THE HISTORY OF GENOME MAPPING IN FRAGARIA SPP.

Abdel-Rahman Moustafa Abdel-Wahab MOHAMED¹, Tomasz JĘCZ², Małgorzata KORBIN²*  
¹Department of Horticulture, Faculty of Agriculture, University of Minia, Egypt  
²Department of Horticultural Plant Breeding, Laboratory of Unconventional Breeding Methods  
Research Institute of Horticulture, Skierniewice, Poland  
Received: November 25, 2014; Accepted: December 12, 2014

ABSTRACT

This overview summarizes the research programs devoted to mapping the genomes within Fragaria genus. A few genetic linkage maps of diploid and octoploid Fragaria species as well as impressive physical map of F. vesca were developed in the last decade and resulted in the collection of data useful for further fundamental and applied studies. The information concerning the rules for proper preparation of mapping population, the choice of markers useful for generating linkage map, the saturation of existing maps with new markers linked to economically important traits, as well as problems faced during mapping process are presented in this paper.

Key words: woodland strawberry, cultivated strawberry, linkage, physical map

INTRODUCTION

Genome maps, displaying the position of genes along chromosomes within the genome of an organism, are classified as genetic and physical maps (Brown 2002). In theory, both maps should provide the same information concerning chromosomal assignment, and the order of loci. However, the distances between loci measured within each map can be quite different because of diverse approaches used for genetic and physical mapping.

Genetic maps are built on the basis of pedigree analysis and classical breeding experiments. The distances between alleles are calculated according to the linkage information, and measured in centimorgans (cM) (Kole & Abbott 2008). Several statistical tools such as the LOD (Morton 1955) and the Hadane/Kosambi functions (Vinod 2011) are used in the mapping procedure for correcting the distortion of genetic distances caused by double recombination events, and proper identification of each linkage. A physical map shows actual physical distances between alleles/markers measured in base pairs (bp). This type of mapping is based on shotgun or hierarchical sequencing with final assembly of contigs derived from cosmid-based libraries or the yeast artificial chromosome (YAC) cloning vectors. In comparison to a genetic map, providing insights into the relative position of loci on chromosomes, the physical map is more “accurate” representation of the genome (Brown 2002).

Regardless of the strategy used, the essence of all mapping approaches is to localise a collection of genes/molecular markers onto their respective positions on the genome (Jiang 2013). The genome maps should define the relative positions of genetic determinants of features that are of interest, and which can serve as reference points for navigation. In consequence, the maps are important tools for identifying all genes responsible for desirable traits, comparing the genomes of different species, investigation of genome evolution, and complete genome sequencing (Dear 2001; Kole & Abbott 2008).

Over the years, numerous maps of plant genomes have been developed in many research centres. The first maps have been created for plants such as Arabidopsis thaliana (Meinke et al. 2009), rice (Oryza sativa) (Meenakshisundaram et al. 2011) and maize (Zea mays) (Ganal et al. 2011). Later on, maps of fruit plant species were also generated, including apple (Khan et al. 2012), and peach (Verde et al. 2013). The first study on

*Corresponding author:  
e-mail: malgorzata.korbin@inhort.pl
Fr"agaria spp. genome maps was undertaken at the end of 20c. but their actual development falls in recent years.

In this overview, the authors present short history of the Fr"agaria spp. genome mapping, with special emphasis on genetic interactions between diploid and polyploid representatives of the genus, spectacular successes and the numerous obstacles encountered in these research efforts.

