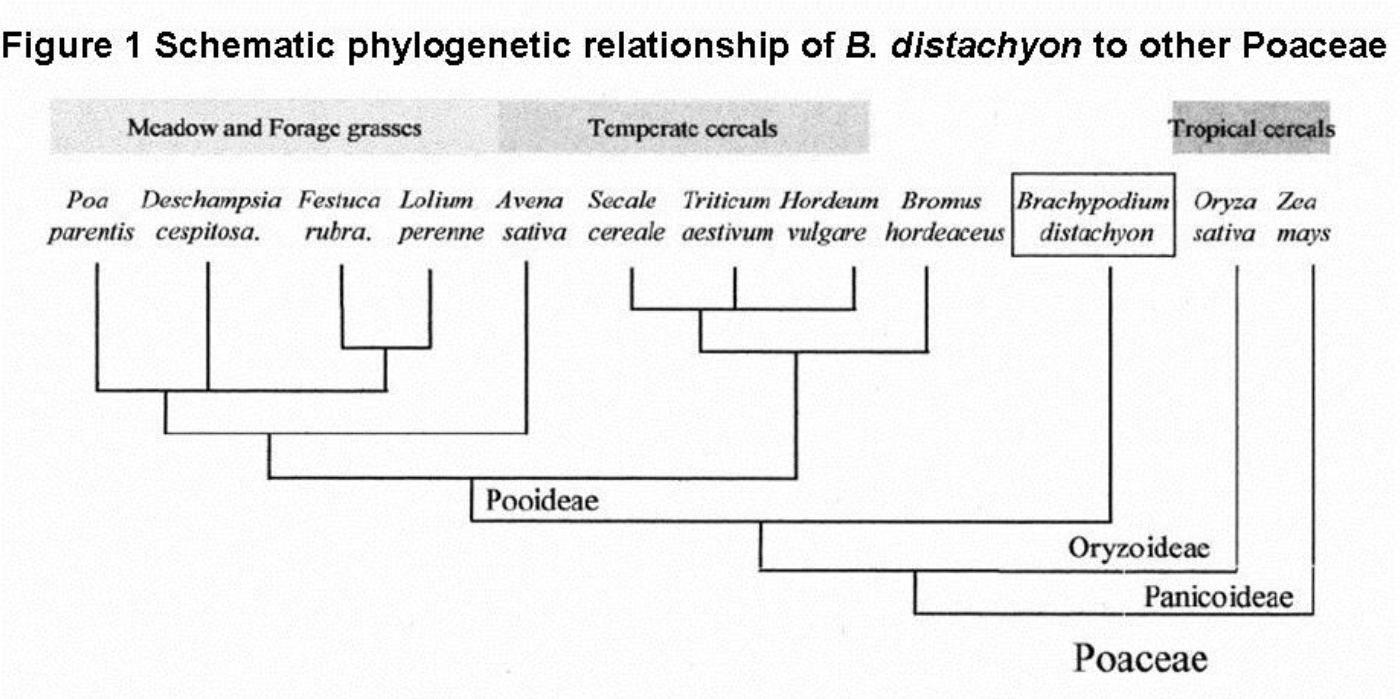
Building the *Brachypodium* – Poaceae Bridge: Anchoring *Brachypodium* genomes to rice pseudomolecules

Student: Christopher Roberts (cjr5@aber.ac.uk)
Supervisor: Dr. Luis Mur (former IBS) and Dr. Ian Armstead (formerly IGER)

1. A new model grass species *Brachypodium distachyon.* *Brachypodium distachyon* possesses the suite of traits desired in a model plant including a small genome (380 Mbp) low degree of repetitive DNA, availability of diploid ecotypes, self-fertility promoting inbreeding, transformability, small stature, and a moderately rapid life cycle. Indeed, these desirable biological features in *B. distachyon* are quite similar to those found in the dominant model plant *Arabidopsis thaliana*.

