Available online: www.notulaebotanicae.ro
Print ISSN 0255-965X; Electronic 1842-4309
Not Bot Horti Agrobo, 2018, 46(1):1-13. DOI:10.15835/nbha46110936
Original Article

Genomic Tools for the Enhancement of Vegetable Crops: A Case in Eggplant

Pietro GRAMAZIO[1]*, Jaime PROHENS[1], Mariola PLAZAS[2], Giulio MANGINO[1], Francisco J. HERRAIZ[1], Edgar GARCÍA-FORTEA[1], Santiago VILANOVA[1]

[1]*Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, Universitat Politècnica de València, Camino de Vera 14, 46022 Valencia, Spain; piegra@upv.es (*corresponding author); jprohens@btc.upv.es; giuman2@upvnet.upv.es; fraherga@upvnet.upv.es; edgarfor@etsiamn.upv.es; sanvina@upvnet.upv.es*

[2]*Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas, Universitat Politècnica de València, Camino de Vera 14, 46022 Valencia, Spain; maplaav@btc.upv.es*

Abstract

Dramatic advances in genomics during the last decades have led to a revolution in the field of vegetable crops breeding. Some vegetables, like tomato, have served as model crops in the application of genomic tools to plant breeding but other important crops, like eggplant (*Solanum melongena*), lagged behind. The advent of next generation sequencing (NGS) technologies and the continuous decrease of the sequencing costs have allowed to develop genomic tools with a greatly benefit for no-model plants such as eggplant. In this review we present the currently available genomic resources in eggplant and discuss their interest for breeding. The first draft of eggplant genome sequence and the new upcoming improved assembly, as well as the transcriptomes and RNA-based studies represent important genomic tools. The transcriptomes of cultivated eggplant and several wild relatives of eggplant are also available and have provided relevant information for the development of markers and understanding biological processes in eggplant. In addition, a historical overview of the eggplant genetic mapping studies, performed with different types of markers and experimental populations, provides a picture of the increase over time of the precision and resolution in the identification of candidate genes and QTLs for a wide range of stresses, and morpho-agronomic and domestication traits. Finally, we discuss how the development of new genetic and genomic tools in eggplant can pave the way for increasing the efficiency of eggplant breeding for developing improved varieties able to cope with the old and new challenges in horticultural production.

Keywords: experimental populations, genetics, genomics, genotyping strategies, QTLs, sequencing, *Solanum melongena*

Abbreviations: MAS: Marker-assisted selection; QTL: Quantitative trait locus; NCBI: National center for biotechnology information; RIL: Recombinant inbred line; UTR: Untranslated region; SNV: Single nucleotide variant; SSR: Simple sequence repeat; SNP: Single nucleotide polymorphism; INDEL: Insertion/deletion; AFLP: Amplified fragment length polymorphism; RAPD: Random amplification of polymorphic DNA; LG: Linkage group; COS: Conserved orthologous set; RFLP: Restriction fragment length polymorphism; PIC: Polymorphic information content; NGS: Next generation sequencing; CSS: Conserved syntenic segments; BC: Backcross; CAPS: Cleaved amplified polymorphic sequence; RAD: Restriction-site-associated; GBS: Genotyping-by-sequencing; SOL: Orthologous genes in *Solanum*; LD: Linkage disequilibrium; GWA: Genome-wide association; CGA: Chlorogenic acid; *PAL*: *Phenylalanine ammonia lyase*; *C4H*: *Cinnamate 4-hydroxilase*; *4CL*: *4-hydroxycinnamoyl-CoA ligase*; *HCT*: *Hydroxycinnamoyl-coA shikimate/quinate hydroxycinnamoil transferase*; *C3'H*: *P-coumaroyl ester 3'-hydroxylase*; *HQT*: *Hydroxycinnamoyl CoA quinate hydroxycinnamoyl transferase*; *PPO*: Polyphenol oxidase enzyme; NIL: Near isogenic line; IL: Introgression line; MAGIC: Multi-parent advanced generation inter-cross; SBG: Sequence based genotyping; SPET: Single primer enrichment technology

Received: 14 May 2017. Received in revised form: 30 Jul 2017. Accepted: 04 Aug 2017. Published online: 19 Sep 2017.

Introduction

Eggplant (*Solanum melongena* L., $2n = 2x = 24$), also known as common eggplant or brinjal eggplant, belongs to the Leptostemonum clade within the *Solanum* genus (the "spiny" solanums) and is the second most important solanaceous fruit crop in total production after tomato (*S. lycopersicum* L.) (Knapp *et al*., 2013). Eggplant has undergone a constant increase in yield (2.7-fold) and total production (8.7-fold) in the last fifty years, although the largest increases have been recorded in the last decade (FAOSTAT, 2017). Common eggplant is widespread and consumed worldwide even though in some areas, mainly in the African continent, two other cultivated eggplants namely scarlet (*S. aethiopicum* L.) and gboma (*S. macrocarpon* L.) eggplants are locally important (Lester and Daunay, 2003; Plazas *et al*., 2014).

The breeding and genomics revolution of the last fifteen years has resulted in the development of large amounts of information of interest for breeding, allowing the enhancement of many vegetable crops. Even though eggplant is the sixth most important vegetable in production in the world for many years it has lagged behind in the development and use of genomic tools compared to other important Solanaceae crops like tomato, potato or pepper (Barchi *et al*., 2011; Hurtado *et al*., 2013). Tomato, being the closest model crop species with many enhanced genomic tools, has been used as reference for eggplant in many studies (Doganlar *et al*., 2002; Wu *et al*., 2009; Barchi *et al*., 2012). In this respect, the scientific genomic achievements in tomato have allowed fishing in its gene pool and taking advantage of the extraordinary genetic wealth of its wild relatives (Víquez-Zamora *et al*., 2013; Aflitos *et al*., 2014). In this way, tomato breeders have been able to find genetic solutions to some biotic and abiotic threats and achieving improvements to cope with unfavorable agricultural environments in a climate change scenario (Bolger *et al*., 2014; Thapa *et al*., 2015). On the contrary, up to now in common eggplant, as well as in scarlet and gboma eggplant, the use of wild relatives for the breeding purpose have been negligible and often purely academic (Rotino *et al*., 2014). The availability of a large array of molecular markers obtained by genomic techniques allows a very efficient marker-assisted selection (MAS), by selecting the genes or quantitative trait loci (QTLs) of interest and facilitating the removal of undesirable traits of wild relatives and saving time and resources compared to the conventional breeding approach (Morrell *et al*., 2012; Brozynska *et al*., 2016). Even though the gap between eggplant and the crops with more genomic resources is still wide, a clear turnaround in this trend has occurred in the last few years (Hirakawa *et al*., 2014; Portis *et al*., 2015; Kouassi *et al*., 2016; Plazas *et al*., 2016; Salgon *et al*., 2017).

In this paper, we review the available genomic resources in eggplant gene pool, including experimental populations that are under development, as well as the future perspectives and directions to exploit the full potential of these tools for basic and applied research in this crop.

Genome assemblies

Sequencing of the genome of a crop is essential to dramatically accelerate crop improvement (Davey *et al*, 2012). A high quality reference genome opens the way to access to information of the complete set of genes, to the different layers of regulatory elements and to the basic genomic architecture, allowing precise structural and functional comparisons between species (Feuillet *et al*, 2011).

As of June 2017 there is only one genome draft publicly available (SME_r2.5.1) in the National Center for Biotechnology Information (NCBI) (SRA accession: DRR014074 and DRR014075) (Hirakawa *et al*., 2014). The accession used for the whole-genome shotgun sequencing was 'Nakate-Shinkuro', an important traditional Asian-type cultivar that has been used in the development of some modern commercial cultivars. For sequencing (~144X), a combined genomic libraries approach was employed, consisting of paired-end (insert size of 200–300 bp) and 2 Kb Illumina mate-pair insert size (Table 1). The transcriptome of two other accessions 'AE-P03' and 'LS1934' was also sequenced to improve the de novo assembly (Li *et al*., 2010). The final assembly consisted of 33,873 scaffolds, covering about 74% (833.1 Mb) of the estimated length of the genome (~1.1Gb) (Arumuganathan *et al*., 1991) with an N50 parameter of 64.5 Kb. The total number of genes predicted was 42,035, of which 4,018 were described as being exclusive of eggplant, quite a large number if compared with more completed genomes like the last version of tomato Heinz 1706 (SL3.0 version, 34,879 genes in ITAG3.10, https://solgenomics.net/organism/Solanum_lycopersicum/genome) or the last version of *Arabidopsis thaliana* genome (Araport11 version, 27,655 Protein Coding Genes) (Cheng *et al*., 2016).