**DIVERSITY OF Fr"AGARIA GENUS AND REFERENCE SPECIES**

The genus Fr"agaria, belonging to sub-family Potentilloideae (formerly Rosoideae), family Rosaceae and order Rosales (Eriksso et al. 2003; Potter et al. 2007; Hummer & Janick 2009), contains 24 different species with different levels of ploidy (DiMeglio et al. 2014). Out of eleven diploid species (2n = 2x = 14), ten, namely: F. bucharica Losinsk, F. chinensis Losinsk, F. daltoniana J. Gay, F. hayatae Makino, F. iinumae Makino, F. mandschurica Schlecht. Ex. J. Gay, F. nilgerrensis Schlecht. Ex. J. Gay, F. nipponica Makino, F. pentaphylla Losinsk, and F. viridis Duch. are native to Eurasia, while F. vesca L. is equally distributed in northern Eurasia and North America (Staudt 1999). F. × bifera Duch. is triploid hybrid (20 = 3x = 21), derived from the cross between F. vesca and F. viridis (Staudt et al. 2003). Tetraploid Fragaria (2n = 4x = 28) comprises five species: F. corymbosa Losinsk, F. gracilis A. Los., F. moupinensis Cardot, F. orientalis Losinsk and F. tibetica Staudt and Dichoré, distributed over East Asian regions. An occurrence of pentaploid Fragaria hybrid, F. × brinthurstii Staudt (2n = 5x = 35), derived from the cross between F. vesca and F. chiloensis (Brinthurst & Senanayake 1966), was described in California and China (Lei et al. 2005). The hexaploid (2n = 6x = 42), F. moschata Duch., was the dominant Fragaria species in Europe until the F. × ananassa appearance, due to its attractive-looking and aromatic fruits (Darrow 1966). Two octoploid species (2n = 8x = 56), F. chiloensis Mill and F. virginiana Duch., occurred natively in South and North America, respectively (Staudt 1989). The third octoploid, F. × ananassa, derives from F. chiloensis × F. virginiana cross (Staudt 1962). The decaploid F. iturupensis Staudt distribution is limited to a few colonies in Asia (Hummer et al. 2009). The newly described decaploid F. cascadenensis Hummer is probably related to octoploid F. virginiana subsp. platypetala and F. vesca subsp. bracteata (Hummer 2012).

F. vesca and F. × ananassa, are the most representative within Fr"agaria genus in the aspect of mapping. Diploid F. vesca, commonly called wild or woodland strawberry, has been proposed as a potential model plant not only for the genus, but also the entire Rosaceae family. Its genome has relatively small size (about 164 Mbp) (Akiyama et al. 2001). Additionally, F. vesca is capable to produce large progeny populations with minimum space requirements (Battey et al. 1998), and can be successfully cross-pollinated with other diploid species, resulting in highly polymorphic hybrid population (Bors 2000).

Octoploid F. × ananassa, described by Duchesne in 1766 (Staudt 1962) as an effect of accidental cross between F. chiloensis and F. virginiana, is most widely cultivated species within Fr"agaria genus, with annual world production over 4.5 million tons (FAOSTAT 2012). The broad range of distribution and cultivation of dessert strawberry is a consequence of its genotypic diversity and highly heterozygous nature, resulting in high adaptive capacity of this species (Kumar et al. 2011). Although still unclear the interpretation of phylogenetic relations between F. × ananassa and diploid Fragaria (Potter et al. 2000) allow to classify diploid F. vesca and F. iinumae (Davis et al. 2006), as well as F. viridis, F. bucharica and F. pentaphylla among ancestors of cultivated strawberry (Davis & Yu 1997; Davis et al. 2006). Complicated allopolyploid composition of strawberry genome makes it an interesting but very challenging object for genetic mapping.

**BREEDING POPULATIONS IN Fr"AGARIA GENOME MAPPING**

The choice of mapping population is a crucial step for successful genetic map constructing. The
genomes of parental forms have to display the significant level of polymorphism. On the other hand, they must be similar enough to prevent the formation of an aberrant chromosome pairing and severely suppressed recombination rates, resulting in reduced linkage distances and distorted linkage disequilibrium (Kole & Abbott 2008).

The expected degree of locus heterozygosity in mapping population can be determined by the breeding strategy (Young 2001). In genetic mapping of diploid Fragaria plants, F₂ population obtained from self-pollination or intercross of F₁ has been used most frequently (Williamson et al. 1995; Yu & Davis 1995; Davis & Yu 1997; Deng & Davis 2001; Sargent et al. 2004, 2006, 2007, 2008, 2011; Ruiz-Rojas et al. 2010; Illa et al. 2011; Zorrilla-Fontanesi et al. 2011b; Koskela et al. 2012). However, some outline maps of F. vesca and F. viridis were elaborated on the basis of backcross population (BC), generated by crossing F₁ individual(s) with inbred parent used in the initial cross (Nier et al. 2006; Cekic et al. 2001; Albani et al. 2004). Also F₃ population, obtained by crossing between F₂ individual(s), was successfully used for generating diploid Fragaria map (Yu & Davis 1995).

