Genetics

Robbie Waugh

Genetics contains around 90 staff, including four fully integrated research groups from the University of Dundee, Division of Plant Sciences. Our work focuses on three groups of crops: potatoes, soft fruits and barley, and incorporates work on model and related species. A dedicated bioinformatics capability underpins our research and we maintain strong interactions with Biomathematics and Statistics Scotland (BioSS – page 57). Genetics manages the Institute Sequencing and Microarray Facility and in 2008 established an Institute Functional Genomics Facility.

This year we welcomed new appointments in barley pathology (Mark Loosely), abiotic stress in potatoes (Ankush Prashar), informatics support for next generation sequencing (Micha Bayer) and developmental genetics in barley (Arnis Druka). We celebrated the joint appointment of John Brown as the head of the Division of Plant Sciences at the University of Dundee and have seen interactions between Genetics and Plant Sciences staff flourish through joint publications and funding. In 2008 we were reviewed twice: once for the RERAD Programme 1 mid term review and once for a rolling series of programme scientific reviews conducted by an international panel of experts. I am pleased to report that we continue to produce measurable outputs in all of our main areas of activity, including PhD-trained staff, software, publications, and new plant cultivars. Some examples of our research achievements in 2008 are summarised below.

Potatoes are rapidly becoming a staple for the world's poor and now have higher production in the developing world than the developed one. As an estimated six billion people in the world suffer from malnutrition caused by micronutrient deficiency, the micronutrient balance of potatoes has increasing importance. Gavin Ramsay and colleagues have been investigating the genetics of micronutrients in potato tubers in two projects. In the first, natural variation in the levels of tuber carotenoids is being explored through collaboration with Mark Taylor (PPFQ). *crtR-b2,* a gene encoding a β-carotene hydroxylase 2, has been confirmed as the major determinant of high carotenoid

Robbie Waugh

levels in tubers (Fig. 1). Small insertions and deletions in the gene enable the discrimination of alleles on the basis of size which is important for the application of marker assisted selection in crop improvement. Microarray experiments have identified several other genes associated with the trait including transcription factors that may coordinate pathway regulation. In the second, the genetic control of mineral content is under investigation with Philip White (EPI) and colleagues at Nottingham University. Germplasm collections and segregating populations have been explored, and genetic variation for several traits has been found. For example calcium, one of three mineral deficiencies in many diets, varies in a four fold range in a population of plants derived from Andean tetraploid potatoes. Levels of variation for other nutritionally significant minerals such as iron and zinc suggest that breeding for enhanced levels would be worthwhile.

Figure 1 Potato tubers showing different carotenoid contents.

Ingo Hein, Glenn Bryan and colleagues have identified a collection of 37 bacterial artificial chromosome (BAC) clones from a library made from a diploid potato clone that expresses high levels of quantitative resistance against late blight, and anchored them to the potato genetic map. Full length *Rpi-blb3*-like candidate resistance genes (*R* genes) have been amplified from these BAC clones, from genomic deoxyribonucleic acid (DNA) and complementary DNA (cDNA). With colleagues in the Netherlands, an *Agrobacterium*-based transient expression assay (ATTA) has been developed to functionally test *R* gene candidates in *Nicotiana benthamiana*. Candidate *R* genes are expressed alongside the positive control *Rpi-blb3* (from Dr Edwin van der Vossen) using binary expression vectors. Inoculated sites are subsequently challenged with *Phytophthora infestans* and the outcome compared between the positive control and the candidate genes (Fig. 2). Several genes are currently being cloned into a suitable ATTA vector for this functional analysis.

Along with partners in Ireland (TEAGASC), Imperial College London and the University of Dundee, Glenn Bryan, David Marshall and colleagues aim to

Figure 2 Candidate resistance genes are over-expressed in *Nicotiana benthamiana* via *Agrobacterium tumefaciens* utilising binary vectors (left) alongside a positive control, *Rpi-blb3* that governs an Avr2 specific response. Two days post Agrobacterium inoculation, areas expressing the *R* genes are subsequently challenged with *P. infestans*. *Rpi-blb3* mediates resistance towards the *P. infestans* isolate Py23 (top panel) but is ineffective towards *P. infestans* isolate 01/29 (bottom panel).

sequence potato chromosome 4 as a contribution to the International Potato Genome Sequencing Consortium. They are making use of a physical map that is genetically anchored to an Ultra High Density (UHD) genetic map that was constructed previously in 'APOPHYS', a collaborative EC FP6 project. Glenn Bryan along with colleagues from Imperial College, the University of Nottingham, and the Wellcome Trust Sanger Institute, is also involved in sequencing tomato chromosome 4. 19Mb (approximately 85% of the euchromatin) is currently completed. The tomato and potato sequences will be powerful resources for comparative genetics within the Solanaceae and for cross species gap closure in the assembled sequences.