A large number of different simple sequence repeats (SSRs) motifs and repeats (83,401) was identified by Hirakawa *et al*. (2014), as well as 4,536 single nucleotide polymorphisms (SNPs) from a microarray among the accessions used for the study. This first genomic resource, which is quite far to be a comprehensive work, paved the way to understanding the genomic architecture of the eggplant and allowed to perform comparisons with other important Solanaceae crops, like potato, tomato and pepper (Potato Genome Sequencing Consortium, 2011; The Tomato Genome Consortium, 2012; Qin *et al*., 2014), whose genomes are more complete and annotated, and also with model species like *Arabidopsis thaliana* (Cheng *et al*., 2016).

The development of another eggplant genome by the Italian Eggplant Genome Sequencing Consortium has been presented in several scientific meetings (Barchi *et al*., 2016), but at the time of writing this work has not yet been published. According to the authors, a high-quality reference genome has been de novo assembled through Illumina sequencing (~155X), using different genomic libraries sizes (from 270 bp to 10 Kb), of the inbred eggplant line "67/3", which was used as a male parent of a 157 F6 recombinant inbred lines (RILs) mapping

population (Table 1). In addition, to improving the assembly, its transcriptome was also sequenced (Barchi *et al*., 2016). Moreover, the hybrid assembly was combined with the sequencing of the female parent "305E40" (35X) and the rest (1X) of the RILs mapping population and with a high-resolution restriction optical map. The final result consisted of 12 pseudomolecules, spanning ~1.2Gb with an L50 of >3Mb. The functional annotation resulted in ~ 40K protein-coding genes, confirming the estimation of Hirakawa *et al*. (2014).

Transcriptomes and RNA-based studies

RNA sequencing is another essential genomic resource (Ozsolak and Milos, 2001). Probably due to the reduced economic and bioinformatic efforts compared to genome assembly and the variety of study approaches and aims, RNA-based studies are in general more abundant. In fact, in RNA-based studies the results depend on the genetic expression at a certain stage of development in a specific plant tissue (Waterhouse and Helliwell, 2003).

Regarding common eggplant, up to now, few studies using RNA sequencing have been carried out. One of them was the de novo assembly of the whole transcriptome from root, stem and young leaves samples (SRA accession: SRR1104129) (Yang *et al*., 2014). From paired-end 2x100 bp libraries, these authors retrieved 15 M of raw reads, which were assembled after a filtering process in 44,672 transcripts and 34,174 unigenes, a similar number to the

Table 1. Metrics for genome and RNA-based studies assemblies performed in the eggplant genepool

Species	Aim of the study	Genetic material	Plant material	Library	Sequencing platform	Raw reads (M)	Coverage	Final assembly	Marker discovery	NCBI accession	Source
S. melongena	Genome assembly	DNA	Leaves	PE (200–300 bp) MP (2Kb)	Illumina HiSeq 2000	1,323	144.3X	33,873 scaffolds	4,536 SNPs 83,401 SSRs	DRR014074 DRR014075	Hirakawa *et al*., 2014
S. melongena	Genome assembly	RNA	Leaves, roots, fruits, flowers	454 GS FLX library	Roche 454 GS FLX	1.4	0.48X				
S. melongena	Genome assembly	DNA	Leaves	Several library sizes from PE (270 bp) to MP (10Kb)	Illumina	-	155X 35X 1X	12 pseudo molecules	-	-	Barchi *et al*., 2016
S. melongena	Genome assembly	RNA	Leaves	-	Illumina	-	-				
S. melongena	Whole transcriptome assembly	RNA	Leaves, root, stem	PE (72 bp)	Illumina HiSeq 2000	30	-	38,185 unigenes	-	SRR1104129	Yang *et al*., 2014
S. torvum	Whole transcriptome assembly	RNA	Leaves, root, stem	PE (72 bp)	Illumina HiSeq 2000	54	-	34,174 unigenes	-	SRR1104128	
S. melongena	Identify putative allergens	RNA	Fruit	PE (100 bp)	Illumina HiSeq 2000	89	-	149,224 transcripts	-	SRR1291243	Ramesh *et al*., 2016
S. aethiopicum	Whole transcriptome assembly	RNA	Leaf, floral bud, fruit	PE (300 bp)	Illumina HiSeq 2000	114	-	87,084 unigenes	164,127 SNVs 976 SSRs	SRR2229192	Gramazio *et al*., 2016a
S. incanum	Whole transcriptome assembly	RNA	Leaf, floral bud, fruit	PE (300 bp)	Illumina HiSeq 2000	105	-	83,905 unigenes	12,396 SNVs 1,248 SSRs	SRR2289250	
S. aculeatissimum	Whole transcriptome assembly	RNA	Root	-	Illumina HiSeq 2000	28	-	69,824 unigenes	-	SRS1383901 SRS1383902	Zhou *et al*., 2016
S. melongena	Identification of MiRNA	RNA	Seedlings	Small RNA libraries	Illumina MiSeq	30	-	5,940 miRNA	-	SRR833801 SRR833802	Yang *et al*., 2013
S. melongena	Identification of MiRNA	RNA	Pistil	Small RNA libraries	-	-	-	686 miRNAs	-	SRR3479276 SRR3479277	Wang, 2017

protein-coding genes from the genome annotation (Hirakawa *et al*., 2014). In addition to performing structural and functional annotation, a comparison was performed using a set of 4,900 orthologs with other 11 plant species, including model plants and Solanaceae crops like tomato and potato, allowing to estimate the time of divergence between them. Furthermore, using the Plant Resistance Gene database (http://prgdb.crg.eu/wiki/Main_Page) a set of 621 resistance genes were identified.

In the same study (Yang *et al*., 2014), the transcriptome of the wild eggplant relative *S. torvum* Sw., a species belonging to the tertiary eggplant gene pool (Syfert *et al*., 2016), was also assembled (SRA accession: SRR1104128). *Solanum torvum*, also known as turkey berry, is of great interest for eggplant breeding since it is resistant to a wide range of soil-borne diseases like root-knot nematodes, *Ralstonia solanacearum*, *Verticillium dahliae*, and *Fusarium oxysporum* f. sp. *melongenae* (Gousset *et al*., 2005; Yamaguchi *et al*., 2010). For *S. torvum*, the assembly statistics and annotation were slightly different to that of common eggplant, being the most relevant difference the number of unigenes (38,185 for *S. torvum* versus 34,174 for common eggplant) (Table 1).

Another de novo transcriptome assembly was released in 2016 in order to identify putative allergens in eggplant fruit (Ramesh *et al*., 2016; SRA accession: SRR1291243). In this study, total RNA was extracted from fruit peel and flesh, a different plant material compared to Yang *et al*. (2014) whole transcriptome, and a total of 48 putative allergens and 526 B-cell linear epitopes were identified from the 149,224 transcripts assembled (Table 1). Of these 40,752 showed significant similarity with predicted proteins in tomato and potato.

In addition to common eggplant, the whole transcriptome of some eggplant relatives has been released. One of them was scarlet eggplant (*S. aethiopicum* accession BBS135) (Gramazio *et al*., 2016a; SRA accession: SRR2229192), the second most important cultivated eggplant, which is common in sub-Saharan Africa, as well as, in some areas of Brazil, Caribbean and south of Italy (Sunseri *et al*., 2010). In the same study, a *de novo* whole transcriptome of *Solanum incanum* L. accession MM577 (SRA accession: SRR2289250), a wild relative of common eggplant which is considered a powerful source of phenolics and tolerant to some abiotic stresses such as drought (Knapp *et al*. 2013), was also assembled.

Due to the high amount of reads obtained (more than 100 M), the number of assembled unigenes in Gramazio *et al*. (2016a) was relatively high (87,084 for *S. aethiopicum* and 83,905 for *S. incanum*), probably due to high representation of 3′ or 5′ untranslated regions (UTRs) and intron sequences from non-mature mRNAs as a consequence of a deep coverage (Table 1). On the other hand, the number of annotated unigenes in protein databases was similar to other *Solanum* crops (34,231 and 30,630 for *S. aethiopicum* and *S. incanum* respectively). In this study, molecular marker discovery was performed, identifying a total of 1,248 microsatellites for scarlet eggplant and 976 for *S. incanum*. In addition, intraspecific and interspecific single nucleotide variant (SNV), SNPs and insertions/deletions (INDELs), were identified not only in *S. aethiopicum* and *S. incanum* but also in *S. melongena* and *S. torvum*, using for comparison the raw reads generated in the Yang *et al*. (2014) study.

