

# CONSTRUCTION OF A HIGH-DENSITY GENETIC LINKAGE MAP AND QTL ANALYSIS FOR HAZELNUT BREEDING

Torello Marinoni Daniela, Valentini Nadia, Portis Ezio, Acquadro Alberto, Beltramo Chiara and Botta Roberto\*

DISAFA - Dipartimento di Scienze Agrarie, Forestali e Alimentari, Università degli Studi Torino, L.go Paolo Braccini 2 - 10095 Grugliasco (TO), Italy

\*Corresponding author: [roberto.botta@unito.it](mailto:roberto.botta@unito.it)

## Abstract

Hazelnut breeding is a slow activity due to the long lifecycle of the plant, the long time between pollination and reliable phenotype observation in the progeny, the presence of self-incompatibility and the high level of heterozygosity. Improved knowledge on the hazelnut genome and the development of marker-assisted selection would greatly facilitate breeding programs. Objective of this study was the construction of high-density genetic maps and the detection of QTL (Quantitative Trait Loci) related to phenological, vegetative and productive traits, by using an F1 progeny obtained by crossing 'Tonda Gentile delle Langhe' (female parent, hereafter TGdL) with 'Merveille de Bollwiller' (syn. 'Hall's Giant' male parent, MB). Genotyping-by-Sequencing (GBS), followed by SNP mining using a Samtools-based pipeline, yielded 7,980 SNP markers. Independent linkage maps were constructed for each parent on the basis of the double pseudo-testcross mapping strategy. The linkage map for TGdL consisted of 1,236 markers covering a total genetic length of 900.4 cM, with a mean inter-marker distance of 0.83 cM. A reciprocal translocation was detected in TGdL between two non-homologous chromosomes. The linkage map for MB was based on segregation at 1,211 markers covering a total genetic length of 899.1 cM, with a mean inter-marker distance of 0.82 cM. Plant and nut traits were recorded during five years and statistically analyzed with molecular data to identify marker/QTL associations for flowering and bud burst time, sucker habit and nut traits, such as nut and seed size, roundness index, percent kernel and ease of pellicle removal.

**Keywords:** *Corylus avellana*, breeding, SNP, GBS, MAS

## INTRODUCTION

The European hazelnut (*Corylus avellana* L.) is the most economically important nut species in the *Betulaceae* family. It is estimated that over 90% of the hazelnut crop is destined to processing. There is a strong demand of plant material for new plantings in several countries and interest in new hazelnut cultivars with higher yield and tolerance/adaptation to particular pathogens or environmental conditions is very high.

In hazelnut, the conventional breeding process is slow, due to the length of juvenile phase and the high heterozygosity level of the species. The construction of genetic linkage maps and identification of molecular markers linked to traits of interest would greatly facilitate breeding programs and the development of marker-assisted selection (MAS). Many important agronomic and quality traits, such as time of bud burst, flowering time, yield and

fruit quality are controlled by many genes; genomic regions containing these genes are known as Quantitative Trait Loci (QTL) (Collard et al. 2005).

Hazelnut is a diploid species ( $2n=2x=22$ ) with an estimated genome size of 378 Mb (<http://www.cavellanagenomeportal.com>). Observations carried out during meiosis indicate the presence of reciprocal translocations in some cultivars, including 'Tonda Gentile delle Langhe', 'Barcelona' and 'Tonda di Giffoni' (Salesses 1973; Salesses and Bonnet 1988).

Objective of the study was the construction of high-density genetic maps and the detection of QTL (Quantitative Trait Loci) related to phenological, vegetative and productive traits.

## **MATERIAL AND METHODS**

The F<sub>1</sub> progeny obtained by crossing 'Tonda Gentile delle Langhe' (female parent, hereafter TGdL) with 'Merveille de Bollwiller' (syn. 'Hall's Giant' male parent, MB) were obtained by controlled cross in February 2008 (Beltramo et al. 2016). The progeny segregated for several phenotypic traits including phenological, vegetative and productive traits. The following traits were evaluated over the 2011-2016 period: time of bud burst, male and female flowering, dichogamy, nut maturity; nut and kernel size (weight and calibre), nut shape (roundness index, RI), shell thickness, percent kernel, ease of pellicle removal after roasting (blanching).