Functional genomics studies in potato have progressed through use of the community Potato Oligo Chip Initiative (POCI) microarray platform. The role of auxin-response genes in the molecular regulation of potato somatic embryogenesis has been investigated and, in collaboration with PPFQ, key expression differences between potato germplasm differing with respect to carotenoid content and other quality traits have been identified. Current work is focused on genetic and functional analysis of a subset of the identified candidate genes which have been prioritised by knowledge of the relevant biochemical pathways (for example, terpenoid synthases, pectin methylesterases).

Exploring and understanding plant biodiversity in natural and managed agricultural systems is a major challenge. Joint development of two 1536-plex single nucleotide polymorphism (SNP)-based barley genotyping platforms (BOPA1 and BOPA2) with Tim Close in the University of California has enabled quick, cheap and accurate genotypic characterisation of any given barley variety, accession, breeding or mutant line. Using this platform Joanne Russell, Luke Ramsay, Jordi Comadran, Andy Flavell (University of Dundee) and others have focused on clarifying the relationships between different barley populations and investigating the link between sequence variation, recombination and linkage disequilibrium. They have assembled extensive collections of barley germplasm including cultivars, globally distributed landraces and wild progenitor collections. Using BOPA1 and 2, they have demonstrated that whole genome association scans can be successfully applied to locate the genes

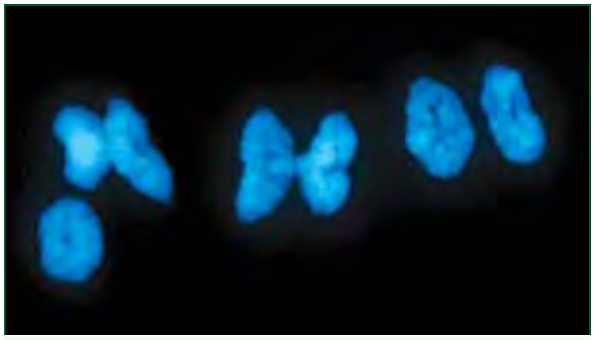

Figure 3 Barley chromosomes at meiosis.
(Courtesy of Dr. Sue Armstrong, University of Birmingham).

controlling a range of phenotypes. These studies have highlighted a severe lack of recombination in the barley genetic centromeres, which contain 30–50% of all barley genes. Consequently, these areas appear almost immune to crossing over during meiosis, impacting upon the release of genetic variation in breeding (Fig. 3). These observations have led to a significant, BBSRC-funded collaboration between Luke Ramsay, Claire Halpin (University of Dundee) Sue Armstrong and Chris Franklin (University of Birmingham) and Glyn Jenkins (University of Wales, Aberystwyth).

Arnis Druka also used the BOPA platform to characterise a unique population of plants known as the 'Bowman collection'. The collection was developed by Jerry Franckowiak at the University of North Dakota by repeated crossing followed by phenotypic selection of 977 mutant lines to the same recurrent parent – the cultivar Bowman – effectively generating a set of Nearly Isogenic Lines. These 977 lines represent the majority of the characterised morphological and developmental variation described in barley over the past 80 years. A comparison between Bowman and each individual 'Bowman line' identified polymorphic markers encompassing the mutant gene that could be