The most recent transcriptome released in the eggplant genepool was from *S. aculeatissimum* Jacq., a wild relative of eggplant resistant to verticillium wilt (Zhou *et al*., 2016). Two different libraries were constructed, for the control and for the infected roots with *Verticillium dahliae*, each of them giving 28 M of raw reads. A total of 64,413 and 71,291 unigenes were obtained for the control and infected roots, respectively, which were functionally annotated, resulting in 17,645 of them differentially expressed (11,696 upregulated and 5,949 downregulated).

Apart from the whole transcriptomes, in common eggplant, other RNA-based studies were performed. Yang *et at*. (2013) identified miRNA from eggplant involved in the process of infection by *V. dahliae* through deep-sequencing of two small RNA libraries, using control and infected seedlings. From 5,940 MiRNA identified, 220 belonged to Solanaceae species and two new miRNA were eggplant specific. The authors identified a total of 33 differentially expressed miRNA between the two libraries (28 down-regulated and 5 up-regulated), which were strongly involved in the *V. dahliae* infectious process.

The most recent RNA-based study (SRA: SRR3479276, SRA: SRR3479277), which is still unpublished and only available in NBCI database (https://www.ncbi.nlm.nih.gov/sra/SRR3479277/), identified miRNA involved in the difference in heterostyly in short-morph and long-morph eggplant pistils using high-throughput small RNA microarray and degradome sequencing. From the 686 miRNAs identified, 10 were differentially expressed and were determined as pistil development-related miRNAs.

Mapping studies, experimental populations, and genotyping methods

Gene mapping establishes a connection between a trait under study and one or more chromosomal regions (Sevon *et al*., 2005). Although genetic maps have been constructed since the first decade of the 20th century (Brown, 2006), the first genetic map in eggplant was released only in 2001 (Nunome *et al*., 2001). This intraspecific map was built using 181 dominant markers (93 amplified fragment length polymorphism (AFLPs) and 88 random amplification of polymorphic DNA (RAPDs)) using a population of 168 F2 progenies, resulting in 21 linkage groups (LGs) and spanning 779.2 cM (Table 2). The aim of this study was to identify genetic regions involved in fruit shape and color development, which were associated with regions in LG2 and LG7 respectively. Low frequency of DNA polymorphism in eggplant (Doganlar *et al*., 2002a; Wu *et al*., 2009) and the tendency of some markers to cluster, like AFLPs and RAPDs (Alonso-Blanco *et al*., 1998; Nilsson *et al*., 2007), generated a high number of LG, which did not correspond to the basic chromosome number in eggplant.

One year later (2002), an interspecific genetic map was developed, using *S. melongena* and the wild relative *S.*

linneanum Hepper & P.-M.L.Jaeger as parents, in order to increase the general low DNA polymorphism observed in mapping populations derived from an intraspecific cross (Doganlar *et al*., 2002a). The aim of this study was to compare tomato and eggplant maps in order to evaluate synteny and identify rearran-gements between the two *Solanum* species occurred during the domestication process. To achieve that objective, single-copy tomato cDNA, genomic tomato DNA, and tomato conserved orthologous set (COS) restriction fragment length polymorphism (RFLP) markers were assessed, which were previously mapped in a tomato map and used to establish a synteny between the tomato and potato genome (Tanksley *et al*. 1992; Fulton *et al*. 2002).

The map, which spanned 1,480 cM along 12 LG (Table 2), confirmed the high collinearity between the tomato and eggplant and identified 25 rearrangements, as well as 125 QTLs related to 22 domestication traits (fruit size, shape, and color and plant prickliness) (Doganlar *et al*., 2002b) and 18 morphological traits related to leaf, flower, and fruit size, shape, appearance, and development (Frary *et al*., 2003).

Another interspecific genetic map using the eggplant wild relative *S. sodomeum* L. [=*S. linnaeanum*], was developed by Sunseri *et al*. (2003). This map, that was built to achieve markers linked to *Verticillium* tolerance using 48 F2 progenies derived from an interspecific cross between a susceptible and tolerant parental, consisted of 273 markers (156 AFLPs and 117 RAPDs) distributed along 12 LG and 736 cM with a small average distance between them (2.7 cM, Table 2).

An improved version of the Nunome *et al*. (2001) genetic map was released in 2003 (Nunome *et al*., 2003a), where seven SSRs markers were added through screening an eggplant genomic library of dinucleotide motifs, in order to merge and reduce LGs. In fact, the number of LGs decreased from 21 to 17, while the map length and the average marker density remained almost the same (Table 2). At the same time, an eggplant genomic library was developed for screening trinucleotides motifs (Nunome *et al*., 2003b). SSRs, also known as microsatellites, are highly polymorphic, genomic (gSSRs) more than genic (EST-SSRs), within and across the species, generally displaying a higher polymorphic information content (PIC) compared to the other molecular markers (Kalia *et al*., 2011). In addition, microsatellites are codominant, robust, highly reproducible, abundant and quite-well spread along the genome (Varshney *et al*., 2005), Nevertheless, before the advent of next generation sequencing (NGS) their identification through the development of genomic libraries required a quite high degree of expertise, which implied a significant investment of time and resources (Fernandez-Silva *et al*., 2013). Nowadays, the identification of thousands of gSSRs and EST-SSRs by DNA and RNA sequencing has become cost-effective and effortless by virtue of the overwhelming advancements in sequencing platforms (Xiao *et al*., 2013; Goodwin *et al*., 2016). Other genomic libraries to discover microsatellites in eggplant were developed by Stàgel *et al*. (2008), Vilanova *et al*. (2012) and Nunome *et al*. (2009), the latter in order to improve the previous versions of their genetic map (Numone *et al*., 2001, 2003) by adding 245 SSRs, thus reducing to 14 the number of LGs and increase the total length to 959 cM (Table 2).

An enhanced version of Doganlar *et al*. (2002) genetic map was obtained by adding 110 COSII and 5 tomato-derived markers and performing a detailed synteny with tomato through inferred the position of additional 522 COSII markers (Wu *et al*., 2009) (Table 2). The high number of common markers allowed to estimate the time of divergence in 12 million years and from genomic structural standpoint a minimum of 24 inversions and 5 chromosomal translocations occurred between the two species, and 37 conserved syntenic segments (CSSs) where the order of genes/markers have been well preserved were detected.

In order to map the resistance to *F. oxysporum* (gene *Rfo-sa1*), Barchi *et al*. (2010) developed an intraspecific genetic map using a 141 F2 population from 305E40, a double haploid line obtained through another culture of a backcross material (BC7) using *S. aethiopicum* as a donor parent carrying the *Rfo-sa1* gene for resistance to *Fusarium*, and the line 67/3, derived from an intraspecific cross between an Italian and Chinese cultivars and susceptible to *F. oxysporum*. The map spanned 718.7 cM across 12 LGs with an average marker density of 3.0 cM and was composed of 212 AFLPs, 22 SSRs developed by Stàgel *et al*. (2008), 1 RFLP and three *Rfo-sa1* cleaved amplified polymorphic sequence (CAPS) markers which co-segregated in LG1 (Table 2).

The same F2 population was also used to identify QTLs related to anthocyanin content (Barchi *et al*., 2012), as well as, 105 QTLs associated with twenty yield, fruit, and morphological traits (Portis *et al*., 2014) and 29 QTLs with seventeen traits (fruit qualitative and health-related compounds) (Toppino *et al*., 2016), this time employing a large set of SNPs identified by restriction-site-associated (RAD) tags sequencing (Barchi *et al*., 2011), which led the total map length to 1390 cM, quite similar to that of Wu *et al*. (2009) and Doganlar *et al*. (2002) maps (Table 2).

In fact, when markers that tend to cluster, like AFLPs, are replaced with markers that are more dispersed, like SNPs, the total length of genetic maps usually increases (Rafalski, 2002). SNPs, like SSRs, are codominant, robust and easy to identify by NGS platform through sequencing (Davey *et al*., 2011; Scheben *et al*., 2016), and in addition have the advantage that are more abundant, ubiquitous and easy to automate than SSRs (Thomson *et al*., 2014; Kim *et al*., 2016), although are less informative (Filippi *et al*., 2015; Gonzaga, 2015). SSRs are generally considered as dispersed markers, even though in some species tend to concentrate more frequently in heterochromatic regions and it is unlikely to cover all the genomic regions assessing just SSRs (Hong *et al*., 2007; Shirasawa *et al*., 2010). On the other hand, it has been reported that the validation of SNPs by high-throughput SNP genotyping, like genotyping-by-sequencing (GBS), RAD tag sequencing or similar, can be 100-fold faster and 75% less expensive than SSRs detection through an agarose or polyacrylamide gels or capillary sequencing (Jones *et al*., 2007; Yan *et al*., 2010). For all these reasons SNPs markers have quickly replaced SSRs in the last few years.