Heterozygosity of cultivated octoploid strawberry limits the generation of segregating populations such as F₂, F₂₃, BC or RIL, commonly used for mapping in other plant species. Therefore, all F. × ananassa maps were built on F₁ progeny obtained from inter/intraspecific cross pollination (CP) of highly heterozygous cultivars (Haymes et al. 1997; Lerceteau-Köhler et al. 2003, 2012; Rousseau-Gueutin et al. 2008; Sargent et al. 2009, 2012; Goldberg et al. 2010; Isobe et al. 2012; Gaston et al. 2013; van Dijk et al. 2014; Mohamed 2014).

The majority of segregating populations used for mapping genomes of Fragaria species contained 50 to 100 individuals (Yu & Davis 1995; Davis & Yu 1997; Haymes et al. 1997; Sargent et al. 2004, 2006, 2007; Nier et al. 2006; Rousseau-Gueutin et al. 2008; Koskela et al. 2012). Since more abundant progeny carries a greater number of meiotic recombination events, thus, it determines the resolution and accuracy of the linkage map more precisely (Paterson 1996), larger segregating populations of over 100 individuals were introduced in the study on F. vesca (Koskela et al. 2012), F. virginiana and F. chiloensis (Spigler et al. 2008, 2010; Goldberg et al. 2010) as well as F. × ananassa (Lerceteau-Köhler et al. 2003, 2012; Rousseau-Gueutin et al. 2008, 2010; Weebadde et al. 2008; Sargent et al. 2009, 2012; Isobe et al. 2012; Gaston et al. 2013; van Dijk et al. 2014). Nevertheless, as some economic factors may limit the progeny number, the bin (selective) mapping appears to be a strategy of choice (Vision et al. 2000). For instance, six seedlings selected from F. vesca × F. bucharica cross composed sufficient population to generate bin map for preliminary investigations of diploid genome (Sargent et al. 2008). However, despite the undeniable advantages of this procedure, only the use of extended set of segregating progeny allow to locate new markers in the regions of genome that were previously unmapped (Sargent et al. 2011).

MOLECULAR MARKERS USED FOR GENERATION OF MAP SKELETON

Different types of molecular markers have been used for mapping genomes of diploid and octoploid Fragaria plants. Scaffolds of the maps were built based on RAPD (Random Amplified Polymorphic DNA) (Davis & Yu 1997; Sugimoto et al. 2005), AFLP (Amplified Fragment Length Polymorphism) (Lerceteau-Köhler et al. 2003, 2005; Weebadde et al. 2008; Sargent et al. 2009), and ISSR (Inter-Simple Sequence Repeat) markers (Cekic et al. 2001; Albani et al. 2004). However, markers described above have many drawbacks, such as low reliability and insufficient level of polymorphism (RAPD), dominant inheritance (RAPD, AFLP) (Paran & Michelmore 1993), and labour-intensive data scoring (ISSR) (Cekic et al. 2001). The usefulness of some of RAPD and AFLP markers for strawberry genome mapping increased after their conversion in SCARs (Sequenced Characterized Amplified Regions) (Davis et al. 1995; Haymes et al. 2000; Albani et al. 2004; Sargent et al. 2004, 2006; Lerceteau-Köhler et al. 2005).