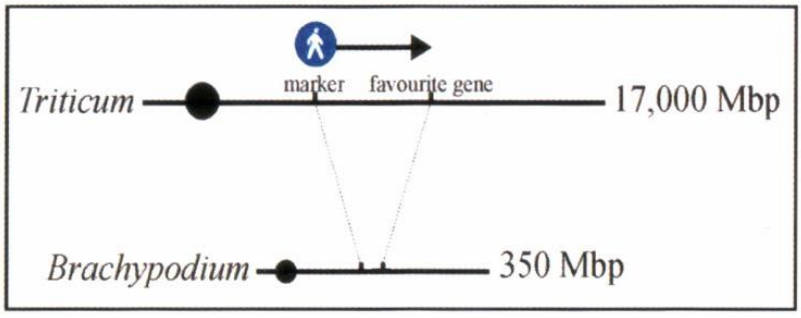
Figure 1 Schematic phylogenetic relationship of *B. distachyon* to other Poaceae



Molecular phylogenetic analyses have demonstrate that the genus *Brachypodium* diverged from the ancestralstock of Pooideae immediately prior to the radiation of the modern "core pooids" (Triticeae, Bromeae, Poeae, and Aveneae), which includes the majority of important temperate cereals and forage grasses

A molecular map of Brachypodium is essential for many purposes, but perhaps most importantly such a map will facilitate the identification of valuable genetic traits in elite commercial cereal species.

Figure 2 : The potential of *Brachypodium* as a genomic bridge



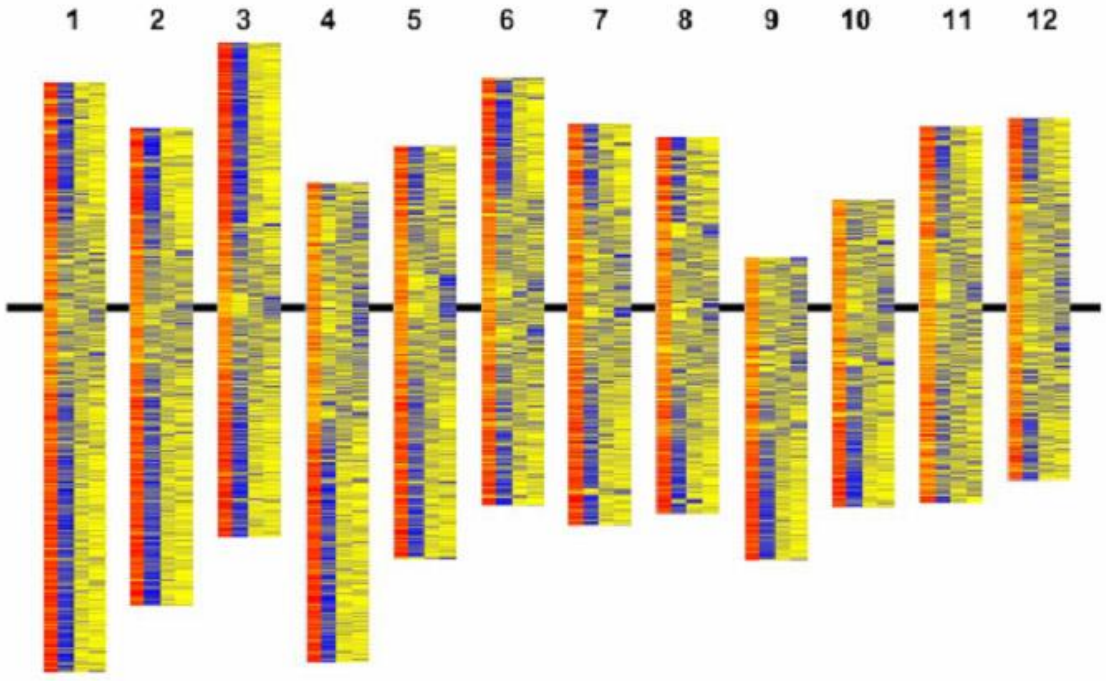
In Figure 2, a hypothetical genetic trait in wheat (*Triticum*) is separated by a considerable physical distance from a genetic marker, perhaps due to a paucity of available polymorphic markers in that region due to repetitive DNA sequences. However, in *B. distachyon* the interval between the marker and genetic traits is likely to be less due to its reduced amounts of repetitive DNA. Hence, *B. distachyon* can act as a "bridge" species

aiding genetic analysis in cereal species. The genome of *B. distachyon* accession Bd21 is currently being sequenced and this will provide the definitive description of genome organisation and the position of orthologous (i.e. conserved cross-species) markers. However, not all genetic traits found within *B. distachyon* can be contained within the sequence of a single accession. Hence, primary molecular map of Brachypodium is being developed based on an F2 population from the cross between *B. distachyon* accession Bd3–1 × Bd21, with markers primarily derived from simple sequence repeat (SSR) length polymorphisms and subsequently with a large number of single nucleotide polymorphism (SNP) markers.

