

First TriticeaeGenome Project Results

1 Summary description of the project objectives

TriticeaeGenome (TG) aims at achieving significant progress in Triticeae (wheat and barley) genomics to support efficient breeding of improved varieties for the European agriculture through the accomplishment of 7 main objectives, which are:

- The construction and anchoring of physical maps from the wheat and barley group 1 and 3 chromosomes that carry a large number of important agronomic traits
- The isolation of 5 genes and QTLs underlying disease resistance, yield and quality traits in wheat and barley
- The identification and exploitation of new alleles for the isolated genes through the use of natural and mutant populations as well as wild germplasm
- The development of new markers and association panels to support the selection of new varieties that meet farmer and consumer needs
- The establishment of new bioinformatics databases and tools to structure, relate and comprehensively analyse the large scale genomics data gathered within the project
- The training of partners and collaborators in emerging technologies, the dissemination of the results of the project and the transfer of technology to industry and international collaborators.
- The coordination and integration of Triticeae genomics research at the international level

TriticeaeGenome is developed as a main contribution to the international consortia efforts in constructing physical maps of barley and hexaploid wheat for improving plant breeding, accelerating gene and QTL isolation and setting up the foundation for future genome sequencing. It aims at delivering novel information and tools to breeders and scientists for a better understanding of the Triticeae genomes organization, evolution, and function thereby, providing a better understanding of the biology of these essential crops and enabling significant improvement of their composition and characteristics to satisfy the needs of consumers, processors and producers.

2 Description of the work performed since the beginning of the project and the main results achieved so far

All 7 work packages have begun in the first period of the TriticeaeGenome project and have achieved their milestones and deliverables on time. Progresses in the different WPs are described below:

The general objective of **WP1** is to produce the genomic resources needed to construct the physical maps of the wheat chromosomes 3B, 3D, 1A and 1B and develop markers from the BAC ends sequences of the Minimal Tiling Paths for these chromosomes. The objectives of year 1 were (1) the construction of BAC libraries for chromosome 3B (version 2) and from chromosome arms 3DS, 3DL, 1BS and 1AL, (2) the fingerprinting of the 3Bv2 and 3B MTP1 BAC clones to complete the physical map of chromosome 3B as well as fingerprinting of the 3DL and 3DS BAC libraries and, (3) the assembly of the 3Bv2 and 3DL

physical maps and the design of the corresponding MTPs. The 3 objectives have been achieved with the construction ahead of schedule of all 7 chromosome specific BAC libraries (3Bv2, 3DS, 3DL, 1BS, 1BL, 1AS and 1AL) at 12 to 15X coverage and, the fingerprinting of 190'992 BAC clones from the 3Bv2, 3DL and 3DS libraries using an improved HICF SNaPshot protocol. Contig assembly was performed for chromosome 3B using the first version of the physical map (Paux et al, Science 2008) and the new fingerprints. The new 3B physical map consists of 1205 contigs with an average length of 819 kb that represent 986 Mb (99% of the chromosome) and a coverage of 19.2x. A new MTP comprised of 9314 clones has been established and will be used for rearraying and sequencing. An initial build of chromosome 3DL resulted in the assembly of 1,000 contigs with an average size of 404 kb and a total of 405 Mb that represents 99% of the chromosome. A Minimal tiling path of 5826 clones, covering 99% of the physical map was selected for rearraying and further anchoring of the physical map. Task1.4 and 1.5 that relate to the rearrangement and pooling of the BAC libraries on the basis of the results obtained in Task 1.3 as well as to the sequencing of BAC ends form the MTPs will start in the next period.

The general objective of **WP2** is to anchor the physical contigs produced in WP1 and associated project (*i.e.* barley physical map) to the genetic maps using a large number of markers and deliver anchored physical maps of the wheat and barley group 1 and 3 chromosomes. The objectives of the first year were to (1) establish a database with all potential markers for group 1 and group 3 chromosomes and (2) initiate the development of high-resolution mapping populations in wheat and barley for meiotic anchoring of the physical maps with these markers. This has been achieved through (i) a survey of existing marker resources and an evaluation of their suitability for efficient anchoring, (ii) the assessment of the usefulness of the Brachypodium genome information for anchoring information the Triticeae genome physical maps and (iii) pilot high throughput anchoring experiments using BAC pools. Barley and wheat publicly available datasets were identified and surveyed for several thousand markers. Markers for which sequence information is available will be focused on first and a marker data repository has been initiated to make the information accessible to all partners. The development of new marker resources based on repetitive DNA and gene-based sequence information has been initiated in barley and about 300 Conserved Orthologous Sets (COS) markers have been designed from sequence comparison of mapped wheat and barley EST markers to reference genome sequences of rice, sorghum and Brachypodium. Anchoring the barley physical map, which is under development (current status: 9.3 x genome coverage) in the framework of the German genome program GABI, has been initiated by screening multidimensional BAC pools for ~800 markers of chromosome 1H and 3H. To increase the resolution of meiotic mapping, new populations originating were initiated from reference parental lines for physical mapping and sequencing projects in wheat and barley (propagated up to F6 generations so far). In addition, new mapping populations based on intermated genetically distant inbred lines are explored for agronomic trait mapping and more efficient anchoring in barley. Thus, all planned objectives of WP2 have been achieved in the first year.