In parallel, microsatellites (SSRs, Simple Sequence Repeats), which are co-dominant, highly polymorphic, transferable, easy to amplify and very reliable markers (Hokanson et al. 1998), became the
convenient alternative in generation and saturation of *Fragaria* maps. Over 370 SSR markers were developed from diploid *F. vesca, F. viridis* and *F. bucharica* (James et al. 2003; Cipriani & Testolin 2004; Hadonou et al. 2004; Cipriani et al. 2006; Monfort et al. 2006; Sargent et al. 2003, 2004, 2006, 2011, 2012). Furthermore, 326 SSRs were identified in octoploid *F. × ananassa* and *F. virginiana* (Nourse et al. 2002; Ashley et al. 2003; Lewers et al. 2005; Rousseau-Gueutin et al. 2008; Sargent et al. 2008; Spigler et al. 2010). However, SSRs are basically anchored in non-coding regions of the genome, thus are not tightly linked to genes or traits of interest (Sargent et al. 2007). To overcome this obstacle, EST-SSR (Expressed Sequence Tag-SSR) primers were designed for coding regions of genes, resulting in development of polymorphic functional markers, informative not only within species, but also among genera. In total, 6125 EST-SSR markers were developed for *Fragaria* species and used for providing a linkage framework (Folta et al. 2005; Bassil et al. 2006a, b; Gil-Ariza et al. 2006; Keniry et al. 2006; Sargent et al. 2006, 2007, 2008; Spigler et al. 2008; Vilanova et al. 2008; Zorrilla-Fontanesi et al. 2011b; Isobe et al. 2012). Lately, SNP-based markers (Single Nucleotide Polymorphism), although diallelic and thus less polymorphic than SSRs, were recognised as providing valuable genetic information due to their abundance and ubiquity along the genome (Mammadov et al. 2012). SNP-derived CAPS (Cleavage Amplified Polymorphic Sequence) markers were successfully positioned on diploid *Fragaria* linkage map by Ruiz-Rojas et al. (2010). Recently, SNPs were also successfully used for mapping of male-sterility loci in diploid strawberry (Tennessen et al. 2013).

**EXISTING GENETIC MAPS OF *FRAGARIA* SPECIES**

The first genetic map of diploid *Fragaria* was constructed based on *F. vesca* ssp. *vesca* ‘Baron Solemacher’ (a highly homozygous inbred line) × *F. vesca* ssp. *americana* ‘WC6’ population (Davis & Yu 1997). The map consisted of 80 markers (75 RAPDs, 1 gene-specific, two isoenzyme and two morphological markers) gathered in 7 linkage groups (LGs), covering the total length of 445 cM. Sargent et al. (2004) created a genetic map based on *F. vesca f. semperflorens* 815 × *F. bucharica* 601 progeny. This map, known as FV × FB reference diploid *Fragaria* map, displayed 76 different markers (66 SSR, 1 SCAR, 6 gene-specific markers and 3 genes controlling morphological traits) located in 7 LGs with a total genetic distance of 448 cM. Subsequently, this reference map has been improved (Sargent et al. 2008; Illa et al. 2011; Zorrilla-Fontanesi et al. 2011b) and it latest version contains a total of 700 sequence tag sites derived from SSRs, gene-specific markers, RFLPs, SNPs and ESTs, spanning along all seven linkage groups and covering a genetic distance of 442.8 cM (Sargent et al. 2011).

The first linkage map of cultivated strawberry was created for ‘Capitola’ × CF1116 progeny with the use of 789 AFLP and two putative gene markers. The map covered 1604 cM (female) and 1496 cM (male) in 43 co-segregation groups for both parental forms. In association testing with repulsion-phase markers, 30 LGs in ‘Capitola’ genome and 28 LGs in CF1116 were detected (Lerceteau-Köhler et al. 2003). After saturation with new AFLP, SCAR and SSR markers (Rousseau-Gueutin et al. 2008), 367 loci in 28 LGs were located in the female map (total length: 2582 cM), whereas the male map displayed 440 markers, located in 26 LGs (total length: 2165 cM). The resulting map constituted the first comprehensive reference map of the octoploid strawberry. Based only on AFLP markers, the genetic map of 127 lines derived from the cross between ‘Tribute’ × ‘Honeoye’ encompassed 43 linkage groups with total genome length of 1541 cM (Weebadde et al. 2008). Spigler et al. 2008 developed the first SSR-based genetic map of octoploid *Fragaria* species (*F × virginiana*). The resultant map contained 42 linkage groups spanning genome length of 2373 cM, and displayed 210 SSR markers. Later, this map was refined by Spigler and co-workers (Spigler et al. 2010), each parental form map contained 30 linkage groups. The genetic linkage map of *F × ananassa* based on 232 × 1393 population was constructed and improved subsequently by Zorrilla-Fontanesi and co-workers (Zorrilla-Fontanesi et al. 2011a, 2012) The total genome length of the final map was 1400.1 cM, generated by allocating 363 SSRs in 39 linkage groups.
In parallel, Sargent et al. (2012) generated a genetic map of cultivated strawberry using the ‘Redgauntlet’ × ‘Hapil’ progeny. The RG × H map consisted of 549 loci (490 of which were transferable SSR and gene-specific markers), distributed among 28 LGs, spanning 2140.3 cM and covering about 91% of F. × ananassa genome. Recently, van Dijk et al. (2014) reported a genetic map of ‘Holiday’ × ‘Korona’ (H × K) population based on 186 SSRs, revealing 508 loci deployed in 28 LGs, as well. The total estimated length of the integrated map was spanning 2050 cM. The integrated strawberry map based on ‘Elsanta’ × ‘Senga Sengana’ mapping population was also created in the Research Institute of Horticulture in Skiernevice. The map contained 116 marker loci distributed in 30 LGs with total length of 1450.4 cM (Mohamed 2014).

