FINAL TECHNICAL REPORT

R6454

Dr John Newbury, University of Birmingham

Executive Report

Germplasm collections underpin national and international breeding activities for all food crops. Effective and representative germplasm collections, however, tend to be very large and are difficult to maintain and organise efficiently. Recent development in molecular marker technology have allowed us to address problems such as the identification of duplicate material within such collections and the assembling of 'core collections' which represent a high percentage of all variation available within a single species. This work has been carried out using the large rice germplasm collection held at IRRI as a model system. The collection now comprises more than 100,000 samples of rice and is a major supplier of rice germplasm to research groups and breeders throughout the world. The germplasm has made important contribution to the production of new rice varieties. Pressure on germplasm distribution will further increase over the next 30 years as plant scientists strive to meet the demands for increased rice production. In this programme we have carried out research which will enhance the efficiency with which rice germplasm can be selected for use in breeding programmes and contribute to the anticipated gain of 1.5% per annum in rice yield defined as a DFID target.

We have developed a range of molecular marker techniques that allow the rapid estimation of genetic relationships between rice samples. These techniques have been compared and decisions made about the most appropriate methods for use with rice germplasm. Using these molecular marker methods, protocols for the identification of duplicate accessions and the development of core collections have been developed. The use of markers unmapped and mapped markers for such purposes have been compared and the limitations and advantages of these classes of marker have been documented. A molecular marker laboratory has been set up in the International Rice Germplasm Collection at IRRI. Associations between performance for quantitative traits and the presence/absence of specific markers have been established allowing prediction of rice field performance using marker data. Key AFLP markers have been sequenced and converted to single-locus PCR markers which are currently being used at IRRI for trial predictions of performance of rice accessions.

The results of the research have been widely disseminated during visits between collaborators between the three groups involved (at Birmingham, JIC and IRRI), at conferences and lecture tours, and in a series of publications in international journals.

Background

The demand for this work originated from the Genetic Resources Center at IRRI. The genebank at IRRI currently holds over 100,000 accessions of rice. This material is freely available to rice breeders who have made extensive use of it in the past for the production of new varieties. However, the maintenance, characterisation, evaluation and multiplication of this material is an enormous task. The scale of this task itself represents a constraint on the efficient utilisation of appropriate material by rice breeders who are striving to increase yield and contribute to the anticipated gain of 1.5% per annum in rice yield. Characterisation of all the available germplasm for key traits, especilly those which require assessment under carefully controlled conditions and/or using equipment requiring very skilled operatives, is not possible and yet a genebank manager needs to be able to provide the most appropriate material to breeders in response to their specifications. The central thesis of this research programme is that clear information on rice accessions can be efficiently obtained using newly-introduced molecular marker technologies and that this information can be used in a variety of ways to enhance the efficiency with which the collection is managed.

Market studies were carried out before commencement of the programme. These included many intensive discussions with the Head of the Genetic Resources Center at IRRI over a period of four years to define the major problems encountered in genebank management. Over the past decade, visits have also taken place to a number of other CGIAR centres housing germplasm collections in order to discuss management problems: these included the transit center of INIBAP in Leuven, ICARDA in Syria, NIAR in Tsukuva, Wakehurst Place at Kew, ICRISAT in India, etc. Because of the traditional interest of staff at Birmingham University with plant genetic resources and the location at The John Innes Centre of the UK national collections of small grain cereals and peas, together with the special cereal collections, we receive a regular flow of international visitors from crop genebanks with whom we have held discussions.

Although the results of our studies will have their most immediate impact on those who work with rice germplasm, the species serves as a clear example for those working with other crop genetic resources collections. Plant genetic resources, which represent crop biodiversity, have been actively conserved for four decades in 'genebanks' across the globe. The Consultative Group on International Agricultural Research (CGIAR) centres maintain over 500,000 accessions of more than 30 crops, while the United States Plant Germplasm System stores 380,000 samples of over 8,000 plant species. The results obtained, and protocols developed, using rice will be transferable to a large number of other genetic resources collections.