WP3 aims at the isolation of five genes and QTL involved in disease resistance, quality and yield traits. Target loci include fungal disease resistance genes in bread (*Qsng.sfr-3BS*) and durum wheat (*YrH52*), a QTL for grain yield in durum wheat (*QYld.idw-3B*), a Pentosan Viscosity (*PV*) QTL for a relevant quality trait in bread wheat and finally a gene (*cul4*) involved in the determination of plant architecture in barley. For the first 12

months of the project, WP3 partners have focused on establishing the genetic and molecular resources needed for fine genetic mapping of the targeted genes and QTL through (1) the construction of high-resolution mapping populations and (2) the collection of molecular markers in the target regions and the evaluation of their polymorphism in the parents of the mapping populations. At the beginning of the project, approximate map positions were known from the analysis of small populations. After a first year of activity, genetic intervals of 4 to 16 cM have been defined for all target genes/QTL and populations of 2000 to 4000 individuals are underway for high resolution mapping. Given the early stage of the project and the relatively low mapping resolution at the project start, the identification of candidate genes is still in progress. However, a few candidates have been proposed already for the PV QTL on 1BL using synteny with rice and progress is expected rapidly from the sequence information generated by INRA in the 3BS regions spanning the peaks of two of the other QTL. Fine mapping and candidate gene identification will be accelerated by the marker repository and physical maps generated from WP2. In preparation of the validation of the candidate genes, a survey of transformation and mutant population facilities has been conducted. It shows that a large panel of resources is already available for the functional validation of the candidate genes in WP3.

WP4 aims at developing markers and populations for breeding traits of interest and develop new wheat varieties that meet consumer needs (quality, environmental impact...) through breeding schemes which can relate on the availability of perfect markers (tightly linked to the genes or cloned genes themselves). For year 1, the main objectives were (1) the development of BAC pools for the whole genome 'Chinese Spring' (CS) BAC library and the development of bioinformatics tools for the generation of gene-based markers from BAC end sequences and, (2) the assessment of general adaptation and genetic diversity of a primary diversity panel of 743 winter wheat varieties from UK, France and Germany and the identify a sub-set of 384 varieties for multi-location yield components trials in 2010. All objectives have been achieved on schedule with the completion of the CS BAC pools that are now available for the project and the international community at CNRGV and, the creation of the 743 elite varieties primary panel which represents the foundation for association mapping studies planned by the project partners. This primary panel will be reduced to a panel of 384 lines using both phenotypic data gathered in the summer of 2009 and genotypic data for some SSR and major genes. Finally, an improved strategy based on the use of new sequencing technologies has been designed to efficiently generate genome specific markers.

WP5 aims at developing bioinformatics tools and resources to enhance the analysis and construction of the wheat and barley physical maps, assist in contig assembly and anchoring, analyse and annotate Triticeae genomic sequences and, integrate data and analysis tools in a platform for Triticeae genomics. During the first year, databases at URGI and HGMU were optimized to prepare their implementation with marker data that will be useful for the physical map anchoring and map-based cloning projects. Marker types and common exchange format were discussed. In addition, a new analytical framework for Linear Topology Contig (LTC) assembly has been developed and its elements have been tested on data from the wheat 3B physical map. The results show that LTC has many advantages relative to standard FPC in particular in the capacity to detect chimeric contigs and repair gaps thereby leading to longer and more robust contigs. A comparative viewer called CrowsNest has been developed by HMGU to visualize and compare both physical and

genetic maps of individual plant genomes. Finally, the semi-automated online annotation pipeline TriAnnot developed by INRA has been implemented with modules for the annotation of Tranposable Elements and a new version (5.1) has been released.