2. *Brachypodium distachyon* accession typing by anchoring microarray spot hybridisation intensities on to rice – pseudogenes.

Currently only rice and *Arabidopsis thaliana*, offers a physically ordered complete genome sequence. In order to exploit these resources for analyses in other plant species; Dr. Ian Armstead has recently developed an approach where sequences of cDNA transcripts of a range of grass species are assessed for overall homologies and then anchored to positions of rice orthologous genes on pseudomolecules, corresponding to chromosomes. This made it possible to establish an overall picture of areas of rice genome where genes with differing levels of conservation are clustered. The identification of such area suggest genes/ regions of the rice genome (and by implication other genomes) under differing levels of selection as well a good areas for genetic marker conservation. This approach first entailed anchoring different types of rice gene sequences on to 12 pseudomolecules corresponding to the 12 rice chromosomes (Fig. 3).

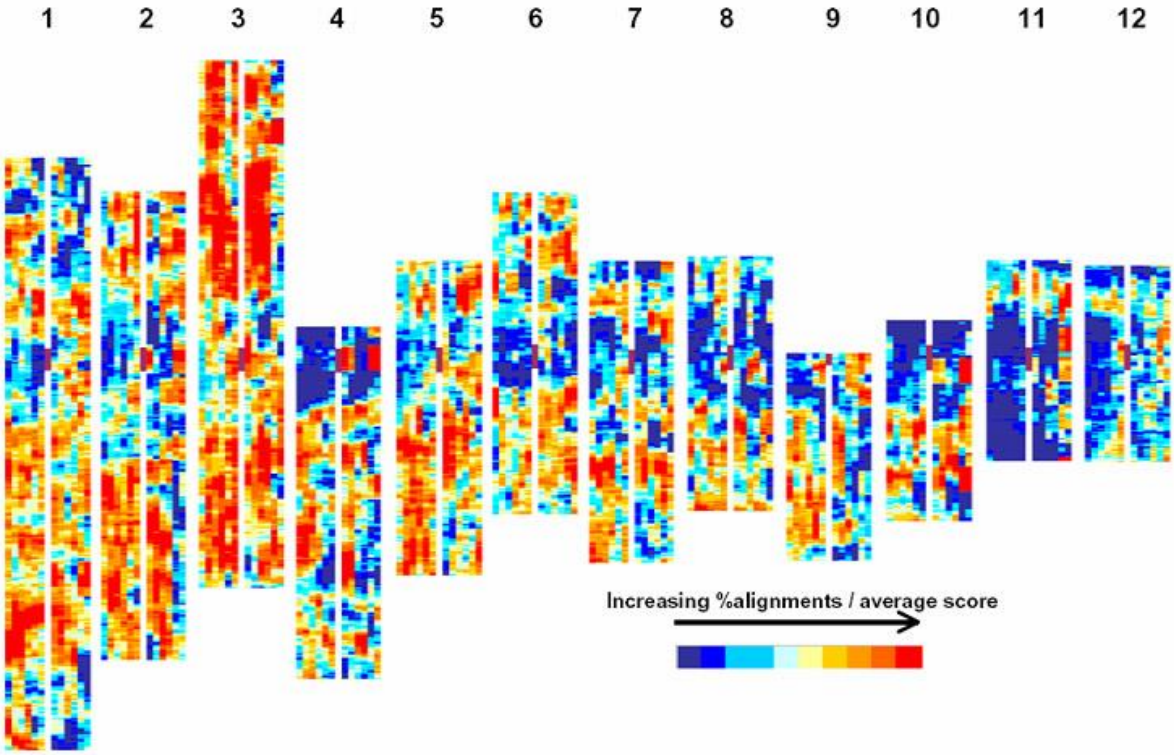
Figure 3 Distribution of differently-annotated TIGR rice loci on rice pseudomolecules



Linear order of differently annotated types of TIGR rice loci (TRL) on each of the rice pseudomolecules (1–12). For each rice pseudomolecule: column 1 (red) = combined test database significant alignments; column 2 (blue) = functionally annotated TRL or expressed protein, column 3 (blue) = hypothetical protein, column 4 (blue) = retro/transposon-related sequence. Pseudomolecules are aligned along the centromere

Subsequently cDNA sequence alignments from other species were mapped on to these pseudomolecules (Figure. 4). This revealed hot spots within the genomes where gene sequences were well or poorly conserved between species.