The 275 seedlings and three individuals, obtained from rooted suckers of each of the two parents were planted in November 2010 at the campus of University of Torino, Department of Agricultural, Forest and Food Sciences (45°07'N; 7°58'E; 293 m a.s.l.). The seedlings and parents were planted at a spacing of 4 x 4 m and trained in an open vase system.

The Genotyping-by-Sequencing (GBS) approach (Elshire et al. 2011) was used to discover single nucleotide polymorphism (SNP) markers. The mapping population was a set of 213 (of the 275) F<sub>1</sub> individuals of the progeny TGdL x MB, and included 50 new individuals that were added to the 163 plants already analysed with SSR markers by Beltramo et al. (2016). In October 2014, approximately 3 µg of genomic DNA from each individual and the two parents were sent to the Genomic Diversity Facility at Cornell University - Institute of Biotechnology (USA) (<http://www.biotech.cornell.edu/brc/genomic-diversity-facility>) for GBS. Raw reads were analyzed for filtering out contaminant substrings and removing reads with poor quality ends ( $Q < 30$ ). SNP mining was conducted by adopting a Samtools-based pipeline. Independent framework linkage maps were constructed for each parent using the double pseudo-testcross mapping strategy (Weeden 1994) through JoinMap v4.0 (Van Ooijen 2006). A set of 24 microsatellite primer pairs were used to facilitate the identification of the linkage groups (LG).

The two separate parental maps were used to assign putative QTL locations by performing both the simple interval mapping (Lander and Botstein 1989) and multiple QTL mapping (Jansen and Stam 1994), procedures implemented within MapQTL v4.0 software (Van Ooijen et al. 2002).

## **RESULTS AND DISCUSSION**

### **Map construction**

The GBS approach generated a total of 46.2 Gb of DNA sequences. Single nucleotide substitutions were identified, with a frequency of one SNP every 206 nucleotides. Of the total segregating SNP markers discovered, 34% SNP markers were heterozygous in both parents, while the 32% and the 34% were heterozygous only in TGdL and MB, respectively. Only SNP markers segregating in only one of the parents were retained for map construction. In order to produce more accurate linkage maps, a further stringent selection was applied, considering only markers grouped at threshold logarithm of odds (LOD) ratio of 12.0.

Finally, a set of 1,236 and 1,211 non-redundant markers (SNPs and SSR) were used for constructing the maps of TGdL and MB, respectively. A graphical representation of the genetic maps is showed in figure 1.

A reciprocal translocation, previously observed by Salesses 1983, was detected in TGdL between two non-homologous chromosomes. The strategy reported by Farré et al. (2011) was applied for the construction of the TGdL map. We were able to group the markers into four well separated groups (at LOD>4): two of them, designated TGdL\_09a and 09b, included markers in common with MB\_09, and the other two (TGdL\_10a and 10b) with markers in common with MB\_10.

The linkage map for TGdL consisted of 13 LGs, for a total genetic length of 900.4 cM, with a mean inter-marker distance of 0.8 cM; the majority (77.3%) of map intervals were less than 1 cM; only eight gaps > 5cM were present. The linkage map for MB covered a total genetic length of 899.1 cM, with a mean inter-marker distance of 0.8 cM. The majority (77.1%) of map intervals were less than 1 cM; only five gaps > 5cM were present.

### **QTL detection**

Separated QTL analyses were performed in each season and for each parent (TGdL and MB). In this paper are reported only the results obtained for the female parent. The QTLs that explained more the 10% of the phenotypic variance (PV) are hereafter referred to as 'major' QTL. Overall, 19 major QTL was discovered for phenological traits and 8 for nut traits.

One major QTL for leaf budburst, stably expressed in all seasons (2012-2016) and responsible for 30 to 55% of PV, was discovered in the vicinity of the SSR locus AJ417975b on LG\_02. In the same region one major QTL for female flowering (44-55% of PV) and a major one for dichogamy (13-30% of PV) were also found in all seasons of observation (2013 - 2016). One stable major QTL for male flowering time was detected on LG\_10a responsible for 15 to 21% of PV (2013 - 2016), while another major QTL was detected on LG\_03 for two seasons. For time of nut maturity, a major QTL were detected for two seasons on LG\_02 and another major QTL on LG\_03 (only one season).

