From the model legume Medicago truncatula to alfalfa

Bernadette Julier

From the model legume *Medicago truncatula* to alfalfa

by Bernadette JULIER

Abstract: Large efforts have been devoted to *Medicago truncatula* genomics. They offer data and tools to identify *M. truncatula* genes involved in genetic variation for traits related to agronomic value in crop species. This strategy is based on the hypothesis that the same genes are involved in trait variation in both species. These tools are now used to develop alfalfa genomics in order to carry out genomic studies in the target species. Identification of genes or markers in the crop species will contribute to develop tools and theory to implement marker assisted selection in alfalfa.

Key words: genomics, marker, model species, QTL, sequence

Many efforts have been devoted to *Medicago truncatula* Gaertn. genetics and genomics since 1990. A draft sequence has been obtained after sequencing the gene-rich genome portions (euchromatin) (9). Its analysis showed that *M. truncatula* genome is more complex than expected from its short genome size. The genome has been shaped by a whole-genome duplication but also by numerous rearrangements and local duplications. These genome changes allowed the emergence of specific genes involved in rhizobial nodulation. The number of genes is over 62,000, and the genome contains numerous transposons, retrotransposons and transcription factors. Resequencing of a set of 26 accessions has given access to sequence polymorphism (1). A large diversity was found, but linkage disequilibrium was shorter than expected in a selfing species. Transcriptomics data were used to construct a genechip array (https:www.affymetrix.com/products_services/arrays/specific/medicago.affx) that can be used to detect the role of genes in specific functions and organs. In parallel, accessions collected from sites of the Mediterranean Basin were made available to the scientific community in gene banks (http://www1.montpellier.inra.fr/BRC-MTR/).

INRA, UR4, URP3F, Lusignan, France

Natural populations have been used to establish mapping populations of recombinant inbred lines, which have been mapped with microsatellite markers developed from EST sequences and used to detect quantitative trait loci (QTL) for traits having agronomic interest in crop species. Fine mapping was used to detect candidate genes explaining the QTL. Collections of tilling or transposon-tagging mutants have been established to detect new alleles of known genes or new genes involved in phenotypic variants.

From *M. truncatula* to alfalfa: strategies

Exploiting genomic knowledge from *M. truncatula* for the genetic improvement of alfalfa is not simple. Detecting genes involved in the variation for breeding traits is often needed to improve breeding programs. Two non-exclusive strategies can be devised. In the first one, candidate genes are detected in *M. truncatula* and are then tested for their effect in crop species (Figure 1). In the second one, the genomics of alfalfa is developed by taking into account the knowledge from *M. truncatula* genomics, and used to detect genes or markers involved in traits. These genes or markers become available to be incorporated into breeding programs (Figure 1).

Identification of genes in *M. truncatula*

In the first strategy, the QTL detection in mapping populations of *M. truncatula*, followed by a fine mapping step, gives access to a small genomic region. Using an *in silico* analysis of the genes annotated in the QTL region and a literature review of the genes known to be involved in the trait under study, a list of candidate genes can be established. This method was successfully used for flowering date. One of the candidate genes, a Constans-like gene, contained sequence polymorphism and showed differential expression between the two parents (7). Analysis of mutants with altered phenotypes (forward genetics), or phenotypic analysis of genotypes showing differential gene expression (reverse genetics) are other ways to detect candidate genes. Association genetics in *M. truncatula* could also help to identify genes or genomic regions involved in traits. The efficiency of association genetics will depend on the linkage disequilibrium (LD) of the population under study, which can be long in this selfing species.

