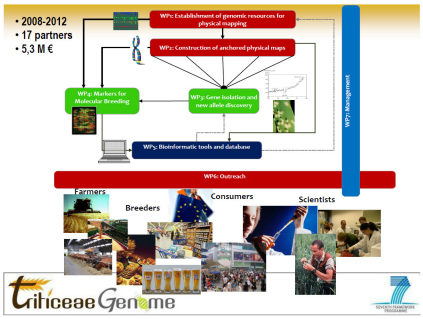
Gene-based marker development from group1 and 3 chromosomes in wheat

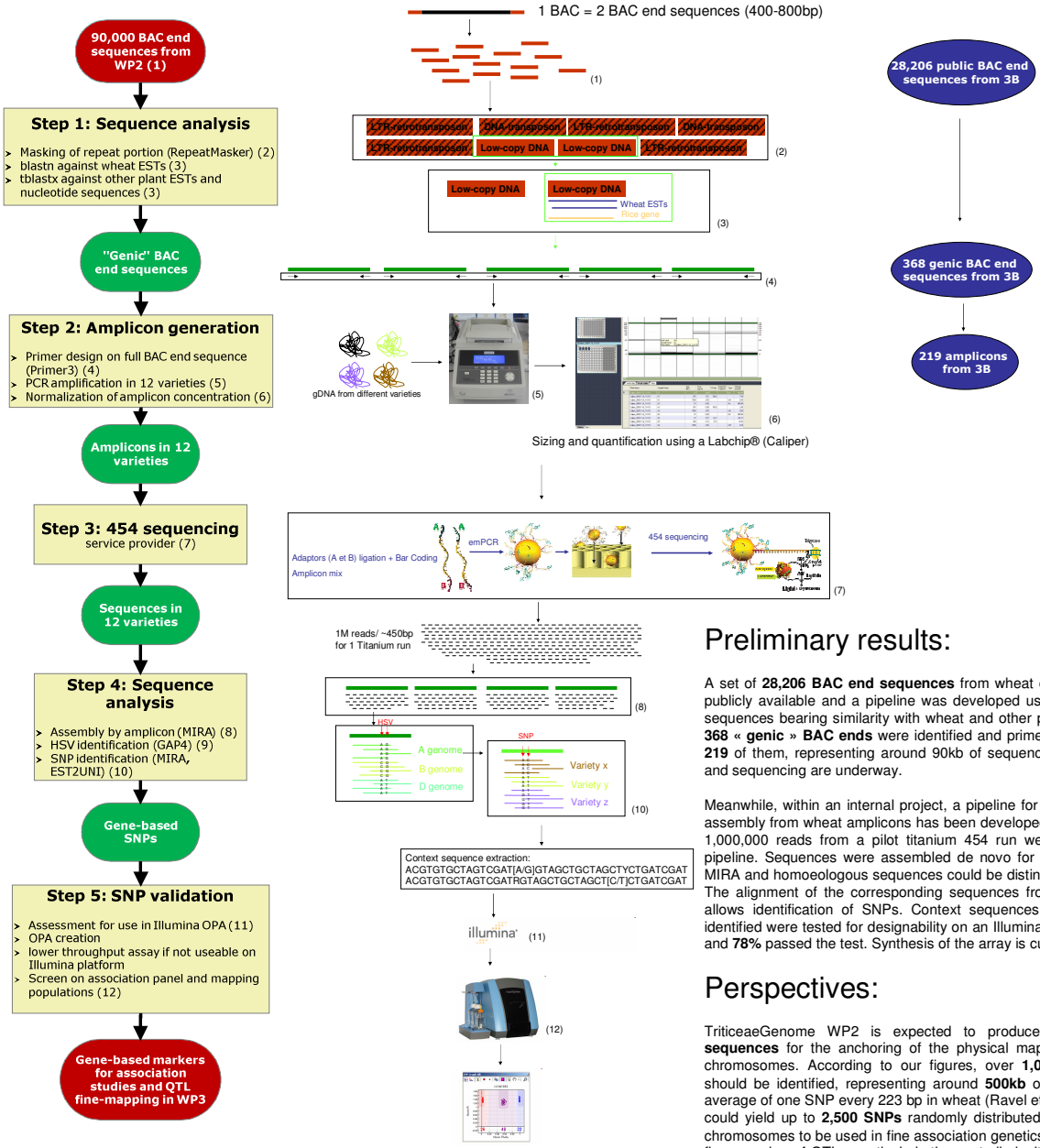
Sébastien Faure[1], Mickaël Throude[1], Jorge Duarte[1], Etienne Paux[2], Catherine Feuillet[2], Sébastien Praud[1]
[1]Biogemma, 8, rue des frères Lumière 63028 CLERMONT FERRAND cedex 2
[2]UMR INRA-UBP 1097 Génétique, Diversité et Ecophysiologie des Céréales, 234 av. du Brézet 63100 CLERMONT FERRAND

Introduction:

The Triticeae Genome program has the ambition to construct and anchor physical maps for wheat and barley group 1 and 3 chromosomes (WP1 and 2), isolate genes and QTLs underlying disease resistance, yield and quality traits (WP3) and build comprehensive tools for molecular breeding (WP4) as well as bioinformatic resources for the Triticeae (WP5).
A particularly interesting tool for molecular breeding would be a large set of gene-based markers to use for association mapping.
Within WP2, BAC end sequences will be generated for the anchoring of the physical maps. Paux et al (2006) showed that BAC end sequences could be a great source of « genic » sequences, representing 1.2% of the 11Mb of BAC end sequences obtained in this study. Therefore, we have developed a set of tools to use this resource for SNP identification.



Strategy:



Preliminary results:

A set of **28,206 BAC end sequences** from wheat chromosome 3B was publicly available and a pipeline was developed using this set to detect sequences bearing similarity with wheat and other plant ESTs. A total of **368 « genic » BAC ends** were identified and primers were designed for **219** of them, representing around 90kb of sequence. PCR amplification and sequencing are underway.

Meanwhile, within an internal project, a pipeline for 454 sequence reads assembly from wheat amplicons has been developed at Biogemma. Over 1,000,000 reads from a pilot titanium 454 run were used to train the pipeline. Sequences were assembled de novo for each amplicon using MIRA and homoeologous sequences could be distinguished using GAP4. The alignment of the corresponding sequences from each variety then allows identification of SNPs. Context sequences for 453 SNPs thus identified were tested for designability on an Illumina GoldenGate® array, and **78%** passed the test. Synthesis of the array is currently under way.

Perspectives:

TriticeaeGenome WP2 is expected to produce **90,000 BAC end sequences** for the anchoring of the physical maps of group 1 and 3 chromosomes. According to our figures, over **1,000 genic BACends** should be identified, representing around **500kb** of sequence. With an average of one SNP every 223 bp in wheat (Ravel et al 2007), this project could yield up to **2,500 SNPs** randomly distributed over group 1 and 3 chromosomes to be used in fine association genetics and to speed up the fine-mapping of QTLs, particularly those studied within TriticeaeGenome WP3.

References:
Paux et al 2006 Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B *The Plant Journal* **48**: 463-474
Ravel et al 2007 DNA sequence polymorphisms and their application to bread wheat quality *Euphytica* **158**: 331-336

"The research leading to these results has received funding from the European Community's Seventh Framework Program (FP7/ 2007-2013) under the grant agreement n°FP7-212019"