Fukuoka *et al*. (2012) in an effort to represent genomic

region overlooked in the previous maps, mapped a considerable set of SNPs identified from a set of 4,754 orthologous genes in *Solanum* (SOL) developed from 16K eggplant, 47K tomato, and 57K potato unigenes. An integrated intraspecific map was built from two F2 populations, LWF2 and ALF2, and 952 markers (639 SNPs and 313 SSRs) along 12 LGs, resulting in 1,285 cM and an average marker density of 1.4 cM, covering 1.5 times the genomic region represented in Nunome *et al*. (2009) (Table 2).

The same set of SNPs developed by Fukuoka *et al*. (2012) and the SSRs developed by Nunome *et al*. (2009) were used to build two intraspecific maps from two F2 populations, ALF2 and NAF2, to identify QTLs involved in parthenocarpy (Miyatake *et al*., 2012). The two main QTLs detected (*Cop3.1* and *Cop8.1*) were identified in both maps, which presented different LGs (12 versus 15), length (1,414 versus 1,153), and markers mapped (132 SNPs and 118 SSRs versus 125 SSRs and 49 SNPs), although shared one parent (parthenocarpic line AE-P03) (Table 2).

The first intraspecific population of RILs in eggplant was developed from the resistant bacterial wilt *R. solanacearum* line AG91-25 (MM960), derived from the Turkish line MM127 and a *S. aethiopicum* Aculeatum Group accession, and the commercial type line MM738 (Lebeau *et al*., 2013), which were previously used by Doganlar *et al*. (2002a). That population was used to dissect the genetic control of resistance to *R. solanacearum* to four strains of phylotype I and identify genes/QTLs through developing a genetic map, which spanned 884 cM along 18 LGs (Table 2). Although, the map had a low saturation due to extremely low polymorphism rate, with only 119 markers mapped and most of which AFLPs, a major gene was identified (*ERs1*) in LG2. The low level of polymorphism in the RILs population showed with AFLPs drastically changed when the individuals were sequenced by GBS, identifying 1,779 filtered SNPs and allowing to build a high-density genetic map for screening four new strains of *R. solanacearum*, apart from the previous ones of Lebeau *et al*. (2013), belonging to phylotypes I, IIA, IIB and III (Salgon *et al*., 2017). The overall map length and the average marker density had been increased to 1,085 and 4.4 cM, respectively, and the LGs number had been reduced from 18 to 14 (Table 2), which lead to identifying a major QTL at the bottom of LG 9 that controls three phylotype I strains, corresponded to the previously identified major gene *ERs1* of Lebeau *et al*. (2013), and two other minor QTLs on LG 2, associated with partial resistance to strains of phylotypes I, IIA, III, and on LG 5, controlling the strains of phylotypes IIA and III. This was a clear example of the high potential of the recent advances in genomics which improves extremely the resolution and precision of the genetic studies.

An additional study about of *Fusarium* wilt resistance (*F. oxysporum* f. sp. *melongenae*) was performed for mapping a resistant locus using different F2 and F3 populations, LWF2, EWF2, ALF2, LWF3, and EWF3, a backcross inbred lines population, ALBIL, and three genetic maps, AL2010, LW2012, and EW2012, previously developed by Fukuoka *et al*. (2012) and Hirakawa *et al*. (2014) (Table 2). A major resistance locus (*FM1*) was identified at the end of chromosome 2, at the exact same position of *Rfo-sa1* from *S. aethiopicum* gr. Gilo (Toppino *et al*., 2008), suggesting they might be orthologous.

The first attempt of association mapping based on linkage disequilibrium (LD) was performed with 141 eggplant accession from different countries and 105 SSRs developed by Nunome *et al*. (2009) to investigate nine fruit traits (Ge *et al*., 2013, Table 2). The analysis performed revealed a total of 49 marker associations related to eight phenotypic traits and 24 SSRs, being the total variation explained ranged from 4.5 to 22.8%.

A larger set of 191 accessions, including breeding lines, old varieties, and landraces from the Mediterranean basin and Asia, were investigated by Cericola *et al*. (2014) based on genome-wide association (GWA) approach for anthocyanin pigmentation and fruit color at two locations over two years using 314 SNPs developed by Barchi *et al*. (2011). A total of 56 associations were found between SNPs and anthocyanin content and fruit color-related traits, which were clustered into 12 groups and scattered over nine chromosomes, being eight of the groups overlapping with known QTL and demonstrating in that way the effectiveness of GWA approach. In addition, synteny with tomato allowed the identification of the genomic regions associated with anthocyanin accumulation in LGs 2, 5, and 12. Using the same association panel of 191 accessions and set of 314 SNPs (Table 2), Portis *et al*. (2015) examined the phenotype/genotype associations related to 33 traits (fruit, plant and leaf morphology traits) identifying 194 association to 30 traits, which involved 79 SNP loci in 39 distinct regions distributed across the 12 LGs.

A further improvement of Doganlar *et al*. (2002a) and Wu *et al*. (2009) maps was performed by increasing the number of markers to 864, by adding 400 AFLPs and 117 RFLPs, and using a larger F2 population (108 individuals) (Doganlar *et al*., 2014). In that way, the overall map length remained almost the same but the marker average distance decreased from 6.1 cM to 1.8 cM (Table 2). On the other hand, the improved map precision led the authors to revise the number of rearrangements between eggplant and tomato from 29 (five translocations and 24 inversions) (Doganlar *et al*., 2002a; Wu *et al*., 2009) to 33 (19 translocations and 14 inversions). Simultaneously, Frary *et al*. (2014) revised the QTLs identification by Doganlar *et al*. (2002b) and Frary *et al*. (2003), taking the opportunity of the improved map of Doganlar *et al*. (2014) to interrogate 58 F2 individuals (*S. linnaeanum* line MM195 and *S. melongena* line MM738) for 42 traits identifying 71 QTLs, twenty-two of which were novel.

In order to exploit the great genetic diversity of the wild relatives, a new interspecific map (named SMIBC) was developed using *S. incanum* MM577, which, as has been mentioned previously, is considered a powerful source of phenolics and tolerant to some abiotic stresses such as drought (Knapp *et al*., 2013). In fact, with the objective of increasing the content of chlorogenic acid (CGA) in eggplant, which is usually the main phenolic compound (Whitaker *et al*., 2003; Prohens *et al*., 2013), a first set of introgression lines in the eggplant gene pool has been developed using *S. incanum* as a donor parent (Gramazio *et*

al., 2016b). Thus, to track and easily introgress the alleles involved in the content of CGA (*phenylalanine ammonia lyase* (*PAL*), *cinnamate 4-hydroxilase* (*C4H*), *4-hydroxycinnamoyl-CoA ligase* (*4CL*), *hydroxycinnamoyl-coA shikimate/quinate hydroxycinnamoil transferase* (*HCT*), *p-coumaroyl ester 3'-hydroxylase* (*C3'H*), and *hydroxycinnamoyl CoA quinate hydroxycinnamoyl transferase* (*HQT*)) in the genetic background of eggplant, they were successfully mapped in different LGs using 91 BC1 individuals (Gramazio *et al*., 2014). In addition, five polyphenol oxidase enzymes (*PPO1*, *PPO2*, *PPO3*, *PPO4*, *PPO5*), which may be involved in the browning of the fruit flesh by oxidation of GCA and other phenols, mapped in a cluster in LG 8, as well as, candidate genes important in domestication such as fruit shape (*OVATE*, *SISUN1*) and prickliness. The mapping was assisted by the use of synteny of the orthologous genes in tomato using Tomato-EXPEN 2000 map (Fulton *et al*., 2002). Furthermore, apart from tomato, SMIBC established a macro-synteny with four other eggplant maps (Nunome *et al*., 2009; Wu *et al*., 2009; Barchi *et al*., 2012; Fukuoka *et al*., 2012) by using shared markers. SMIBC spanned 1,085 cM along 12 LGs with a total of 243 markers (42 COSII, 99 SSRs, 88 AFLPs, 9 CAPS, 4 SNPs and one morphological marker) (Table 2).

A few months later, at the end of 2014, the first draft of eggplant genome was finally released online (Hirakawa *et al*., 2014), where, apart from the genomic sequence, an integrated linkage map was also constructed from two F2 population (EWF2 and LWF2), which were previously used for mapping by Nunome *et al*. (2001) and Fukuoka *et al*. (2012). The map presented a total of 795 markers (574 SNPs and 221 SSRs) for an overall map length of 1,280 cM along 12 LGs and achieving the highest average marker density so far (0.7 cM) (Table 2).