The applicants had previously obtained funding from DFID to apply some of the available techniques to rice germplasm with a view to enhancing the management of the huge collection of rice material held at IRRI. This built on progress made over the last 15 years, during which DNA-based molecular markers have been used to help address many questions in biology. More recently, the marker technologies have made use of PCR technology; this offers advantages in the small amounts of plant material required and the large number of polymorphic markers that can be obtained. Before this project was initiated, such markers were already being applied in plant research programmes either in genetic mapping studies, in measurements of diversity or for DNA fingerprinting purposes. In this programme, marker technologies have been compared for their relative efficiencies and appropriate marker types have been used to address specific problems associated with the handling of rice genetic resources.

Project Purpose

Establishment of a strategy for the establishment of core collections and the development of a core collection for the NBPGR Assam O. sativa material.

Establishment of a strategy for the recognition of duplicate accessions.

Development of molecular marker protocols for application to rice germplasm.

Establishment of the genetic basis for predictions of quantitative trait performance using molecular markers: if thi is due to linkage, then identification of key markers for morphological and physiological performance.

Development of a strategy for establishing the size and structure of wild Oryza species collections.

Comparative mapping of the genomes of key wild 'A-genome' species in order to identify whether evolutionary translocations will restrict the introgression of useful genes in wild crosses.

Research Activities

A strategy was agreed with IRRI for establishing core collections using optimal combinations of random/stratified sampling and diversity testing using molecular markers. Core collection strategies were applied to sets of 200 accessions in the UK (using marker technologies) and IRRI (using morphological characters). Appropriate protocols have been transferred to IRRI from the UK.

Marker systems appropriate for the prediction of rice quantitative traits have been transferred to IRRI where, if they prove successful, they can be incorporated into their developing rice database.

Methods for identifying duplicate accessions have been developed. In particular, the DNA Fingerprint Linkage Blocks (DFLBs) system was developed. Methods for using marker data for the establishment of core collections were developed. Technologies developed in 2.1–2.3 transferred to IRRI (but see problems with NBPGR liaison later). A range of molecular marker systems that can be used to address specific questions important in the management of rice germplasm were evaluated. Progeny of existing rice crosses were obtained and used to establish the underlying causes of associations between marker presence/absence and performance for quantitative traits. Some of the markers used to predict performance for quantitative traits among diverse accessions were demonstrated to be useful for predicting performance among segregating progeny of existing crosses.
Crosses were made between selected accessions of diverse germplasm in order to monitor the segregation of key markers and of selected quantitative traits. Markers that explain a high proportion of the variation for performance for quantitative traits were selected and tested to determine whether the associations are due to genetic linkage. Diversity was assessed within collections of some wild *Oryza* species using molecular markers. The contribution of novel alleles made to total *Oryza* genepool by wild species was not properly assessed as planned because other aspects of the programme were given priority over this section of the work. Similarly, a strategy for optimising wild species collection sizes was not developed.
O. sativa was crossed with other 'A-genome' species to allow production of skeletal maps using F2s so that translocations, inversions etc can be identified. Wild *Oryza* species carrying other genomes (e.g. E,F) were used in intra-specific crosses to allow production of skeletal maps using F2s so that translocations, inversions etc. can be identified relative to O. sativa.