The most saturated integrated strawberry map to date, derives from three mapping populations: 02-19 × ‘Sachinoka’, ‘Kaorino’ × ‘Akihime’, and inbred lines: 0212921 × 0212921 (Isobe et al. 2012). The map comprises 1856 EST-SSR loci distributed among 28 LGs, spanning 2364.1 cM.

**GENES AND QTL REGIONS MAPPED TO FRAGARIA LGS**

Identification of major gene loci and QTL regions linked with important physiological and economic traits, such as disease resistance, fruit quality, photoperiodic flowering, and sterility, is one of the most significant objectives of genetic linkage mapping.

The first markers linked to disease resistance gene were reported by Haymes et al. (1997). The authors identified locus of a single dominant gene conferring resistance to Phytophthora fragariae var. fragariae, the causal agent of red stele root in Fragaria ssp., and 7 markers linked to this gene (Haymes et al. 2000). Strawberry resistance to anthracnose caused by Colletotrichum acutatum pathogenicity group 2, was identified as trait determined by a single Rca2 gene (Denoyes-Rothan et al. 2005) and two SCAR markers Rca2_240 and Rca2_417, linked to dominant resistance allele of Rca2 gene, were detected (Lerceteau-­Köhler et al. 2005). Cekic et al. (2001), working on population of F. vesca × F. vesca f. semperflorens, characterised two markers linked to a gene controlling seasonal flowering (SFL). Markers linked to SFL-like genes as well as QTLs connected with flowering were also identified and mapped on different segregating populations by other authors (Sargent et al. 2004; Sugimoto et al. 2005; Weebadde et al. 2008; Gaston et al. 2013), together with non-­running trait (r) (Davis & Yu 1997; Albani et al. 2004; Sargent et al. 2004; Gaston et al. 2013).

The location of genes determining fruit quality traits has been so far a major scope of Fragaria mapping projects. A close linkage (1.1 cM) of the yellow fruit colour (c) locus to shikimate dehydrogenase (SDH) and flavonan 3-­hydroxylase (F3H) genes was reported in diploid Fragaria on LG I (Williamson et al. 1995; Davis & Yu 1997; Deng & Davis 2001). Some additional genes of known function, such as chalcone synthase (CHS), chalcone isomerase (CHI), regulatory gene of anthocyanin biosynthetic pathway (RAN) (Deng & Davis 2001), as well as dihydroflavonol 4-­reductase (DFR), anthocyanidin synthase (ANS) (Deng & Davis 2001; Sargent et al. 2007), phosphoglucoisomerase (PgI2) (Davis & Yu 1997), and ‘pale-­green leaf’ gene (pg) (Sargent et al. 2004) were mapped respectively to LG III, VII, V, II, V, and VI. In total, 26 genes were positioned in the FV × FB reference map (Sargent et al. 2007, 2008). Additionally, sex-­determining loci were mapped in LG VI (Spigler et al. 2008), meanwhile male sterility locus was identified in LG IV of F. vesca ssp. bracteata (Tennesen et al. 2013).

In the cultivated octoploid strawberry genetic map, the cellulose bioregulator (COBRA) gene in LG1 (Sargent et al. 2012) and 11 other gene-specific loci (Sargent et al. 2009) have been positioned up to date.