In order to exploit genomic resources and genetic data for key agronomic traits, Rinaldi *et al*. (2016) performed a syntenic relationship and QTL orthology among eggplant, tomato, and pepper using their respective genome assemblies, although the genome sequences of the three species are different in coverage, assembly quality, and percentage of anchorage. While the comparison between tomato and pepper was quite comprehensive due to the high quality of their assembly, the comparison with eggplant and tomato and eggplant and pepper was less exhaustive due to the impossibility to localize the physical position of the most eggplant QTL. Nevertheless, most of the previous rearrangements previous detected were confirmed and new ones were identified, even though an enhanced version of eggplant genome could have improved the precision of the analysis.

Future direction for genetics and genomics tools for eggplant breeding

In the last ten years from the genetic and genomic investigation standpoint, eggplant has shortened the gap with other important major crops like tomato, potato, and pepper. Nevertheless, more genomic resources are needed for an efficient breeding in order to develop new improved varieties that have to deal with a changing climate scenario and new and threatening biotic stresses, and to understand the genetic diversity, evolutionary history, domestication, and ecology in eggplant (Gramazio *et al*., 2016a).

In this light, a high-quality genome sequence from an inbred line is absolutely necessary, in which the scaffolds should be assembled in chromosomes or pseudomolecules from sequence-based genetic and physical maps. The sequencing technologies using long-reads such as PacBio combined with a huge amount of small-reads of Illumina could achieve a satisfactory result (Mavromatis *et al*., 2012), although long-read sequencing technologies combined with physical mapping approaches, like optical mapping, Hi-C and the Dovetail Chicago methods, are offering new and most accurate solution to genome assembly (Yuan *et al*., 2017). An accurate reference genome assembly would encourage many research groups to sequence the genomes and the transcriptomes of their genotypes of interest or wild eggplant related species with a modest amount of economic resources, obtaining a meaningful information for a wide range of studies.

For instance, more attempts to develop experimental populations using wild relatives could be addressed if more information would be available for marker assisted selection to introgress efficiently genomic fragments of allied species. Up to now, the attempts to introgress wild relatives traits has been performed only sporadically and even though 25 allied eggplant species have been employed in interspecific cross with *S. melongena* (Rotino *et al*., 2014), only one set of introgression lines has been developed using *S. incanum* as a donor parent (Gramazio *et al*., 2016b) and no other biparental and multiparental populations are available.

The delay in developing populations using wild relatives is staggering compared with other crops like tomato, in which more than 96 genes and QTLs has been introgressed from 14 different wild relatives, comprising several sets of near isogenic lines (NILs), introgression lines (ILs) and RILs, as well as a multi-parent advanced generation intercross (MAGIC) population (Pascual *et al*., 2015; Redden *et al*. 2015). A reliable reference genome would indeed speed up the development and precision of these populations in eggplant. Barrantes *et al*. (2014) developed most of the lines of an ILs between tomato and *S. pimpinellifolium* L. from a BC3S1 generation, four generations versus eight required to develop the set of ILs with *S. incanum* (Gramazio *et al*., 2016b). Both tomato and *S. pimpinellifolium* parents had advanced assemblies that allowed designing arrays to perform high-throughput genotyping since the early generations of ILs population development.

In the last few years, sequence based genotyping (SBG) technology, which includes methods for the simultaneous polymorphism discovery and genotyping like GBS, RAD, ddRAD and related methods, has been a relative economical way to produce information in a crop and facilitate genotyping before a reference genome was available (Elshire *et al*., 2011). Barchi *et al*. (2011) RAD approach identified ~10,000 SNPs and 2,000 putative SSRs from the parents of a mapping population with a reference genome. Similarly, a GBS was performed in a 180 F6 RILs population identifying 1,779 filtered SNPs and improved drastically the quality of an intraspecific genetic map (Salgon *et al*., 2017). These genotyping through sequencing methods constitute a really powerful tools to increase precision and accelerate

Table 2. Statistics of mapping studies performed in the eggplant genepool

Source	Type of map	Aim of the study	Population		Parentals	Markers	Map lenght (cM)	Linkage groups	Average distance between markers (cM)
Nunome *et al*., 2001	Intraspecific genetic map	Fruit shape, color development	168	F2	EPL-1 (*S. melongena,* Japan) WCGR112-8 (*S. melongena,* India)	93 AFLPs 88 RAPDs	779	21	4.9
Doganlar *et al*., 2002a	Interspecific genetic map	Syntheny with tomato	58	F2	MM195 (*S. linnaeanum*) MM738 (*S. melongena*)	232 RFLPs	1,480	12	7.6
Sunseri *et al*., 2003	Interspecific genetic map	*Verticillium* wilt tolerance	48	F2	*S. sodomeum* [=*S. linnaeanum*] Buia (*S. melongena*)	156 AFLPs 117 RAPDs	736	12	2.7
Nunome *et al*., 2003	Intraspecific genetic map	Fruit shape, color development	120	F2	EPL-1 (*S. melongena,* Japan) WCGR112-8 (*S. melongena,* India)	101 RAPDs 54 AFLPs 7 SSRs	716	17	4.9
Nunome *et al*., 2009	Intraspecific genetic map	SSR development	94	F2	EPL-1 (*S. melongena,* Japan) WCGR112-8 (*S. melongena,* India)	245 SSRs	959	14	4.3
Wu *et al*., 2009	Interspecific genetic map	Syntheny with tomato	58	F2	MM195 (*S. linnaeanum*) MM738 (*S. melongena*)	232 RFLPs 110 COSII 5 tomato-derived markers	1,535	12	6.1
Barchi *et al*., 2010	Intraspecific genetic map	*F. oxysporum* resistance	141	F2	305E40 (DH, from *S. melongena* and *S. eathiopicum*) 67/3 (*S. melongena*)	212 AFLPs 22 SSRs 1 RFLP 3 CAPS	718	12	3.0
Barchi *et al*., 2012	Intraspecific genetic map	Anthocyanin content	156	F2	305E40 (DH, from *S. melongena* and *S. eathiopicum*) 67/3 (*S. melongena*)	339 SNPs 33 SSRs 27 COSII 11 RFLPs 3 CAPS 2 HRM	1,390	12	3.8
Fukuoka *et al*., 2012	Intraspecific genetic map (LW2010 x AL2010 = LWA2010)	Development of *Solanum* orthologous (SOL) gene sets	90 F2 93 F2	(LWF2) (ALF2)	LS1934 (*S. melongena*) WCGR112-8 (*S. melongena*, India) LS1934 (S. melongena) AE-P03 (*S. melongena*)	639 SNPs 313 SSRs	1,285	12	1.4
Miyatake *et al*., 2012	Intraspecific genetic map	QTL identification for parthenocarpy	135 F2 93 F2 (NAF2)	(ALF2)	LS1934 (*S. melongena*) AE-P03 (*S. melongena*) Nakate-Shinkuro (*S. melongena*) AE-P03 (*S. melongena*)	132 SNPs 118 SSRs 125 SSRs 49 SNPs	1,414 1153	12 15	- -