Regarding nut traits, carpological analysis were made for three years (2013-2015). The number of major QTLs was lower than the one detected for phenological traits. Yet, some QTLs were found stable across years also for nut traits. In fact, two stable QTL for percent kernel were detected on LG\_01 and LG\_02, as well as two stable QTL for shell thickness were detected on the same LGs (LG\_01 and LG\_02). For nut RI, two major QTL were detected on the same region on LG\_03, and a minor but very stable QTL on LG\_11. Additional major and minor QTLs were found for the following traits: nut weight, nut calibre, kernel RI, kernel calibre and blanching (table 1).

### **CONCLUSIONS**

Two well-saturated maps for TGdL and MB hazelnut cultivars were constructed using SNP markers discovered by GBS. A problem of reciprocal translocation yielding a merged LG in TGdL was solved using a statistical approach. Several solid major QTLs were established for phenological traits. A set of QTLs for nut traits was identified and will be validated through an association mapping approach. The search of genes on the scaffolds linked to the QTLs of interest has started and will be a further development of the work.

### **ACKNOWLEDGEMENTS**

The research was carried out in collaboration with and thank to the financial support of Ferrero Hazelnut Company. The research was also supported by Fondazione CRT.

## Literature cited

- Beltramo, C., Valentini, N., Portis, E., Torello Marinoni, D., Boccacci, P., Sandoval Prando, M.A., Botta, R. (2016). Genetic mapping and QTL analysis in European hazelnut (*Corylus avellana* L.). *Mol Breeding* 36, 27. <http://dx.doi.org/10.1007/s11032-016-0450-6>.
- Collard, B.C.Y., Jahufer, M.Z.Z., Brouwer, J.B., Pang, E.C.K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica* 142, 169-196.
- Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., Mitchell, S.E. (2011). A robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. *PLoS One* 6, e19379.
- Farré, A., Lacasa Benito, I., Cistué, L., de Jong, J.H., Romagosa, I., Jansen, J. (2011). Linkage map construction involving a reciprocal translocation. *Theor Appl Genet* 122, 1029–1037. <http://dx.doi.org/10.1007/s00122-010-1507-2>.
- Jansen, R.C., Stam, P. (1994). High-resolution of quantitative traits into multiple loci via interval mapping. *Genetics* 136, 1447–1455.
- Lander, E.S., Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121, 185–199.
- Salesses, G. (1973). Etude cytologique du genre *Corylus*. Mise en évidence d'une translocation hétérozygote chez quelques variétés de noisetier cultivé (*C. avellana*) a fertilité pollinique réduite. *Ann Amélior Plantes* 23, (1), 59-66.
- Salesses, G., Bonnet, A. (1998). Etude cytogénétique d'hybrides entre variétés de noisetier (*Corylus avellana*) porteuses d'une translocation à l'état hétérozygote. *Cytologia* 53, 407-413.
- Van Ooijen, J.W. (2006). JoinMap<sup>®</sup>4: software for the calculation of genetic linkage maps in experimental populations. Kyazma BV, Wageningen.
- Van Ooijen, J.W., Boer, M.P., Jansen, R.C., Maliepaard, C. (2002). MapQTL 4.0: software for the calculation of QTL positions on genetic maps. *Plant Research International*, Wageningen.
- Weeden, N. (1994). Approaches to mapping in horticultural crops. In: Gressho P. (ed) *Plant genome analysis*. CRC Press Boca Raton, pp 57-68.

