Barley Transcriptome Resources

L. Ramsay, P. Hedley, D. Caldwell, H. Liu, N. McCallum, S. Mudie, L. Cardle, B. Harrower, G. Machray, D. Marshall & R. Waugh

As part of the SEERAD/BBSRC-funded 'Investigating Gene Function' (IGF) programme, SCRI is creating and extending genomics resources, expertise and infrastructure available to the UK cereal community and beyond. SCRI's contribution to the project is analysis on barley, with our collaborators (JIC and IACR-LARS) concentrating on wheat. These are the two cereals of most importance to the UK economy. However, due to extensive genetic similarity, information from other species such as maize and rice are also of great significance. A major part of the cereal IGF programme is the development of first class transcriptome resources.

The transcriptome is a recently coined 'omics'-derived term that is used to denote the population of mRNA transcripts in the cell, weighted by their expression levels. It essentially gives a snapshot of the genome in action within a particular cell type or tissue under particular conditions. The transcriptome represents regions of the genome that code for proteins and thus the portion of the genome of primary importance.

Comparisons of the transcriptome from different tissues, or from the same tissue at different times or under different conditions, gives an indication of the relative importance of certain genes in the expression of traits of interest. Development of transcriptome resources at SCRI is, therefore, a primary goal of the Genome Dynamics Programme.

In particular, within the IGF programme the development of barley transcriptome resources entails:

A. Creation of a wide range of cDNA libraries from various grain developmental stages and abiotic-stressed root material.

B. Generation of EST sequences from ~40,000 clones.

C. Utilisation of a high-density array containing unique selected ESTs.

This work on the barley transcriptome closely mirrors similar work on wheat at IACR-LARS within the cereals IGF project, allowing us to take advantage of the extensive genetic similarity between the two species.

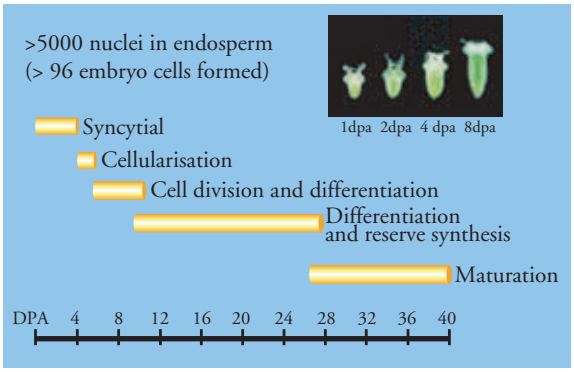

Figure 1 Stages of barley grain endosperm development.

Construction of wide range of cDNA libraries The main focus of the transcriptome programme is the molecular dissection of important events of grain development (Fig. 1) and germination, which are processes of primary biological and commercial interest (e.g. for the Scottish malting and whisky industries). For these studies we have used the cultivar Optic that is currently the most widely grown malting quality spring barley grown in the UK. A secondary focus of our work is on various abiotic-stressed and normal barley root tissue grown in hydroponic environments (Fig. 2).

Directional cDNA libraries have been generated in plasmid vectors utilising customised protocols for rapid and reliable library production. Aliquots of each

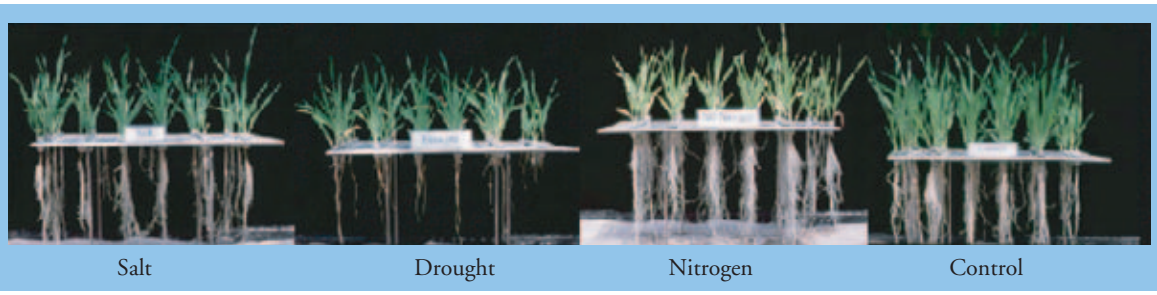

Figure 2 Comparison of abiotic stress on development of barley roots.

library have been plated, picked and immortalised using the high-throughput robotic facilities now available on-site at SCRI.