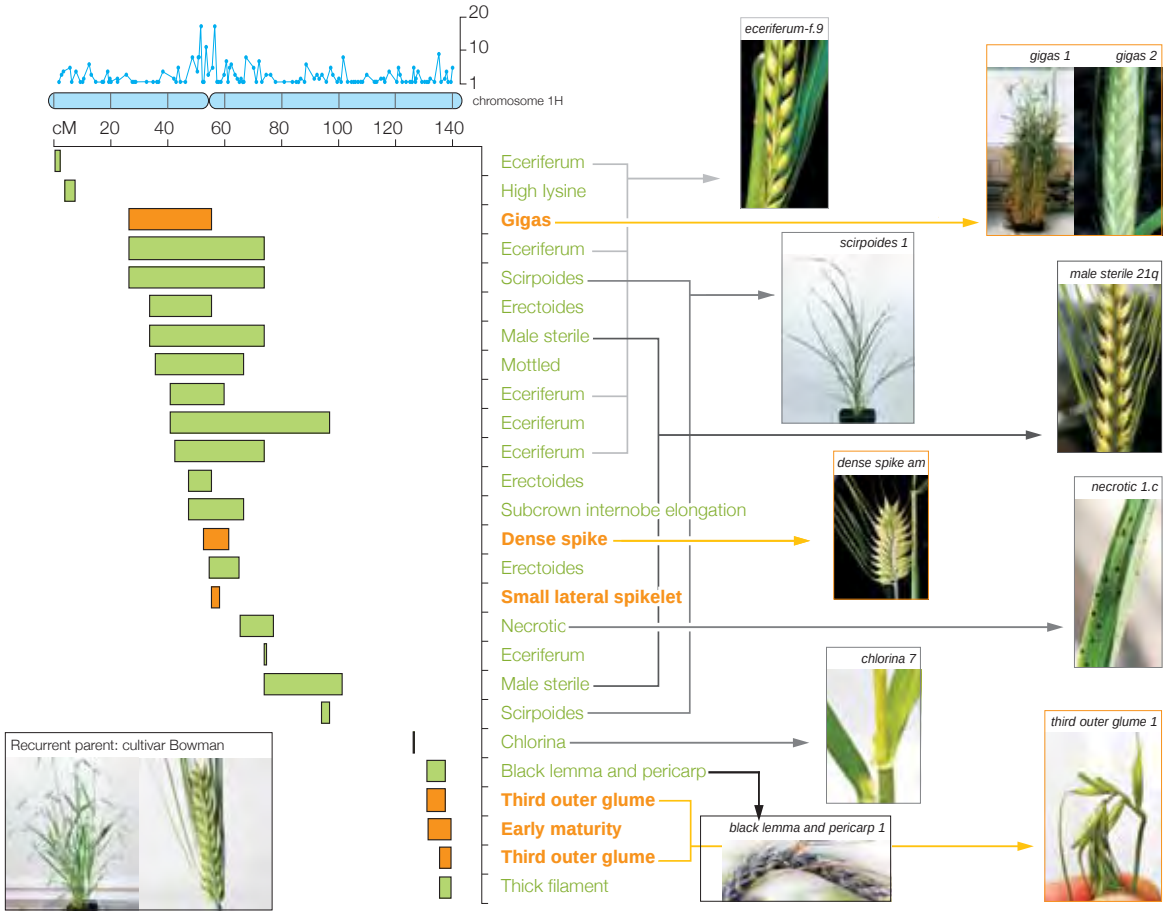

Figure 4 Twenty-six Bowman lines were selected based on polymorphisms located on barley chromosome 1H. The bars on the graph represent the genetic interval containing mutant alleles. The mutant names are highlighted in orange or green text to the right of the graph. Our targets for cloning are highlighted in orange and images of selected phenotypic variants are shown.

arranged according to their genetic map location. This exercise addressed a number of important objectives: it defined the location of the mutated genes, it assessed their distribution in relation to the gene-based SNP marker map and it identified how successful increasing rounds of backcrossing had been in eliminating donor genomes.

Quantitative trait locus (QTL) analysis of malting quality has been practiced widely, but has generally revealed different locations for the contributing genetic factors. While this has often been attributed to differences in the germplasm studied, effects of growing and processing environments are also important. Growing conditions can promote a friable grain structure, give greater likelihood of dormancy or reduced enzyme production. Similarly, processing regimes are determined by end user specifications and may emphasise different attributes of the barley grain. Stuart Swanston and Bill Thomas conducted QTL analysis on data from a population derived from a cross between the German variety Triumph and the US variety Morex grown in Scotland, Spain and the USA. QTL were distributed across all seven chromosomes. However, positive alleles for quality were observed in both parents and consistent effects were correlated with the position of genes known to be involved in cell wall biosynthesis and grain structure. Some of the progeny performed better than the original parents for several traits illustrating the potential of bringing together positive alleles from each germplasm source.

