The wheat genome: swimming in an ocean of transposable elements
Frédéric Choulet, Thomas Wicker, Etienne Paux, Cyrille C. Saintenac, Pierre Sourdille, Beat Keller, Rudi Appels, Catherine C. Feuillet

To cite this version:
Frédéric Choulet, Thomas Wicker, Etienne Paux, Cyrille C. Saintenac, Pierre Sourdille, et al.. The wheat genome: swimming in an ocean of transposable elements. ICTE 2008: International Congress on Transposable Elements, Apr 2008, Saint Malo, France. hal-02820796

HAL Id: hal-02820796
https://hal.inrae.fr/hal-02820796
Submitted on 6 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
The *wheat* genome: swimming in an ocean of Transposable Elements

Frédéric CHOULET¹, Thomas WICKER², Etienne PAUX¹, Cyrille SAINTENAC¹, Pierre SOURDILLE¹, Beat KELLER², Rudi APPELS³ and Catherine FEUILLET¹

¹ UMR INRA/UBP 1095 Génétique, Diversité et Ecophysiologie des Céréales, 63100 Clermont-Ferrand, France
² Institute of Plant Biology, University Zurich, 8008 Zurich, Switzerland
³ Molecular Plant Breeding CRC, State Agriculture Biotechnology Centre, Murdoch University, Perth, WA 6150, Australia

With 17 Gb and an hexaploid structure, the bread wheat (*Triticum aestivum* L.) genome composition and organization is largely unknown and its molecular characterization remains a challenge. Our current knowledge is limited to the analysis of small fragments (150 kb) and random BAC end sequences.

Using a chromosome-specific approach, we have constructed an integrated physical map of the largest wheat chromosome, 3B (1 Gb), to enhance our knowledge of the structure, function and evolution of the wheat genome.

### Establishment of a physical map of the 3B chromosome

![Map of chromosome 3B](image)

- 57,000 fingerprinted BACs (8x coverage)
- 1,033 contigs (82% of the chromosome)
- 678 cytogenetically-anchored contigs
- 1,400 molecular markers

### Sequencing, annotation and marker development at the Rph7-Fhb1 locus

![Annotation of Rph7-Fhb1 locus](image)

- TEs represents 60% of the sequence, the majority of them being LTR-retrotransposons (72%) and CACTA family DNA transposons (20%). The curated annotation has led to detect 609 TEs and truncated derivatives showing variable levels of nested structure.
- The genes and TEs are not randomly distributed along the locus: the length of intergenic regions varies from 500 bps to 283 kb.

This project aims at sequencing and annotating the first Mb-sized contigs available from the hexaploid wheat genome. Preliminary analyses of this 3.2Mb region revealed the presence of 61 different families of repeated elements, of which 15 are represented by more than 10 copies. Furthermore, the 2 major families (*Fatima* and *Sabrina*, LTR-retrotransposons) alone represent 84 TE copies and cover 15% of the contig.

In addition, these first large wheat genome sequences are an important resource for new TE discovery. Here, about 30 annotated features are potentially unknown TEs. Deeper analyses now remain to be done to confirm, classify and characterize these new elements. Thus, the sequencing of large contigs provides additional insights, compared to random BAC sequencing, into the genome structure and composition and supports the discovery of new TEs.

The high TE content of the wheat genome can be exploited as an endless source of **molecular markers** (ISBP: Insertion Site Based Polymorphism). Here, a potential of more than 1200 ISBPs (2 markers per TE) can be designed along the 3.2 Mb.

The data generated by this project provides original and new data on the wheat genome structure, function and evolution at a scale of Mb-sized contig that has never been studied in wheat so far. In addition, it will allow the cloning of economically important disease resistance genes and QTL and enable the development of perfect markers for marker-assisted selection of these genes.

---

**The R locus:**

A cluster of at least 6 fungal disease resistance loci has been identified on the short arm of chromosome 3B in hexaploid wheat.

15 contigs (~20 Mb in total) are located in this region and 2 contigs spanning disease resistance loci are currently under sequencing and annotation. One of them represents 3.2 Mb and spans the *Rph7* and *Fhb1* resistance loci.

---

---