Outputs
Work has been carried out on the development of core collections from large plant germplasm collections using rice as an example. Protocols have been proposed which balance the relative expense and cost of stratified eco-geographical sampling of material, field morphological characterisation, and molecular marker data. Consideration has also been given to the ways that a germplasm managers would handle a core collection especially bearing in mind that germplasm is held in trust and none of it will be discarded. Discussion of these points can be found in refs. 6, 10, 20 and (especially) 19.
A wide range of DNA-based marker techniques have been developed and applied to rice germplasm (refs. 4, 5, 8, 9, 10, 11, 17). In each case the advantages and disadvantages have been assessed particularly with regard to use at IRRI. The techniques applied include RAPD, AFLP, REPAIR, SSR and ISSR and these have been used for genetic mapping and studies of diversity in rice. AFLP markers have proved to be particularly useful because of the large number of polymorphic bands produced and the reproducibility of results (including transferability to other laboratories). Some time has been spent ensuring that AFLP markers observed in studies of one cross represent the same locus as co-migrating bands observed in other crosses. The successful demonstration that this is the case has involved the production of AFLP maps for three segregating rice populations and a doubled haploid population (ref. 15).
Time has also been spent assessing the advantages and disadvantages of using genetically mapped markers for the assessment of patterns of diversity and genetic relationships among rice accessions. This has resulted in the recognition that the use of mapped markers is sensitive to the genetic distance between the parents of the cross that allowed the original mapping (ref. 18). If parents that are closely related are used, then the use of mapped markers for revealing patterns of diversity among diverse germplasm can lead to misleading results. On the other hand, use of map-based DNA fingerprinting provides a fast method for scanning the rice genome. This has resulted in the development of DNA Fingerprint Linkage Blocks (DFLBs) and the degree of similarity or divergence within specific chromosomal regions can be calculated for rice varieties (refs. 5, 15-16). This provides a strategy for

the analysis of the pedigree of rice varieties; information gained in this way can be very useful to breeders when considering future breeding programmes. Genomic regions that are associated with a) the indica/japonica differentiation in rice, and b) segregation distortion in a doubled-haploid mapping population have also been identified (ref. 12, 14).
Research has also been directed toward confirming and explaining the statistical associations between molecular markers and performance for quantitative traits in rice. This phenomenon, originally discovered in our laboratories using RAPD markers and diverse rice germplasm, allows one to accurately predict key characteristics of plant performance using only marker data (ref. 1-3). The potential value of this strategy to the manager of an enormous germplasm collection are obvious. We have been able to show that such associations between markers and performance exist for beet as well as rice and that, in rice, there are associations between performance and a range of different marker types (RAPD, AFLP and isozymes). We have identified and mapped AFLP markers that are associated with performance for a range of quantitative traits. It has become clear that these associations are often the result of genetic linkage between marker loci and QTL. However, on many occasions there are strong associations, across diverse germplasm, between loci that are located on different chromosomes. This novel finding is consistent with the concept of 'adaptive gene complexes' and leads us to believe that blocks of DNA on different chromosomes have remained in association (presumably because of selection) over the thousands of years that rice landraces have been domesticated and differentiated into the diverse material grown today. (ref. 12, 21).
Key AFLP fragments, the presence or absence of which allows accurate prediction of performance for such characters as leaf length, days to flowering, culm number, plant height, grain width etc have been identified. In some cases, these marker fragments have been cloned and converted to simpler, single-locus PCR marker systems (ref. 22). Some of these marker systems have now been transferred to IRRI where trails are being carried out to predict culm number in uncharacterised germplasm.
Genetic diversity has been assessed in the IRRI collections of two wild species of Oryza (O. meridionalis and O. glumaepatula) along with some accessions of O. nivara and O. rufipogon. It has been discovered that some material held within the IRRI wild species collection has been mis-identified (probably by the collector). This problem can easily be identified and corrected using molecular marker data (ref. 7).

Need comments from JIC about how far we got with the following:

O. sativa was crossed with other 'A-genome' species to allow production of skeletal maps using F2s so that translocations, inversions etc. can be identified.
Wild *Oryza* species carrying other genomes (e.g. E,F) were used in intra-specific crosses to allow production of skeletal maps using F2s so that translocations, inversions etc. can be identified relative to O. sativa.