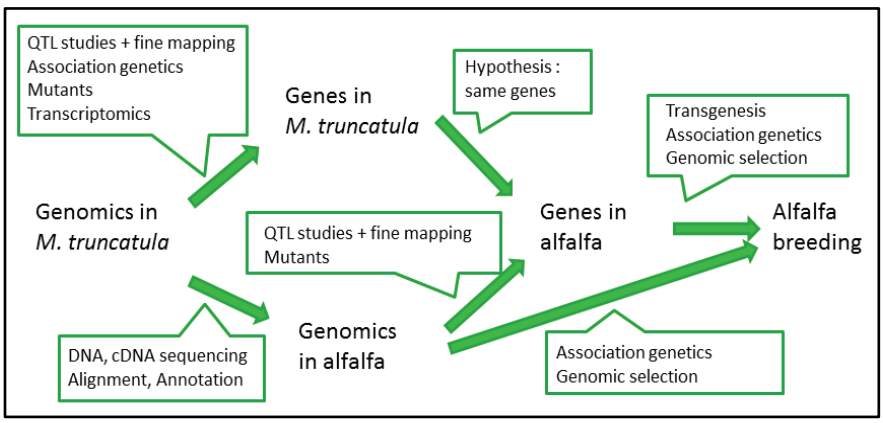

Figure 1. Strategies to transfer genomics information from *M. truncatula* to the breeding program of a legume species; several methodologies (boxes) are proposed to achieve each step

Once a gene has been identified in the model species, it remains to verify the presence of a homeologue of it in alfalfa which has an effect on the same trait. This demonstration can be achieved and exploited by different methods, such as transgenesis, introgression through backcross, or association genetics (Fig. 1). With transgenesis, the modification of expression of an allele in alfalfa is supposed to confer a new trait variation (8), eventually becoming available in a transgenic variety. Introgression through backcross has the same objective, but the allele originates from alfalfa. With association study, the demonstration results in the identification of alfalfa allele(s) that confer valuable variation, as found for a Constans-like gene (5). These alleles could then be used in a breeding program assisted by molecular markers, to increase their frequency in the synthetic varieties.

Development of alfalfa genomics

The second strategy implies the study of alfalfa genomics by DNA or cDNA sequencing using the Next Generation Sequencing (NGS) technologies (Fig. 1). Read assembly is complex because each individual possibly carries 4 different copies of each gene, in addition to the presence of gene families or repeated elements. The assembly of alfalfa sequences can be considered as a re-sequencing, with *M. truncatula* sequence acting as a reference. However, extensive sequence evolution occurred in the two species, leading to large differences in non-coding sequences between *M. truncatula* and alfalfa. Alfalfa sequencing is currently a scientific challenge that is addressed by a few groups in Europe and the USA. In the meanwhile, genetic mapping and QTL detection are feasible. Theories and software (TetraploidMap, http://www.bioss.ac.uk/knowledge/tetraploidmap/) were developed in autotetraploid species. Microsatellite markers were developed in alfalfa which mostly originated from EST of *M. truncatula* (6). These codominant markers with additional dominant markers such as AFLP or DArT (3) or HRM (4) markers are numerous enough to build maps of any mapping population. The comparison of *M. truncatula* and alfalfa maps proved that both macro-synteny and micro-synteny between the two species are high. Due to autotetraploidy, the precision of QTL position is lower in alfalfa than in a diploid species, so the subsequent use of QTL into breeding is still limited to the genetic background of the mapping population. Fine QTL mapping has not been reported in alfalfa, probably because it would require very large mapping populations and *in silico* analysis of genomics regions is not feasible yet. However, this *in silico* analysis, by using the same anchor markers in alfalfa and *M. truncatula*, can be carried out in the model species. On the whole, gene or marker detection directly performed in alfalfa (through bulk segregant analysis (BSA) or genomic selection) is scarce up to now, but the increasing availability of high throughput markers could change this situation in the future. These markers will now be mainly developed in alfalfa from sequencing and resequencing programs. However, any analysis of their function, position on maps, sequence polymorphism of their genomic regions will benefit from genomic resources in *M. truncatula*.