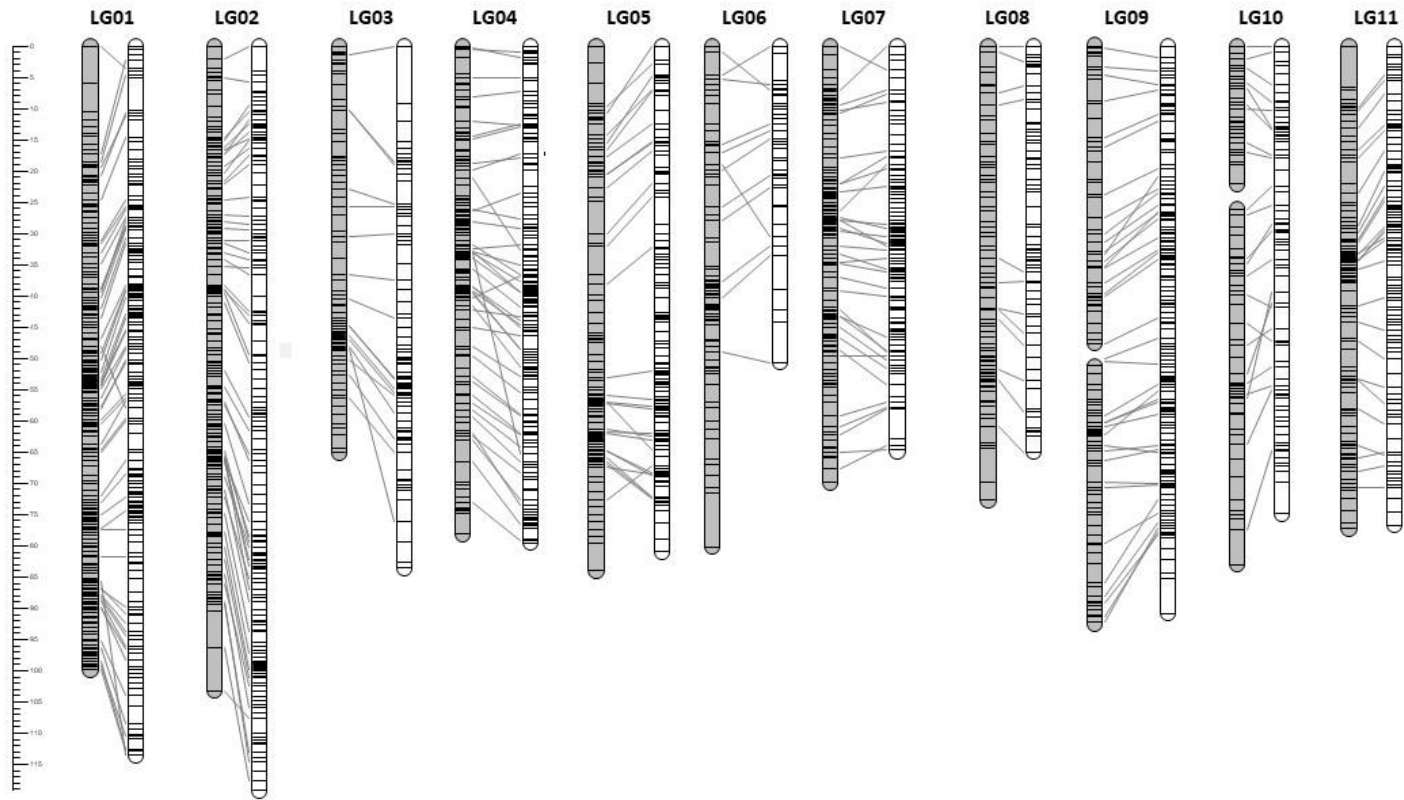


Figure 1. Genetic maps of the *Corylus avellana* cultivars 'Tonda Gentile delle Langhe' (female parent, on the left) and 'Merveille de Bollwiller' (male parent, on the right) aligned on the base of markers developed on common scaffolds.

Table 1. QTLs found in *Corylus avellana* cultivar ‘Tonda Gentile delle Langhe’ for nut weight, nut and kernel calibre, kernel roundness index and blanching: number of major and minor QTL, and linkage group where they were detected during each season of observation.

TRAIT	SEASON	MAJOR QTL	MINOR QTL	Linkage Group (LG)
Nut Weight	2013		1	LG_07
	2014	1	3	LG_07
	2015		3	LG_07
Nut Calibre	2013		3	LG_02, LG_04, LG_06
	2014		4	LG_01, LG_02, LG_04, LG_06
	2015		4	LG_01, LG_02, LG_04, LG_06
Kernel Calibre	2013		3	LG_05
	2014		2	LG_05
	2015		3	LG_05
RI Kernel	2013		3	LG_04, LG_09a, LG_11
	2014		4	LG_04, LG_08, LG_09a, LG_11
	2015		5	LG_04, LG_08, LG_09a, LG_11
Blanching	2013	1	2	LG_01, LG_02, LG_07
	2014		3	LG_01, LG_02, LG_05
	2015		2	LG_01, LG_10b