Some modification to planned activities have been necessary. The full planned interaction with NBPGR did not occur. This was because the Headship of NBPGR changed twice during the first two years of this research programme and this completely disrupted the planning processes. Because of this problem more effort was directed into other aspects of our programme.

Contribution of Outputs

Relevant publications:

P. Virk, B.V. Ford-Lloyd, M.T. Jackson, H.S. Pooni, T.P. Clemeno and H.J. Newbury (1996) Marker-assisted prediction of agronomic traits using diverse rice germplasm. In: International Rice Research Institute, Rice Genetics III. Proceedings of the Third International Rice Genetics Symposium. 16-20 October 1995. Manila, Philippines. pp. 307-316.

P. Virk, B.V. Ford-Lloyd, M.T. Jackson, H.S. Pooni, T.P. Clemeno and H.J. Newbury (1996) Predicting quantitative variation within rice germplasm using molecular markers. Heredity, 76, 296-304.

P.S. Virk, H.J. Newbury, Y. Shen, M.T. Jackson and B.V. Ford-Lloyd (1996) Prediction of agronomic traits in diverse germplasm of rice and beet using molecular markers. Plant Genome IV conference, San Diego, USA. Abstracts page 37.

B.J. Parsons, B.V. Ford-Lloyd, M.T. Jackson and H.J. Newbury (1996) Mapping of molecular markers used in diversity analysis reveals that their chromosomal positions influence patterns of diversity. Plant Genome IV conference, San Diego, USA. Abstracts page 36.

J.H. Zhu, M. Gale and G. Bryan (1996) Map-based DNA fingerprinting in rice: use of AFLP for genetic diversity studies. Plant Genome IV conference, San Diego, USA. Abstracts page 46.

J.A. Callow, B.V. Ford-Lloyd and H.J. Newbury (eds) (1997) Biotechnology and Plant Genetic Resources: Conservation and Use. Biotechnology in Agriculture Series, No. 19. CAB International.

C. Martin, A. Juliano, H.J. Newbury, B-R. Lu, M.T. Jackson and B.V. Ford-Lloyd (1997) The use of RAPD markers to facilitate the identification of Oryza species within a germplasm collection. Genetic Resources and Crop Evolution, 44, 175-183.

B. Parsons, H.J. Newbury, M.T. Jackson and B.V. Ford-Lloyd (1997). Contrasting genetic diversity relationships are revealed in rice (Oryza sativa L.) using different marker types. Molecular Breeding, 3, 115-125.

J. Zhu, M.D. Gale, S. Quarrie, M. Jackson and G.J. Bryan (1997) AFLP markers for the study of rice biodiversity. Theoretical and Applied Genetics, 96, 602-611.

B.V. Ford-Lloyd, M.T. Jackson and H.J. Newbury (1997) Molecular markers and the management of genetic resources in genebanks: a case study of rice. In 'Biotechnology and Plant Genetic resources: Conservation and Use'. Eds. J.A. Callow, B.V. Ford-Lloyd and H.J. Newbury. CAB International. pp. 103-118

H.J. Newbury and B.V. Ford-Lloyd. (1997) Estimating Genetic Diversity. In 'Plant Conservation: The *In Situ* Approach, eds. N. Maxted, B.V. Ford-Lloyd and J.G. Hawkes. Chapman and Hall. pp. 192-206.

P.S. Virk, H.J. Newbury and B.V. Ford-Lloyd (1998) Analysis of segregation distortion in an *indica-japonica* rice cross using AFLP markers. Plant and Animal Genome VI conference, San Diego, USA. Abstracts page 114

B.V. Ford-Lloyd, P.S. Virk, M.T. Jackson and H.J. Newbury (1998) AFLP analysis reveals extensive co-adaptive gene complexes in rice land-races. Plant and Animal Genome VI conference, San Diego, USA. Abstracts page 51

P.S. Virk, B.V. Ford-Lloyd and H.J. Newbury (1998) Mapping AFLP markers associated with subspecific differentiation of *Oryza* sativa and an investigation of segregation distortion. Heredity 81, 613-620.