Figure 4: Heat maps for % sequence alignments and average scores



Colour coded expressed TIGR rice loci (for each rice pseudomolecule (1–12). For each pseudomolecule: column 1–6 = alignments between rice and databases for *Oryza sativa* (coding sequences) *Lolium perenne* (methylation filtered), *Zea mays* (methylation filtered) *Zea mays* (Transcript assemblies) *Hordeum_vulgare* (Transcript assemblies) *Glycine max* (Transcript assemblies) (Transcript assemblies) *Arabidopsis_thaliana* (coding sequences) . Pseudomolecule representations are aligned along the centromeres.

As part of their on-going *Brachypodium* project the Mur group has recently collaborated with Nottingham Arabidopsis Stock Centre (NASC) to probe rice Affymetrix microarrays with genomic DNA from *Brachypodium* accessions Bd21 and Bd3-1 (which are also being used to develop the basic molecular map; see above). These data have indicated that such rice Affymetrix arrays could be used with *Brachypodium* RNA thereby aiding transcriptomic analyses. However, these hybridisation data also show the levels of conservation between genes within the rice genome and elements with the Bd21 and Bd3-1 genomes. Further genomic insight can be gained if these rice-*Brachypodium* gene homologies can be anchored on to the rice pseudomolecules. These will indicate areas of *Brachypodium* –rice genomes which are conserved and may suggest homologous blocks within the genomes; the so-called syntenic regions.

Programme: The student will begin by anchoring the rice array features on to the rice pseudomolecules as shown in Figure. 3 for TIGR rice loci. The heat map developed for each pseudomolecules (Figure 4) was devised based on BLAST (Basic Local Alignment Search Tool),however in this project, instead of BLAST scores, array intensity scores will be used.

Depending on the speed of progress, there could be opportunities to obtain further rice Affymetrix data screened with genomes isolated from other *Brachypodium* accessions at NASC. Assuming that the project achieves a successful outcome, the work will be written up as a short communication paper- in the first instance at least, by the student.

Training value for the student.

Microarray technologies currently form an integral part of many biological research programmes. Although in this project Affymetrix microarrays have been used in an unorthodox manner, the student will become familiar with microarray and design and analysis. Training will also be provided in the novel bioinformatic techniques required to anchor microarray data on to rice pseudomolecules.

Christopher's Career Aspirations

My love of Science was ignited at an early age growing throughout my school life, but after the diversification into the three main branches, biology soon became my passion so it was an obvious choice to specialise in after A-Level.

Work appreciation placements in Year 9 of High School reinforced the idea that science was a career I wished to pursue. I arranged a shadowing position at Bretton Water Treatment Works, where at times I was responsible for chemical tests on drinking water. I also observed science in practise at Green Lane Veterinary Centre in Saltney, standing on consultations, observing operations, and was also fully involved in cleaning the equipment, surgery and animal housing.

Between June and November of last year I volunteered for the Environment Agency, which soon became a paid position. I helped with data collection and tagging of migratory Salmon and Trout at the fish trap at Chester Weir, as well as inputting data into databases at their office in Buckley.

I plan for a career in lab based research, so I am applying for a dissertation that will provide me with the relevant experience. My project of preference, the functional proteomic analyses of tsetse-trypanosome interactions will allow me to gain practical experience in techniques such as gel analysis, ELISA and SDS-PAGE.

I don't want to be too blinkered at this stage in my career, so am still experimenting with other fields of Science and Human Health and have been accepted for a 3 week full time volunteer placement in the Cardio Rehab department at the Countess of Chester Hospital later this year.

Marks to date:

Year One

Average of 81%

Year Two

Semester One

BS22120 EVOLUTION+MOLECULAR SYSTEMATICS	70
BS22320 CELLS, GAMETES & ANIMAL DEVELOPMENT	80

BS23720 APPLIED MOLECULAR BIOLOGY 92

Semester Two

Awaiting Results