Lebeau *et al*., 2013	Intraspecific genetic map	Genetic control of resistance to *R. solanacearum*	178	F6	MM738 (*S. melongena*) AG91-25 (MM960, *S. melongena*)	91 AFLPs 26 SSRs 2 SRAPs	884	18	8.8
Ge *et al*., 2013	Association mapping (GWA)	association analysis of fruit traits	141 eggplant accessions		128 from China and the rest from USA, India, Japan, Italy, Malaysia, Arab Emirates, Thailand, Korea	105 SSRs	-	-	-
Cericola *et al*., 2013	Association mapping (GWA)	Anthocyanin pigmentation and fruit color	191 eggplant accessions		Breeding lines, old varieties, and landraces from the Mediterranean basin and Asia	314 SNPs	-	-	-
Doganlar *et al*., 2014	Interspecific genetic map	Syntheny with tomato	108	F2	MM195 (*S. linnaeanum*) MM738 (*S. melongena*)	400 AFLPs 348 RFLP 116 COSII	1,518	12	1.8
Gramazio *et al*., 2014	Interspecific genetic map (SMIBC)	Mapping chlorogenic acid biosynthesis pathway and polyphenol oxidase genes	91	BC1	MM577 (*S. incanum*) AN-S-26 (*S. melongena*)	99 SSRs 88 AFLPs 42 COSII 9 CAPS 4 SNPs 1 morphological	1,085	12	4.4
Hirakawa *et al*., 2014)	Intraspecific genetic map (AL2010 x EWF2 x LWF2 = LWAE2012)	Draft genome sequence of eggplant	120 F2 90 F2	(EWF2) (LWF2)	EPL-1 (*S. melongen*a, Japan) WCGR112-8 (*S. melongena*, India) LS1934 (*S. melongena*) WCGR112-8 (*S. melongena*, India)	574 SNPs 221 SSRs	1,280	12	0.7
Portis *et al*., 2015	Association mapping (GWA)	Fruit, plant and leaf morphological traits	191 eggplant accessions		Breeding lines, old varieties, and landraces from the Mediterranean basin and Asia	314 SNPs	-	-	-
Miyatake *et al*., 2016	QTL mapping (AL2010, LW2012, and EW2012)	Mapping of a resistance locus against *F. oxysporum* f. sp. *melongenae*	90 F2 (LWF2) 120 F2 (EWF2) 93 F2 (ALF2) 87 F3 (LWF3) 120 F3 (EWF3) 186 BILs (ALBIL)		LS1934 × WCGR112-8 EPL-1 × WCGR112-8 LS1934 × AE-P03 Single-seed descent from LWF2 Single-seed descent from EWF2 (LS1934 × AE-P03) × LS1934	-	-	-	-
Salgon *et al*., 2017	Intraspecific genetic map	Genetic control of resistance to *R. solanacearum*	180	F6	MM738 (*S. melongena*) AG91-25 (MM960, *S. melongena*)	867 SNPs 139 AFLPs 28 SSRs 1 SRAP	1,518	14	1.4

genetic and genomic studies, although the legal dispute for exploited the patent of these methods driven up the prices of these technologies turning them economical prohibitive for many research groups.

Currently, many efforts are being dedicated to developing an alternative to SBG technologies, although many of them required prior knowledge. The information obtained by sequencing is necessary to design primers to interrogate the genomic regions of interest for allelic discovery or genotyping. Each targeting platform differs in throughput, cost, probes, multiplexing, the number of target regions and much more customizable parameters. Among the most-used alternatives are Sequenom MassARRAY iPLEX platform (Gabriel *et al*., 2009), single primer enrichment technology (SPET) (NuGEN, San Carlos, USA), TruSeq Amplicon Sequencing and Nextera Target Enrichment (Illumina, San Diego, USA), KASP SNP genotyping (LGC Genomics, UK).

A new promising approach, rAmpSeq, was announced at the end of 2016 as a robust genotyping platform using repetitive sequences (Buckler *et al*., 2016). This method used conserved regions to design PCR primers for amplifying thousands of middle repetitive regions and interrogate thousands of markers. The authors affirmed that the cost per sample can be less than $2 per sample, which would allow to genotype thousands of samples for a very reasonable cost. Other advantages are the use of PCR without high requirements of DNA quality and quantity and less PCR competition among amplicons due to fairly similar length and composition of the repetitive. On the other hand, compared to the SBG technologies, rAmpSeq identifies fewer markers, required prior information and generally screens intergenic regions, as well as, more challenging bioinformatic analysis. This approach can revolutionize the breeding and conservation biology in the immediate future, even though at moment further improvement are required and optimization in more crops a part of maize.

The genomics revolution, that has led a perspective change in our comprehension of evolution, domestication, genetic architecture and much more aspects, is far from slowing down. The new achievements in sequencing technologies and their decreasing in cost are accelerating the development of high-quality genome reference assemblies, high-throughput genotyping and markers-assisted breeding selection that is reflecting in a greater overall understanding of species and new improved varieties adapted to new upcoming scenarios. In the near future sequencing 100s or 1,000s of samples will become routinary and affordable including for non-model species and for resource and infrastructure-limited institutions in the developing world. This will undoubtly speed up and revolutionize the breeding of eggplant for the development of a new generation of cultivars with dramatically improved yield, quality and resilience.

Acknowledgements

This work has been funded in part by the initiative "Adapting Agriculture to Climate Change: Collecting, Protecting and Preparing Crop Wild Relatives", which is supported by the Government of Norway. This project is managed by the Global Crop Diversity Trust with the Millennium Seed Bank of the Royal Botanic Gardens, Kew and implemented in partnership with national and international gene banks and plant breeding institutes around the world. For further information see the project website: http://www.cwrdiversity.org/. Funding has also been received from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 677379 (G2P-SOL project: Linking genetic resources, genomes and phenotypes of Solanaceous crops) and from Spanish Ministerio de Economía, Industria y Competitividad and Fondo Europeo de Desarrollo Regional (grant AGL2015-64755-R from MINECO / FEDER). Pietro Gramazio is grateful to Universitat Politècnica de València for a pre-doctoral (Programa FPI de la UPV-Subprograma 1/2013 call) contract. Mariola Plazas is grateful to Ministerio de Economía, Industria y Competitividad for a post-doctoral grant within the Juan de la Cierva programme (FCJI-2015-24835). Giulio Mangino is grateful to Conselleria d'Educació, Investigació, Cultura i Esport de la Generalitat Valenciana for a pre-doctoral grant within the Santiago Grisolía programme (GRISOLIAP / 2016/012).

References

Aflitos S, Schijlen E, De Jong H, De Ridder D, Smit S, Finkers R, … Peters S. (2014). Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing. The Plant Journal 80:136-148.

Alonso-Blanco C, Peeters A, Koornneef M (1998). Development of an AFLP based linkage map of Ler Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. The Plant Journal 14:259-271.

Arumuganathan K, Earl ED (1991). Nuclear DNA content of some important plant species. Plant Molecular Biology Reporter 9:208-218.

Barchi L, Lanteri S, Portis E, Stàgel A, Valè G, Toppino L, … Rotino GL (2010). Segregation distortion and linkage analysis in eggplant (*Solanum melongena* L.). Genome 53:805-815.

Barchi L, Lanteri S, Portis E, Acquadro A, Vale G, Toppino L, Rotino GL (2011). Identification of SNP and SSR markers in eggplant using RAD tag sequencing. BMC Genomics 12:304.

Barchi L, Lanteri S, Portis E, Valè G, Volante A, Pulcini L, … Rotino GL (2012). A RAD tag derived marker based eggplant linkage map and the location of QTLs determining anthocyanin pigmentation. PLoS One 7:e43740.

Barchi L, Delledonne M, Lanteri S, Dal Molin A, Minio A, Ferrarini A, ... Rotino GL (2016). A high quality eggplant genome sequence: a new tool for the analysis of the Solanaceae family evolution and for the molecular deciphering of complex traits. In: Kölliker R, Boller B (Eds). Plant breeding: the art of bringing science to life - 20th Eucarpia General Congress. Zurich pp 122.

Barrantes W, Fernández-del-Carmen A, López-Casado G, González-Sánchez MÁ, Fernández-Muñoz R, Granell A, Monforte AJ (2014). Highly efficient genomics-assisted development of a library of introgression lines of *Solanum pimpinellifolium*. Molecular Breeding 34:1817-1831.

Bolger A, Scossa F, Bolger M, Lanz C, Maumus F (2014). The genome of the stress-tolerant wild tomato species *Solanum pennellii*. Nature Genetics 46:1034-1038.

Brown TA (2006). Genomes. Garland Science (3th ed), New York.

Brozynska M, Furtado A, Henry RJ (2016). Genomics of crop wild relatives: expanding the gene pool for crop improvement. Plant Biotechnology Journal 14:1070-1085.

Buckler ES, Ilut DC, Wang X, Kretzschmar T, Gore MA, Mitchell SE (2016). rAmpSeq: Using repetitive sequences for robust genotyping. bioRxiv 096628.

Cericola F, Portis E, Lanteri S, Toppino L, Barchi L, Acciarri N, ... Rotino GL (2014). Linkage disequilibrium and genome-wide association analysis for anthocyanin pigmentation and fruit color in eggplant. BMC Genomics 15:896.

Cheng CY, Krishnakumar V, Chan A, Schobel S, Town CD (2016). Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. bioRxiv 047308.

Davey J, Hohenlohe P, Etter P, Boone J (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Reviews Genetics 12:499-510.

Doganlar S, Frary A, Daunay M, Lester R (2002a). A comparative genetic linkage map of eggplant (*Solanum melongena*) and its implications for genome evolution in the Solanaceae. Genetics 161:1697-1711.

Doganlar S, Frary A, Daunay M, Lester R (2002b). Conservation of gene function in the Solanaceae as revealed by comparative mapping of domestication traits in eggplant. Genetics 161:1713-1726.