J.H. Zhu, P. Stephenson, D.A. Laurie, W. Li, D. Tang, M.T. Jackson and M.D. gale (1999) Towards rice genome scanning by map-based AFLP fingerprinting. Molecular and general genetics, 261, 184-195.

J.H. Zhu JH, P. Stephenson, M.T. Jackson and M.D. Gale (1999) Map-based genome scanning for rice diversity study. Plant and Animal Genomics Congress VII, San Diego, US.

B.J. Parsons, H.J. Newbury, M.T. Jackson and B.V. Ford-Lloyd (1999) The genetic structure and conservation of aus, aman and boro rices from Bangladesh. Genetic Resources and Crop Evolution, In press.

P.S. Virk, H.J. Newbury, M.T. Jackson and B.V.Ford-Lloyd. (1999) Are mapped or anonymous markers more useful for assessing genetic diversity? Theoretical and Applied Genetics, In press.

M.T. Jackson, J.L. Pham, H.J. Newbury, B.V. Ford-Lloyd, and P.S. Virk. (1999) A core collection for rice: needs, opportunities, and constraints. In R.C. Johnson and T. Hodgkin (ed.) Assessing core collections. International Plant Genetic Resources Institute, Rome. In press

Other Dissemination of Results:

H.J. Newbury, BV Ford-Lloyd, P Virk, all visited ICRISAT separately in 1998.
J Zhu, visit to Muenster University for research discussions, Dec. 1998
J Zhu, visit to Sequenom Inc (Hamburg) for research discussions, Dec. 1998
J Zhu, giving lectures in Kuenming, China, Feb 1999.
J Zhu gave a course at the Institute of Crop Germplasm in Beijing (1997)

Plans for further dissemination of results;

The following manuscripts are being prepared:

P. Virk, J-H. Zhu, G. Bryan, B. Ford-Lloyd and H.J. Newbury. The establishment of core collections of rice germplasm using stratified and random sampling.
P. Virk, H.J. Newbury and B.V. Ford-Lloyd. The genetic basis for the observed associations between molecular markers and performance for quantitative traits in rice.
M. Burns, P. Virk, B.V. Ford-Lloyd and H. J. Newbury. The development of single-locus PCR-based markers for the prediction of quantitative trait performance in diverse rice germplasm.

Discussions are being held with staff at IRRI regarding future research a) on the optimisation of heterosis using mapped markers, b) on the production of genotypes with single recombinant chromosomes with overlapping substitutions to allow finer mapping of QTL and the identification of key loci for the improvement of the 'new plant type' and c) analysing differences at QTL between different rice species using the 'overlapping substitutions' approach using cytogenetic physical mapping to monitor synteny on a micro-scale.

The initial beneficiaries of this work will be rice workers at IRRI and in India. The benefits will flow from the CG centres to developing country breeders and eventually to rice farmers, small holders and consumers. Furthermore, whilst the project has been specifically concerned with rice, the techniques and strategies will be of importance to workers in all other crop germplasm collections. The germplasm will contribute an increasing proportion of the anticipated gain of 1.5% per annum in rice yield defined as a DFID target. An important use of the germplasm is to identify resistance genes and help limit the application of chemicals to the crop. The identification of stress tolerance genes will increase yield in less favourable areas of rice production.

Information has been disseminated in a large number of publications in international journals. Many of these contain detailed technical information about molecular protocols. Some have been written with the plant genetic resources community as a distinct target group (e.g. refs. 6, 10, 11 and 19).

Information has also been disseminated during the frequent visits of members of the group to each other's institutes and through visits (often including research seminars) to other institutes such as ICRISAT, ICARDA, Institute of Crop Germplasm in Beijing etc. Staff associated with this programme have also been actively involved in helping set up a molecular marker facility at the GRC at IRRI and transferring technical protocols to that laboratory.