The Tomato Genome Project

The tomato genome has been published in Nature on May 31, 2012, culminating years of work by the Tomato Genome Consortium, a multi-national team of scientists from 14 countries.

Access the genome data on the SGN Tomato Genome Page.

The International Tomato Genome Sequencing Project was beg

un in 2004 by an international consortium including participants from Korea, China, the United Kingdom, India, the Netherlands, France, Japan, Spain, Italy and the United States. The initial approach was to sequence only the euchromatic sequence using a BAC-by-BAC approach, and in total more than 1,200 BACs have been sequenced. In 2009, a complementary whole-genome shotgun approach was initiated, which in conjunction with other data yielded high quality assemblies. The International Tomato Annotation Group (ITAG) annotates the genome builds generated by this combined sequencing approach.

We were manly involved in the production of the first assembly provided to the community the first december 2009.

We also have been involved in the improvement steps of this assembly by feeling gaps with BAC sequences and with alternative assemblies. We also built Superscaffolds using clone-end information (BAC and fosmid ends).

We also participate to the generation of the official annotation for the tomato genome provided by the International Tomato Annotation Group (ITAG), a multinational consortium, funded in part by the EU-SOL project.

Funding

This project was funded by INRA, ANR and EU-SOL project

See the INRA press release.

The initial approach description

The International Tomato Genome Sequencing Project aimed in a first approach to sequence the gene-rich euchromatic portions of the twelve tomato chromosomes. This International program involves 12 different countries among which France is devoted to the sequencing of chromosome 7. With an estimated length of 27 Mbases of gene-dense euchromatin, the coverage of chromosome 7 is expected to require the sequencing of more than 250 BACs. The strategy is to identify and sequence a minimal tiling path of BAC clones through this approximately 27 Mb euchromatin.

Seed BACs selection using FISH and ILs mapping

The initial pool of seed BACs provided by Cornell (USA) was enlarged with new putative seed BACs whose location on chromosome 7 was verified by two methods before proceeding to the sequencing step (performed by Cogenics, Grenoble, France) :

* Screening of the multi-species introgression lines (ILs) for BAC-end sequence polymorphism.
* BAC-FISH technology on mitotic chromosomes (INRA Rennes) or on pachyten meiotic chromosomes (Wageningen).

DNA pools and Macroarray filters generation

In order to increase our capacity to select positive clones for the sequencing pipe, the following new tools were generated:

* Macroarray filters were generated by the French Plant Genomic Resource Centre (CNRGV, INRA-Toulouse) for all the 3 BAC libraries available : *Hind*III, *Mbo*I and *EcoR*I and half of the fosmid library (Giovannoni, Cornell - USA).
* Three-dimensional DNA pools (3D-pools) were produced in collaboration with the CNRGV. The *Hin*dIII 3D-pools, made of half of the *Hind*III library, corresponds to 7.8 x coverage of the tomato genome while the *Mbo*I 3-D pools are made with the entire *Mbo*I library, corresponding to 7.5 x genome equivalent. This lay-out allows the screening of *Hind*III and *Mbo*I libraries respectively with 448 and 384 PCR reactions. The 3D-pools have been tested and validated by PCR screening using specific primers for chromosome 7 markers (seed BAC selection) or designed from the BAC-End Sequences (overlapping BAC selection). The results confirmed the high sensitivity and specificity of the method. These tools are now available for all the partners of the Multinational Tomato Genome Sequencing Project and can be ordered at the CNRGV.