Simultaneously, an extensive exploration of 232 × 1392 octoploid Fragaria map allowed to identify 33 QTLs responsible for fruit quality (Zorrilla-­Fontasi et al. 2011a). The QTLs were associated with fruit number (LGs II, III, V), width (LG V), colour (LG V), levels of L-­ascorbic acid (LGs IV, V), total acidity, pH, anthocyanins (LG III), and soluble solids (LG V). The authors also positioned some major genes controlling the fruit quality traits, such as putative R2R3 MYB transcription factor.
(LG II), Fa-Exp2 (encodes a fruit specific auxin-independent expansion; LG VII), SGR (the senescence-inducible chloroplast stay-green protein; LGV), a putative galactosyl-transferase (LG II), putative zinc-binding transcription factor (LG I), FaGalUR (encoding enzyme of D-galacturonate; LG IV), FaMYOX (myo-inositol oxidase; LG V), FvDHAR (dehydro-ascorbate reductase; LG VII), and FvGMP (GDP-mannose pyrophosphorylase; LG VII). Further investigations of the same mapping population resulted in detection of QTLs linked to fruit aroma and controlling the synthesis of 48 different volatiles (Zorrilla-Fontanesi et al. 2012).

In parallel, 87 QTLs controlling fruit quality traits were detected in the study on cultivated strawberry ‘Capitola’ × CF1116 map (Lerceteau-Köhler et al. 2012). The QTLs were associated with fruit diameter (LGs II, III, IV, V), length (LGs I, II, III, IV, V, VI, VII), ripening date (LGs II, III, IV, V, VI, VII), firmness (LGs II, III, IV, V), colour (LGs I, II, III, IV, V, VI), and also the status of organic acids (LGs I, II, III, IV, V), glucose, fructose, sucrose sugar (LGs I, III, IV, V, VI, VII), citrate and malate (LGs I, II, III, V, VI, VII).

PHYSICAL MAP

Till date, only one physical map within Fragaria genus – map of F. vesca ‘Hawaii 4’ (FvH4) genome was developed in the cooperation of huge group, containing over 70 scientific teams (Shulaev et al. 2011). A total of 219 Mb of sequence data were assembled into approximately 3,200 sequence scaffolds. Based on these scaffolds, 34,809 predicted genes were located and identified, of which 25,050 were subsequently preliminarily annotated. Sequencing scaffolds (272) contained the majority (96%) of the sequence data 222, containing a total of 197.7 Mb of nucleotides have been anchored to precise positions on the diploid Fragaria genetic map using the 411 STS markers (Sargent et al. 2011).

CHALLENGES OF FRAGARIA GENOME MAPPING

The genetic and genomic study on diploids v. polyploids, such as those of Fragaria species and many others within Rosaceae genera, are extremely difficult. The polyploidy increases allelic and non-allelic combinations due to multiple alleles that result in complex gene actions and interactions for QTL. In addition, the meiotic behaviour is still debated, considering either bivalent or multivalent formation, or both (Doerge & Craig 2000). The high heterozygosity among strawberry cultivars arises the difficulty in producing inbred lines. For that reason, pseudo test cross strategy (F1) is the most widely applied to construct mapping population of F. × ananassa. Allele dose estimation of polyploids for genetic map construction is still in process. Different techniques have been proposed to solve this problem, such as SDRF (single-dose restriction fragment) (Wu et al. 1992), MAC-PR (micro-satel-lite DNA allele counting-peak ratios) (Esselink et al. 2004), and MACDE (marker allele dose and configuration establishment) (van Dijk et al. 2012). Despite all obstacles, a comprehension of Fragaria genome organisation at the molecular level and the development of well-characterized wild and cultivated strawberry genome maps are essential prerequisite to detect markers linked to agronomic traits and thus improve breeding (Jiang 2013).

Acknowledgment

The paper was prepared within the framework of the study on strawberry map funded partially by Polish Ministry of Agriculture and Rural Development (project MRiRW HOR ln – 801 – 15/132, No.108/ 2011-2013) and partially by the government of Republic of Egypt (student exchange program).

REFERENCES


Esselink G.D., Nybom H., Vosman B. 2004. Assignment of allelic configuration in polyploids using...


colinearity between their genomes and the essentially disomic behavior of the cultivated octoploid strawberry. Genetics 179: 2045-2060. DOI: 10.1534/genetics.107.083840.


Verde I., Abbott A.G., Scalabrin S., Jung S., Shu S., Marroni F., et al. 2013. The high quality draft genome of peach (Prunus persica) identifies unique patterns of
The history of genome mapping in *Fragaria* genetic diversity, domestication and genome evolution. Nat. Genet. 45:487-496. DOI: 10.1038/ng.2586.