WP6 has a twofold objective: i) to provide training in emerging technological approaches and ii) to disseminate results, and transfer technology to industry. The objectives for the first 12 months of the project were: 1) to implement the training program, finalize the schedule of training events, and hold the first training course; 2) develop and implement the project website; 3) begin dissemination activities through posters, seminars, and literature. They have all been achieved on time and the TG project has a good visibility internationally and an active network of collaborators. Three new training courses are planned for 2009 and participants outside the TG project are included in each course. The TriticeaeGenome website has been successfully deployed at the URL www.triticeaegenome.eu. A total of 47 dissemination events in eight countries were held during the first year, with a total audience of more than 6500 people. Both talks and posters were presented, and a printed brochure was prepared and distributed for the first time during month 10. Potential collaborations with a WheatBiotech project in Argentina were initiated by the coordinator during a twinning workshop organised in Argentina by the EC and the Argentinean ministry of Science.

Thus, during the first year of TriticeaeGenome all the planned activities were initiated and significant progress has been achieved on time in the different work packages. The more fundamental WPs (1 and 2) have progressed rapidly and the basis for the application of all the genomic tools and resources developed are solid. This will enable the more applied WPs (3 and 4) to take off in the next reporting period. Bioinformatics tools (WP5) are advancing well to integrate the data from the project and develop new and more efficient tools for physical mapping and annotation. The TG project has already acquired a odd visibility and training is actively going on. Seven deliverables have been achieved already, all on time.

3 The expected final results and their potential impact and use

TriticeaeGenome will have several far-reaching impacts: (i) establish strategies and methods for improving genomics approaches in two of the most challenging crop genomes; (ii) develop new tools to accelerate gene isolation and support the development of molecular breeding in wheat and barley; (iii) contribute to a better understanding of traits underlying yield, quality and disease resistance; (iv) provide the foundation for future sequencing of the wheat and barley chromosomes; (v) strengthen the interactions and coordination with international collaborators and support the Triticeae networks; (vi) contribute to the transfer of know-how between research and industry and to the dissemination of information to the public; (vii) make permanent long-term improvements in social and economic cohesion on a global scale. The strategy of TriticeaeGenome is to capture immediately significant outputs for wheat, barley (and rye) breeding in parallel to continued advancement of basic research on wheat and barley. The main objective of the project is to establish a pipeline for crop improvement that starts with the construction of physical maps and provides improved tools for the development of barley and wheat varieties with enhanced composition and characteristics that meet end user needs. The ultimate goal is to help EU farmers address the challenge of delivering safe, high-quality, and health-promoting food and feed in an economically competitive, environmentally sensitive, and sustainable manner while

improving yield and stability across different environments that will be affected increasingly by climatic change. Moreover, as chromosomal regions involved in controlling phenotypes important to breeding for yield, disease resistance, nitrogen use efficiency, water use efficiency, total biomass and cell wall chemistry are identified by physical and genetic mapping, the structure of TriticeaeGenome will ensure that these findings are taken into breeding programs and will be available for wider use by the international wheat and barley communities. The new varieties that are released are still the product of classical breeding schemes and the current lack of markers does not allow effective molecular breeding for many traits of interest. In the short term, robust markers and marker trait associations obtained within TriticeaeGenome will enable plant breeders to select the best suited combinations of parental lines and facilitate efficient marker-based selection procedures by reducing the number of crosses to be carried out. In addition, gene discovery will be accelerated through association genetics approaches; while the new high-resolution populations that will be developed will allow the isolation of genes that control complex traits. Thus, the aspiration of TriticeaeGenome is to play a major role in the international efforts that aim to create the next generation breeding platforms for wheat and barley varieties to enable these essential species to contribute to the new Green Revolution for the sustainable production of one of mankind's most important foods while preserving the agricultural habitat.

Contact details

Dr. Catherine FEUILLET (Project Coordinator)

Institut National de la Recherche Agronomique (INRA)

UMR Génétique, Diversité et Ecophysiologie des Céréales (Crouël site)

234, avenue du Brézat, 63100 Clermont-Ferrand, France

Tel: +33 (0)4 73 62 46 84

Fax: +33 (0)4 73 62 44 53

E-mail: catherine.feuillet@clermont.inra.fr

Prof. Alan H. SCHULMAN (Dissemination and Technology Transfer Responsible)

MTT Agrifood Research Finland & Univ. Helsinki

MTT/BI Plant Genomics Lab, P.O. Box 56, FIN-00014 Helsinki, Finland

Tel: +358 40 768 2242

Fax: +358 91 915 9930

E-mail: alan.schulman@helsinki.fi

Emmanuelle LAGENDIJK (European Project Manager)

INRA Transfert (IT)

INRA Centre of Clermont-Ferrand (Crouël site)

234, avenue du Brézat, 63100 Clermont-Ferrand, France

Tel: +33 (0)4 73 62 49 44

Fax: +33 (0)4 73 62 44 53

E-mail : emmanuelle.lagendijk@paris.inra.fr

TriticeaeGenome project website: www.triticeaegenome.eu