the utility of the potato invertase mini-exon in splicing studies, invertase genes have been of interest due to their role in carbohydrate metabolism and the variety of expression patterns exhibited by this gene family. The specificity of the expression of one member of the invertase gene family was demonstrated in a joint study with the Units of Cell Biology and Plant Biochemistry on the development of the potato tuber. Discrete expression was observed in the apical hook region of the elongating stolon and associated with the

apical bud region of the mature tuber. This multidisciplinary study was published recently in the *Plant Cell*[2].

Highly parallel gene expression analyses are the focus of studies on the malting barley grain. From an ever-expanding catalogue of gene sequences (ESTs) generated at SCRI and elsewhere, we have manufactured an initial microarray of a thousand elements to examine temporal and developmental expression of this gene set during malting. The same gene set is being exploited for SNP discovery with subsequent genetic mapping of the SNPs on standard mapping populations. These sequences are also being located on a physical map of the barley genome using radiation hybrid technology. Our development of this technology in plants, the first such achievement world-wide, significantly enhances the ability to link physical and genetic maps of the barley genome, and hence will facilitate the move from phenotype to a determination of the gene(s) which control it.

Research on the manipulation of gene expression in transgenic plants has also led to exciting new discoveries in the past year. This work has developed in, and takes cognisance of, a social and regulatory framework that provides important drivers for the research. We are developing novel selection systems coupled to efficient protocols for the transformation of monocots. These will replace antibiotic and herbicide selection systems that will be phased out. For the model dicot, tobacco, we have developed highly efficient male germ line transformation protocols that dispense with the need for any selection at all. Using these protocols, we have shown for the first time that gene targeting, at a frequency allowing knockouts, knockins, and allele replacement, is feasible in higher plants. This research has major potential for application in the preparation of new generations of transgenic plants that address quality and disease-resistance attributes.

Genomics Unit Last year, two exciting new three-year programmes of work were initiated in the Genomics Unit which firmly place our research in an international context. The first of these is to develop high throughput, high resolution physical mapping in crop plants, using an innovative approach known as HAPPY mapping. Over the past year, we have demonstrated that HAPPY mapping is a powerful, physical mapping approach which could have an important role to play in several current or planned projects in the Institute. The second and larger project is the 'Cereal Community Resources' programme funded jointly by BBSRC and SEERAD through the Investigating Gene Function (IGF) initiative. This £3.1 M initiative covers parallel programmes on barley, at SCRI, and wheat, at the John Innes Centre in Norwich (Dr Graham Moore) and the Institute for Arable Crops Research (IACR) in Long Ashton (Dr Keith Edwards). The project has brought five new members of staff to the Unit and is focused on the development of 'biological resources' which will be exploited by SCRI, national and international research communities. There are three basic components in the barley programme. The first is a gene discovery project where we intend to sequence upwards of 40,000 ESTs, largely from the developing grain, and deposit the information into local UK (UKCropNet) and large international databases (dbEST, ITEC). The data we generate will complement information emerging from parallel wheat and barley gene discovery programmes being carried out in the USA, Canada, Japan, Germany, Australia, France and Finland, and will result in a world-wide collection of around 750,000 Triticeae ESTs in the public domain. These EST sequences will represent the majority of the genes present in the cereal genome and act as an entry point into detailed biological discovery. To this end, we are currently involved in discussions with the wider Triticeae genomics community about the development of a microarray expression platform which will unify and integrate barley/wheat expression studies globally. In parallel with the lab-based activities of the gene discovery programme, the second class of resources we have initiated are automated procedures for managing and analysing the vast amount of data being generated, and establishing searchable, informative databases for its long term storage and annotation. This area of Bioinformatics is becoming increasingly important to the work carried out in the Unit and the Institute, and will play a vital role in the way scientific information is utilised and exploited by both local and remote users. It is an area set to expand dramatically over the coming years. The third component of the IGF program, which will serve to link the ESTs to biological function, is the generation of chemically mutagenised barley populations. Mutant populations are a powerful resource which can be exploited to determine gene function by reverse genetics approaches through the identification of gene knockouts and allelic variants of target genes. Mutant plants, in which a target gene has been disrupted, can assist in unravelling the role that the gene plays in the biological process being studied, through phenotypic analyses and gene expression studies. A challenge for

us in the coming year will be to develop a robust and sensitive detection platform which allows us to quickly and accurately identify mutations in target genes.

Funding has also been secured to use molecular methods to study biodiversity and population genetics in a bryophyte (*Anastrophyllum joergensenii*), a pteridophyte (*Athyrium distentifolium*) and two angiosperms (*Koenigia islandica* and *Arabis petrea*) (in collaboration with the Royal Botanic Garden, Edinburgh), and an analysis of linkage disequilibrium in barley. In addition, a divisional effort has been initiated to discover, characterise and map Single Nucleotide Polymorphisms (SNPs) in the barley genome.

While these represent the directions in which the Unit is moving, existing projects are also now coming to fruition. For example, the EU-funded Ultra High Density (UHD) linkage mapping project in potato has progressed well and we now have segregation data for some 10,000 markers. In conjunction with two potato Bacterial Artificial Chromosome (BAC) libraries constructed over the last year, the UHD map is being exploited to build a genetically anchored physical map of a small region (2-10 cM in genetic length) of the potato genome. It is also beginning to impact upon our SEERAD and BPC funded tetraploid genetics projects by allowing us to anchor what were previously anonymous markers to defined positions on the potato genetic map, providing rapid assignment of chromosome segments to defined regions of the potato genome.

References

[1] Simpson, C.G., Hedley, P.E., Watters, J.A., Clark, G.P., McQuade, C., Machray, G.C. & Brown, J.W.S. (2000). *RNA* 6, 422-433.

[2] Viola, R., Roberts, A.G., Haupt, S., Gazzani, S., Hancock, R.D., Marmiroli, N., Machray, G.C. & Oparka, K.J. (2001). *Plant Cell* 13, 385-398.