Barley leaf rust caused by *Puccinia hordei* Otth. is one of the model systems to investigate basal resistance that is inherited in a quantitative manner. In collaboration with Rients Niks in the Netherlands, Xinwei Chen has explored the molecular basis of quantitative resistance to leaf rust in the Steptoe x Morex reference population through an integrated strategy combining microarray based time course experiments, genetical genomics and genetic mapping. After collecting expression data from all lines in the population 18 hours post infection, correlation analysis between gene expression and rust resistance has revealed 128 genes significantly correlated with resistance. The most highly correlated 24 genes exclusively mapped to *Rphq*11 which had the largest effect on the phenotype. These candidates are currently under further investigation.

Resistance to some fungal diseases in raspberry is associated with morphological characters, particularly cane pubescence (Fig. 5), an epidermal cell trait determined by gene H (HH or Hh giving hairy, hh giving hairless canes). Julie Graham, Mary Woodhead and colleagues have mapped gene H to *Rubus* linkage group 2 (LG2), closely associated with resistance to cane botrytis and spur blight. Gene H is also associated with QTLs for cane spininess, and fruit ripening where development from open flowers to green/red fruit is delayed by up to two days in Hh genotypes compared to hh genotypes. Their challenge now is to identify the genes in this region and determine how they contribute to these traits. BAC libraries previously constructed from the resistant parent will be invaluable in this endeavour.

Figure 5 Hairy raspberry cane.

In collaboration with The World Agroforestry Centre with funding from Rothamsted International African Fellows Programme, Joanne Russell has assessed the level and distribution of genetic variation in the indigenous African fruit tree *Allanblackia* (Fig. 6), the subject of increased interest for edible oil production for the global food market. Until recently *Allanblackia* has been largely an overlooked wild tree and very little was known about the biology of the genus to guide conservation and use in the transition from wild harvest to cultivated production. We assessed the genetic composition of populations of five *Allanblackia* species. Data indicated significant differentiation between some taxa. Genetic relatedness between species and geographic proximity sometimes – but not always – corresponded, an observation that likely reflected complex evolutionary processes related

Figure 6 *Allanblackia*

to migration and dispersal in the genus. Our data suggest that Cameroon presents particular challenges for conservation and opportunities for domestication of the genus.

Gene expression is regulated by networks of transcriptional control and further modulated by post-transcriptional processes. One of these, alternative splicing (AS), is an extremely important, but understudied, mechanism affecting at least a third of plant genes. AS is involved in plant development, signalling and response to biotic and abiotic stresses. It generates proteins with different domains which can affect their localisation, activity and ability to interact with other proteins or substrates. John Brown, Craig Simpson and colleagues have developed a system for analysing multiple (>300) plant AS events simultaneously and are using this system to address how AS is regulated in general, its role in controlling expression in specific pathways and processes and how AS and gene expression are influenced by external stimuli such as temperature. They have shown that proteins that bind to the 5' ends of messenger ribonucleic acids (mRNAs) can affect AS of the first intron and that SR protein splicing factors regulate AS of sets of different transcripts. Approximately 15% of AS transcripts are targeted for degradation and thereby regulate gene transcript levels and alternative splicing. Their multiplex AS assay has rapidly become the foundation for numerous interactions with labs across Europe.

Post-transcriptional processing of mRNAs occurs in the nucleus of the cell (Fig. 7). The nucleolus is the major sub-compartment of the nucleus and is now recognised to have multiple functions in processing of diverse RNAs and assembly of RNA–protein complexes. After finding mRNA-associated proteins in the plant nucleolus John Brown and Craig Simpson have recently and unexpectedly discovered mRNAs in the nucleolus. Moreover, the nucleolar mRNAs appear to be enriched for improperly processed variants that are detected by an mRNA quality control mechanism and destroyed. This link between the nucleolus, aberrant mRNA processing and mRNA quality control is unique to plants and demonstrates a new function for the nucleolus.

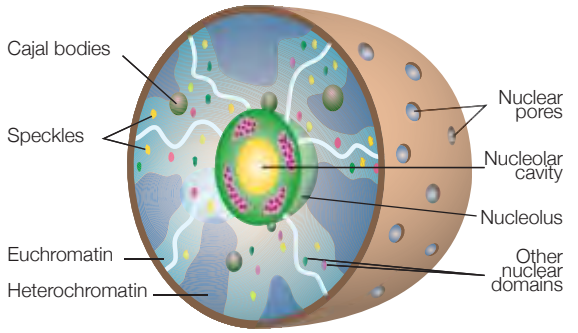

Figure 7 The plant cell nucleus.