Perspectives

Current development of alfalfa genomics, now possible because of extensive knowledge in *M. truncatula*, offers new perspective to directly detect genes or markers in the crop species for agronomics traits. All limitations to genetic analysis in alfalfa are not yet overcome and more specifically, polyploidy and heterozygosity will lastingly reduce their efficiency. So *M. truncatula* is still a reference, mainly for all the -omics data, but also for the identification of genes involved in agronomic traits. ■

References

(1) Branca A, Paape TD, Zhou P, Briskine R, Farmer AD, Mudge J, Bharti AK, Woodward JE, May GD, Gentzbittel L, Ben C, Denny R, Sadowsky MJ, Ronfort J, Bataillon T, Young ND, Tiffin P (2011) Whole-genome nucleotide diversity, recombination, and linkage disequilibrium in the model legume *Medicago truncatula*. Proc Natl Acad Sci USA 108: E864-E870

(2) Castonguay Y, Cloutier J, Laberge S, Bertrand A, Michaud R (2005) A bulk segregant approach to identify genetic polymorphisms associated with cold tolerance in lucerne. Sapporo, Japan, 10-15 July 2004

(3) Ghesquière M, Barre P, Durand JL, Julier B, Litrico I, Maamouri A, Mournet P, Risterucci AM, Sampoux JP, Vignes H (2012) Construction of a DArT marker resource for better adapted forage crops to climate change. Proceedings, 7th International Symposium on the Molecular Breeding of Forage and Turf, Salt Lake City, USA, 4-7 June 2012, 85

(4) Han Y, Khu DM, Monteros MJ (2011) High-resolution melting analysis for SNP genotyping and mapping in tetraploid alfalfa (*Medicago sativa* L.). Mol Breed 29:489-501

(5) Herrmann D, Barre P, Santoni S, Julier B (2010) Association of a CONSTANS-LIKE gene to flowering and height in autotetraploid alfalfa. Theor Appl Genet 121:865-876

(6) Julier B, Flajoulot S, Barre P, Cardinet G, Santoni S, Huguet T, Huyghe C (2003) Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. BMC Plant Biol 3:9

(7) Pierre JB, Bogard M, Herrmann D, Huyghe C, Julier B (2011) A CONSTANS-like gene candidate that could explain most of the genetic variation for flowering date in *Medicago truncatula*. Mol Breed 28:25-35

(8) Shadle G, Chen F, Reddy MSS, Jackson L, Nakashima J, Dixon RA (2007) Down-regulation of hydroxycinnamoyl CoA: Shikimate hydroxycinnamoyl transferase in transgenic alfalfa affects lignification, development and forage quality. Phytochem 68:1521-1529

(9) Young ND, Debelle F, Oldroyd GED, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KFX, Gouzy J, Schoof H, Van de Peer Y, Proost S, Cook DR, Meyers BC, Spannagl M, Cheung F, De Mita S, Krishnakumar V, Gundlach H, Zhou SG, Mudge J, Bharti AK, Murray JD, Naoumkina MA, Rosen B, Silverstein KAT, Tang HB, Rombauts S, Zhao PX, Zhou P, Barbe V, Bardou P, Bechner M, Bellec A, Berger A, Berges H, Bidwell S, Bisseling T, Choisne N, Couloux A, Denny R, Deshpande S, Dai XB, Doyle JJ, Dudez AM, Farmer AD, Fouteau S, Franken C, Gibelin C, Gish J, Goldstein S, Gonzalez AJ, Green PJ, Hallab A, Hartog M, Hua A, Humphray SJ, Jeong DH, Jing Y, Jocker A, Kenton SM, Kim DJ, Klee K, Lai HS, Lang CT, Lin SP, Macmil SL, Magdelenat G, Matthews L, McCorrison J, Monaghan EL, Mun JH, Najar FZ, Nicholson C, Noirot C, O'Bleness M, Paule CR, Poulain J, Prion F, Qin BF, Qu CM, Retzel EF, Riddle C, Sallet E, Samain S, Samson N, Sanders I, Saurat O, Scarpelli C, Schiex T, Segurens B, Severin AJ, Sherrier DJ, Shi RH, Sims S, Singer SR, Sinharoy S, Sterck L, Viollet A, Wang BB, Wang KQ, Wang MY, Wang XH, Warfsmann J, Weissenbach J, White DD, White JD, Wiley GB, Wincker P, Xing YB, Yang LM, Yao ZY, Ying F, Zhai JX, Zhou LP, Zuber A, Denarie J, Dixon RA, May GD, Schwartz DC, Rogers J, Quetier F, Town CD, Roe BA (2011) The Medicago genome provides insight into the evolution of rhizobial symbioses. Nat 480: 520-524