Doganlar S, Frary A, Daunay MC, Huvenaars K, Mank R, Frary A (2014). High resolution map of eggplant (*Solanum melongena*) reveals extensive chromosome rearrangement in domesticated members of the Solanaceae. Euphytica 198:231-241

Elshire R, Glaubitz J, Sun Q, Poland J (2011). A robust simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6:e19379.

FAOSTAT (2017). Food and Agriculture Organization of the United Nations. Retrieved 2017 June 1 from http://www.fao.org/faostat/en/.

Fernandez-Silva I, Whitney J, Wainwright B (2013). Microsatellites for next-generation ecologists: a post-sequencing bioinformatics pipeline. PLoS One 8:e55990.

Feuillet C, Leach JE, Rogers J, Schnable PS, Eversole K (2011). Crop genome sequencing: lessons and rationales. Trends in Plant Science 16:77-88.

Filippi CV, Aguirre N, Rivas JG, Zubrzycki J, Puebla A, Cordes D, ... Lia VV (2015). Population structure and genetic diversity characterization of a sunflower association mapping population using SSR and SNP markers. BMC Plant Biology 15:52.

Frary A, Doganlar S, Daunay M, Tanksley SD (2003). QTL analysis of morphological traits in eggplant and implications for conservation of gene function during evolution of solanaceous species. Theoretical and Applied Genetics 107:359-370.

Frary A, Frary A, Daunay MC, Huvenaars K, Mank R, Doganlar S (2014). QTL hotspots in eggplant (*Solanum melongena*) detected with a high resolution map and CIM analysis. Euphytica 197:211-228.

Fukuoka H, Miyatake K, Nunome T, Negoro S, Shirasawa K, Isobe S, ... Ohyama A (2012). Development of gene-based markers and construction of an integrated linkage map in eggplant by using *Solanum* orthologous (SOL) gene sets. Theoretical and Applied Genetics 125:47-56.

Fulton T, Hoeven R, Van der Eannetta N (2002). Identification analysis and utilization of conserved ortholog set markers for comparative genomics in higher plants. The Plant Cell 14:1457-1467.

Gabriel S, Ziaugra L, Tabbaa D (2009). SNP genotyping using the sequenom massARRAY iPLEX Platform. Current Protocols in Human Genetics http://doi.org/10.1002/0471142905.hg0212s60.

Ge H, Liu Y, Jiang M, Zhang J, Han H, Chen H (2013). Analysis of genetic diversity and structure of eggplant populations (*Solanum melongena* L.) in China using simple sequence repeat markers. Scientia Horticulturae 162:71-75.

Gonzaga ZJ (2015). Evaluation of SSR and SNP markers for molecular breeding in rice. Plant Breeding and Biotechnology 3:139-152.

Goodwin S, McPherson J, McCombie W (2016). Coming of age: ten years of next-generation sequencing technologies. Nature Reviews Genetics 17:333-351.

Gousset C, Collonnier C, Mulya K, Mariska I, Rotino GL (2005). *Solanum torvum* as a useful source of resistance against bacterial and fungal diseases for improvement of eggplant (*S. melongena* L.). Plant Science 168:319-327.

Gramazio P, Prohens J, Plazas M, Andújar I, Herraiz FJ, Castillo E, ... Vilanova S (2014). Location of chlorogenic acid biosynthesis pathway and polyphenol oxidase genes in a new interspecific anchored linkage map of eggplant. BMC Plant Biology 14:350.

Gramazio P, Blanca J, Ziarsolo P, Herraiz FJ, Plazas M, Prohens J, Vilanova S (2016a). Transcriptome analysis and molecular marker discovery in *Solanum incanum* and *S. aethiopicum*, two close relatives of the common eggplant (*Solanum melongena*) with interest for breeding. BMC Genomics 17:300.

Gramazio P, Prohens J, Plazas M, Herraiz FJ, Ziarsolo P, Cañizares J, Vilanova S (2016b). GBS-assisted recovery of "lost" introgressions in advanced backcrosses of *Solanum incanum* to cultivated eggplant (*S. melongena*). In: Book of Abstracts of the 13th Annual Solanaceae Conference 2016. Davis, USA pp 38.

Hirakawa H, Shirasawa K, Miyatake K, Nunome T, Negoro S, Ohyama A, ... Fukuoka H (2014). Draft genome sequence of eggplant (*Solanum melongena* L.): The representative *Solanum* species indigenous to the old world. DNA Research 21:649-660.

Hong CP, Piao ZY, Kang TW, Batley J, Yang T, Hur Y (2007). Genomic distribution of Simple Sequence Repeats in *Brassica rapa*. Molecules and Cells 23:349-356.

Hurtado M, Vilanova S, Plazas M, Gramazio P, Herraiz FJ, Andújar I, Prohens J (2013). Phenomics of fruit shape in eggplant (*Solanum melongena* L.) using Tomato Analyzer software. Scientia Horticulturae 164:625-632.

Jones E, Sullivan H, Bhattramakki D, Smith J (2007). A comparison of simple sequence repeat and single nucleotide polymorphism marker technologies for the genotypic analysis of maize (*Zea mays* L.). Theoretical and Applied Genetics 115:361-371.

Kalia RK, Rai MK, Kalia S, Singh R, Dhawan AK (2011). Microsatellite

markers: an overview of the recent progress in plants. Euphytica 177:309-334.

Kim C, Guo H, Kong W, Chandnani R, Shuang LS, Paterson AH (2016). Application of genotyping by sequencing technology to a variety of crop breeding programs. Plant Science 242:14-22.

Knapp S, Vorontsova S, Prohens J (2013). Wild relatives of the eggplant (*Solanum melongena*: Solanaceae): New understanding of species names in a complex group. PLOS ONE 8:e57039.

Kouassi B, Prohens J, Gramazio P, Kouassi A, Vilanova S, Galán-Ávila A, ... Plazas M (2016). Development of backcross generations and new interspecific hybrid combinations for introgression breeding in eggplant (*Solanum melongena*). Scientia Horticulturae 213:199-207.

Lebeau A, Gouy M, Daunay M, Wicker E (2013). Genetic mapping of a major dominant gene for resistance to *Ralstonia solanacearum* in eggplant. Theoretical and Applied Genetics 126:143-158.

Lester N, Daunay (2003). Diversity of African vegetable *Solanum* species and its implications for a better understanding of plant domestication. Schriften zu Genetischen Ressourcen 22:137-152.

Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, ... Wang J (2010). De novo assembly of human genomes with massively parallel short read sequencing. Genome Research 20:265-272.

Mavromatis K, Land ML, Brettin TS, Quest DJ, Copeland A, Clum A, ... Kyrpides NC (2012). The fast changing landscape of sequencing technologies and their impact on microbial genome assemblies and annotation. PLoS One 12:e48837.

Miyatake K, Saito T, Negoro S, Yamaguchi H, Nunome T, Ohyama A, Fukuoka H (2012). Development of selective markers linked to a major QTL for parthenocarpy in eggplant (*Solanum melongena* L.). Theoretical and Applied Genetics 124:1403-1413.

Morrell PL, Buckler ES, Ross-Ibarra J (2012). Crop genomics: Advances and applications. Nature Reviews Genetics 13:85-96.

Nilsson N, Halldén C, Hansen M, Hjerdin A (1997). Comparing the distribution of RAPD and RFLP markers in a high density linkage map of sugar beet. Genome 40:644-651.

Nunome T, Ishiguro K, Yoshida T, Hirai M (2001). Mapping of fruit shape and color development traits in eggplant (*Solanum melongena* L.) based on RAPD and AFLP markers. Breeding Science 51:19-26.

Nunome T, Suwabe K, Iketani H, Hirai M (2003a). Identification and characterization of microsatellites in eggplant. Plant Breeding 122:256-262.

Nunome T, Suwabe K, Ohyama A, Fukuoka H (2003b). Characterization of Trinucleotide Microsatellites in Eggplant. Breeding Science 53:77-83.

Nunome T, Negoro S, Kono I, Kanamori H, Miyatake K, Yamaguchi H, ... Fukuoka H (2009). Development of SSR markers derived from SSR-enriched genomic library of eggplant (*Solanum melongena* L.). Theoretical and Applied Genetics 119:1143-1153.

Ozsolak F, Milos P (2011). RNA sequencing: advances challenges and opportunities. Nature Reviews Genetics 12:87-98.

Pascual L, Desplat N, Huang BE, Desgroux A, Bruguier L, Bouchet JP, ... Causse M (2015). Potential of a tomato MAGIC population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era. Plant Biotechnology Journal 13:565-577.