Ian Milne and David Marshall have led the development of 'Flapjack', a new visualisation tool that facilitates the analysis of data generated by high throughput SNP genotyping technologies (Fig. 8). Its graphical displays are rendered in real time allowing for rapid navigation and comparisons between lines, markers and chromosomes. Flapjack provides a number of graphical

Figure 8 Screenshot of 'Flapjack' graphical genotyping and analysis software.

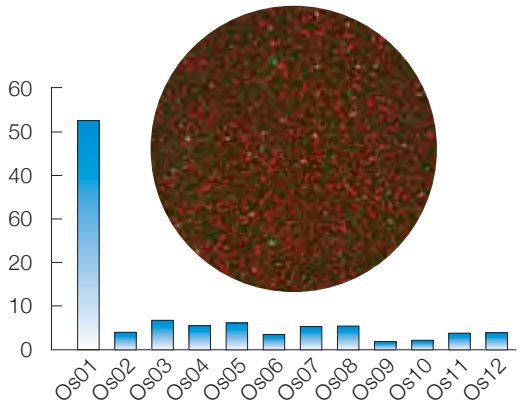

Figure 9 Barley array probed with reference barley RNA (red), which allows spot identification, and a single pool of approximately 400 BACs derived from chromosome 3B of wheat (green). Green/orange spots identify homologous genes within the BAC pool. Success of the approach is confirmed by comparison of positive probe sequences with rice orthologues, where over 50% correspond to rice chromosome 1, which exhibits high synteny with wheat chromosome 3B.

genotype views with individual alleles coloured by state, frequency or similarity to a given standard line. It supports a range of interactions with the data, including moving lines or markers around the display, inserting or deleting data, and sorting or clustering lines by either genotype similarity to other lines, or by trait scores. Any map based information such as QTL positions can be aligned against the displayed graphical genotypes to identify associated haplotypes.

Micha Bayer and Ian Milne have developed OPTIRas, a modular, web based decision support system to assist starch potato growers in the Netherlands and Northern Germany. The system presents both financial and scientific overviews of various scenarios that encompass different combinations of potato cultivars, nematode infestation levels in fields, seed rates, fertiliser regimes, and storage losses after harvest. The system has been developed for use in an internet environment, with all farmer-side interaction being performed via a web browser. Designed to be modular and extensible, the separate components share a common look and feel and are presented to the end user as part of a single, inclusive website.

The Sequencing and Microarray Facility led by Pete Hedley supports genetics based research across the Institute by providing access to state of the art genomics technologies and expertise. In 2008 we installed a BeadXpress (Illumina) platform which processes 96- or 384-plex SNP detection assays and increases genetic marker throughput while minimising costs per datapoint. We have also begun to exploit next generation sequencing technologies through links with UK university service providers. Microarray analysis continues to be used by many groups across the Institute. Highlights include the dissection of flavour and texture components in potato tubers in collaboration with PPFQ and, in a novel approach, the development of an approach that provides high-throughput anchoring of physical and genetic maps (Fig. 9).

Our research in potatoes and soft fruits feeds into downstream breeding programmes managed by SCRI/MRS. The soft fruit programme, led by Rex Brennan and Nikki Jennings, has consistently delivered commercially successful cultivars that dominate their respective industries both within the UK and worldwide. Emerging priorities for soft fruit breeding include the environmental resilience of germplasm, particularly with respect to winter chilling requirements linked to climate change, and identification of genotypes with enhanced health benefits and elevated antioxidant activities from the fruit.

Finally, in 2008, Vales Sovereign, a potato variety bred for Greenvale AP recently won the Tesco Fresh Produce Variety of the Year award, covering all fruit and vegetables. Vales Sovereign was national listed in 2003 and came from a research programme to source parental material having multiple copies of the H1 gene for resistance to the ‘golden’ potato cyst nematode. It also has excellent resistance to blackleg, blackdot and common scab. It is described as an outstanding 'all rounder'. The nationwide launch by Greenvale AP follows successful commercial trials, where it performed very well in taste panels and sold strongly in test stores.