A total of 34 libraries have been produced from a range of tissues at different stages of grain development and from root and shoot tissues subject to a number of abiotic stresses.

Generation of EST sequences Clones from the stored cDNA libraries have been used for plasmid preparation and sequencing. This has taken advantage of the robotics for picking/replicating, liquid handling and capillary sequencing facilities on-site. Over 40,000 cDNA clones have now been sequenced in a single pass from the 5'-end, generating a large number of Expressed Sequence Tags (ESTs), which have been submitted to GENBANK dbEST.

The high throughput sequence generation has necessitated the development of software tools to deal with the information generated. The downstream processing, searching and databasing of the EST sequences has been automated by establishing a high throughput pipeline for the analysis of EST sequence data. This pipeline has been developed using Perl scripts to automate and 'glue together' a range of well established Sun Solaris software tools for the processing and assembly of sequence data (Fig. 3).

The 40,000 sequenced cDNA clones have been analysed and submitted to dbEST, the public access database for expressed sequence tags at NCBI. This represents a major contribution to the cereal genomics community. Placing sequence information in the public domain has already resulted in requests for specific clones and information from both national and international colleagues.

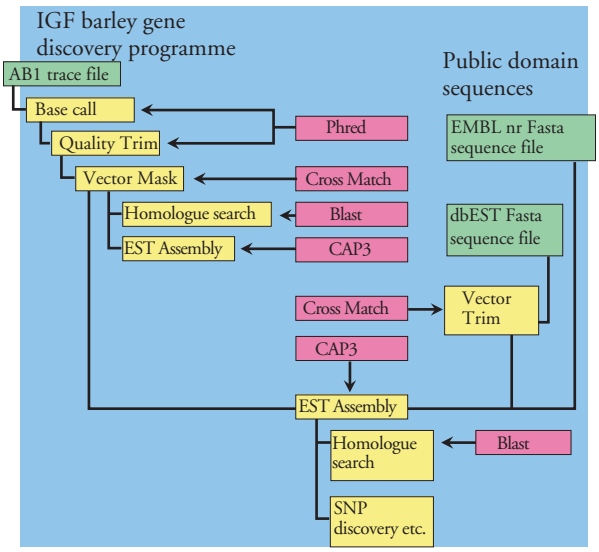

Figure 3 Sequence processing: Tasks and software.

Utilisation of a high-density array The redundancy inherent in the generation of EST data allows the barley sequences generated at SCRI and elsewhere to be assembled into a 'unigene set' that gives a more complete picture of the genes when compared to the single pass ESTs themselves. The targeted tissue approach taken within the IGF programme and the relatively low redundancy within the libraries has resulted in SCRI's sequences playing an essential role in the generation of a 'global' unigene set from all publicly available barley ESTs. Over 5,000 unique clones developed within the IGF programme has been sent to the University of Arizona (Rod Wing) for 3'-end sequencing.

This additional information will be utilised in the development of a barley Affymetrix 'gene chip' constructed with oligonucleotides designed from the assembled EST data. This microarray will allow the barley community to monitor the expression profiles of over 22,000 barley genes in any single experiment.

The use of microarray technology will augment studies already initiated using the EST information generated within the programme. In addition, insight into gene function is being gained through the use of parallel transcriptional analyses (such as cDNA-AFLPs to measure temporal and spatial regulation of expression) and RNA-based *in situ* tissue hybridisations to identify sites of transcriptional activity.

With the initial generation of transcriptome resources nearing completion, the focus of our work is now progressing to their utilisation to answer questions of relevance to the biology and cultivation of the crop.

Acknowledgements
This work was funded as part of the SEERAD/BBSRC jointly funded project "Cereal Community Resources for investigating gene function" as part of the IGF (Investigating Gene Function) Initiative.