Plazas M, Andújar I, Vilanova S, Gramazio P, Herraiz FJ, Prohens J (2014). Conventional and phenomics characterization provides insight into the diversity and relationships of hypervariable scarlet (*Solanum aethiopicum* L.) and gboma (*S. macrocarpon* L.) eggplant complexes. Frontiers in Plant Science 5:318.

Plazas M, Vilanova S, Gramazio P, Rodríguez-Burruezo A, Fita A, Herraiz FJ, ... Prohens J (2016). Interspecific hybridization between eggplant and wild relatives from different genepools. Journal of the American Society for Horticultural Science 141:34-44.

Portis E, Barchi L, Toppino L, Lanteri S, Acciarri N, Felicioni N, ... Rotino GL (2014). QTL mapping in eggplant reveals clusters of yield-related loci and orthology with the tomato genome. PLoS One 9:e89499.

Portis E, Cericola F, Barchi L, Toppino L, Acciarri N, Pulcini L, ... Rotino GL (2015). Association mapping for fruit, plant and leaf morphology traits in eggplant. PLoS One 10:e0135200.

Potato Genome Sequencing Consortium (2011). Genome sequence and analysis of the tuber crop potato. Nature 475:189-95.

Prohens J, Whitaker BD, Plazas M, Vilanova S, Hurtado M, Blasco M, ... Stommel JR (2013). Genetic diversity in morphological characters and phenolic acids content resulting from an interspecific cross between eggplant *Solanum melongena* and its wild ancestor (*S. incanum*). Annals of Applied Biology 162:242-257.

Qin C, Yu C, Shen Y, Fang X, Chen L, Min J, ... Zhang Z (2014). Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. Proceedings of the National Academy of Sciences of the United States of America 111:5135-5140.

Rafalski J (2002). Novel genetic mapping tools in plants: SNPs and LD-based approaches. Plant Science 162:329-333.

Ramesh KR, Hemalatha R, Vijayendra CA, Arshi UZS, Dushyant SB, Dinesh KB (2016). Transcriptome analysis of *Solanum melongena* L. (eggplant) fruit to identify putative allergens and their epitopes. Gene 576:64-71.

Redden R, Yadav S, Maxted N, Dulloo M, Guarino L (2015). Crop wild relatives and climate change. Wiley-Blackwell.

Rinaldi R, Van Deynze A, Portis E, Rotino GL, Toppino L, Hill T, ... Lanteri S (2016). New insights on eggplant/tomato/pepper synteny and identification of eggplant and pepper orthologous QTL. Frontiers in Plant Science 7:1031.

Rotino GL, Sala T, Toppino L (2014). Eggplant. In: Pratap A, Kumar J (Eds). Alien gene transfer in crop plants, Vol 2. Springer, New York pp 381-409.

Salgon S, Jourda C, Sauvage C, Daunay MC, Reynaud B, Wicker E, Dintinger J (2017). Eggplant resistance to the *Ralstonia solanacearum* species complex involves both broad-spectrum and strain-specific quantitative trait loci. Frontiers in Plant Science 8:828.

The Tomato Genome Consortium (2012). The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485:635-641.

Scheben A, Batley J, Edwards D (2017). Genotyping-by-sequencing approaches to characterize crop genomes: choosing the right tool for the right application. Plant Biotechnology Journal 15:149-161.

Sevon P, Toivonen H, Onkamo P (2005). Gene mapping by pattern discovery. Data Mining in Bioinformatics 2005:105-126.

Shirasawa K, Asamizu E, Fukuoka H, Ohyama A, Sato S, Nakamura Y, … Isobe S (2010). An interspecific linkage map of SSR and intronic polymorphism markers in tomato. Theoretical and Applied Genetics 121:731-739.

Stàgel A, Portis E, Toppino L, Rotino G, Lanteri S (2008). Gene-based microsatellite development for mapping and phylogeny studies in eggplant. BMC Genomics 9:357.

Sunseri F, Sciancalepore A, Martelli G, Acciarri N, Rotino GL, Valentino D, Tamietti G (2003). Development of RAPD-AFLP map of eggplant and improvement of tolerance to *Verticillium* wilt. Acta Horticulturae 625:107-115.

Sunseri F, Polignano GB, Alba V, Lotti C, Bisignano V, Mennella G, … Ricciardi L (2010). Genetic diversity and characterization of African eggplant germplasm collection. African Journal of Plant Science 4:231-241.

Syfert MM, Castañeda-Álvarez NP, Khoury CK, Särkinen T, Sosa CC, Achicanoy HA, … Knapp S (2016). Crop wild relatives of the brinjal eggplant (*Solanum melongena*): poorly represented in genebanks and many species at risk of extinction. American Journal of Botany 103:635-651.

Tanksley S, Ganal M, Prince J (1992). High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141-1160.

Thapa S, Miyao E, Davis R, Coaker G (2015). Identification of QTLs controlling resistance to *Pseudomonas syringae* pv. tomato race 1 strains from the wild tomato *Solanum habrochaites* LA1777. Theoretical and Applied Genetics 128:681-692.

Thomson MJ, Alfred J, Dangl J, Kamoun S, McCouch S, Baird N, … McCouch S (2014). High-throughput SNP genotyping to accelerate crop improvement. Plant Breeding and Biotechnology 2:195-212.

Toppino L, Valè G, Rotino GL (2008). Inheritance of *Fusarium* wilt resistance introgressed from *Solanum aethiopicum* Gilo and Aculeatum groups into cultivated eggplant (*S. melongena)* and development. Molecular Breeding 22:237-250.

Toppino L, Barchi L, Lo Scalzo R, Palazzolo E, Francese G, Fibiani M, … Rotino GL (2016). Mapping quantitative trait loci affecting biochemical and morphological fruit properties in eggplant (*Solanum melongena* L.). Frontiers in Plant Science 7:256

Rajeev V, Andreas G, Mark S (2005). Genic microsatellite markers in plants: features and applications. Trends in Biotechnology 23:48-55.

Vilanova S, Manzur J, Prohens J (2012). Development and characterization of genomic simple sequence repeat markers in eggplant and their application to the study of diversity and relationships in a collection of different cultivar types and origins. Molecular Breeding 30:647-660.

Víquez-Zamora M, Vosman B, Geest A, Visser RG, Finkers R, … Voorrips R (2013). Tomato breeding in the genomics era: insights from a SNP array. BMC Genomics 14:354

Waterhouse PM, Helliwell CA (2003). Exploring plant genomes by RNA-induced gene silencing. Nature Reviews Genetics 4:29-38.

Whitaker BD, Stommel JR (2003). Distribution of hydroxycinnamic acid conjugates in fruit of commercial eggplant (*Solanum melongena* L.) cultivars. Journal of Agricultural and Food Chemistry 51:3448-3454.

Wu F, Eannetta NT, Xu Y, Tanksley SD (2009). A detailed synteny map of the eggplant genome based on conserved ortholog set II (COSII) markers. Theoretical and Applied Genetics 118:927-935.

Xiao M, Zhang Y, Chen X, Lee EJ, Barber CJS, Chakrabarty R, … Sensen CW (2013). Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. Journal of Biotechnology 166:122-134.

Yamaguchi H, Fukuoka H, Arao T, Ohyama A, Nunome T, Miyatake K, ... Negoro S (2010). Gene expression analysis in cadmium-stressed roots of a low cadmium-accumulating solanaceous plant *Solanum torvum*. Journal of Experimental Botany 61:423-437.

Yan J, Yang X, Shah T, Sánchez-Villeda H, Li J, Warburton M, … Xu Y (2010). High-throughput SNP genotyping with the Goldengate assay in Maize. Molecular Breeding 25:441-451.

Yang L, Jue D, Li W, Zhang R, Chen M, Yang Q (2013). Identification of MiRNA from eggplant (*Solanum melongena* L.) by small RNA deep sequencing and their response to *Verticillium dahliae* infection. PLOS ONE 8:e72840.

Yang X, Cheng YF, Deng C, Ma Y, Wang ZW, Chen XH, Xue LB (2014). Comparative transcriptome analysis of eggplant (*Solanum melongena* L.) and turkey berry (*Solanum torvum* Sw.): phylogenomics and disease resistance analysis. BMC Genomics 15:412.

Yuan Y, Bayer PE, Batley J, Edwards D (2017). Improvements in genomic technologies: application to crop genomics. Trends in Biotechnology 35:547-558.

Zhou X, Bao S, Liu J, Zhuang Y (2016). De novo sequencing and analysis of the transcriptome of the wild eggplant species *Solanum aculeatissimum* in response to *Verticillium dahliae*. Plant Molecular Biology Reporter 34:1193-1203.