Mark A. Chapman   *Editor*

# The Eggplant Genome

Springer

# Compendium of Plant Genomes

**Series Editor**

Chittaranjan Kole, Raja Ramanna Fellow, Government of India,
ICAR-National Research Center on Plant Biotechnology, Pusa, New Delhi,
India

Whole-genome sequencing is at the cutting edge of life sciences in the new millennium. Since the first genome sequencing of the model plant Arabidopsis thaliana in 2000, whole genomes of about 100 plant species have been sequenced and genome sequences of several other plants are in the pipeline. Research publications on these genome initiatives are scattered on dedicated web sites and in journals with all too brief descriptions. The individual volumes elucidate the background history of the national and international genome initiatives; public and private partners involved; strategies and genomic resources and tools utilized; enumeration on the sequences and their assembly; repetitive sequences; gene annotation and genome duplication. In addition, synteny with other sequences, comparison of gene families and most importantly potential of the genome sequence information for gene pool characterization and genetic improvement of crop plants are described.

**Interested in editing a volume on a crop or model plant?** Please contact Dr. Kole, Series Editor, at ckoleorg@gmail.com

More information about this series at http://www.springer.com/series/11805

Mark A. Chapman
Editor

# The Eggplant Genome

Springer

*Editor*
Mark A. Chapman
Biological Sciences
University of Southampton
Southampton, UK

*This book series is dedicated to my wife Phullara,*
*and our children Sourav and Devleena*

Chittaranjan Kole

# Preface to the Series

Genome sequencing has emerged as the leading discipline in the plant sciences coinciding with the start of the new century. For much of the twentieth century, plant geneticists were only successful in delineating putative chromosomal location, function, and changes in genes indirectly through the use of a number of "markers" physically linked to them. These included visible or morphological, cytological, protein, and molecular or DNA markers. Among them, the first DNA marker, the RFLPs, introduced a revolutionary change in plant genetics and breeding in the mid-1980s, mainly because of their infinite number and thus potential to cover maximum chromosomal regions, phenotypic neutrality, absence of epistasis, and codominant nature. An array of other hybridization-based markers, PCR-based markers, and markers based on facilitated construction of genetic linkage maps, mapping of genes controlling simply inherited traits, and even gene clusters (QTLs) controlling polygenic traits in a large number of model and crop 88 plants have been identified. During this period, a number of new mapping populations beyond $F_2$ were utilized and a number of computer programs were developed for map construction, mapping of genes, and for mapping of polygenic clusters or QTLs. Molecular markers were also used in studies of evolution and phylogenetic relationship, genetic diversity, DNA fingerprinting, and map-based cloning. Markers tightly linked to the genes were used in crop improvement employing the so-called marker-assisted selection. These strategies of molecular genetic mapping and molecular breeding made a spectacular impact during the last one and a half decades of the twentieth century. But still they remained "indirect" approaches for elucidation and utilization of plant genomes since much of the chromosomes remained unknown and the complete chemical depiction of them was yet to be unraveled.

Physical mapping of genomes was the obvious consequence that facilitated the development of the "genomic resources" including BAC and YAC libraries to develop physical maps in some plant genomes. Subsequently, integrated genetic–physical maps were also developed in many plants. This led to the concept of structural genomics. Later on, emphasis was laid on EST and transcriptome analysis to decipher the function of the active gene sequences leading to another concept defined as functional genomics. The advent of techniques of bacteriophage gene and DNA sequencing in the 1970s was extended to facilitate sequencing of these genomic resources in the last decade of the twentieth century.

As expected, sequencing of chromosomal regions would have led to too much data to store, characterize, and utilize with the-then available computer software could handle. But development of information technology made the life of biologists easier by leading to a swift and sweet marriage of biology and informatics, and a new subject was born—bioinformatics.

Thus, the evolution of the concepts, strategies, and tools of sequencing and bioinformatics reinforced the subject of genomics—structural and functional. Today, genome sequencing has traveled much beyond biology and involves biophysics, biochemistry, and bioinformatics!

Thanks to the efforts of both public and private agencies, genome sequencing strategies are evolving very fast, leading to cheaper, quicker, and automated techniques right from clone-by-clone and whole-genome shotgun approaches to a succession of second-generation sequencing methods. The development of software of different generations facilitated this genome sequencing. At the same time, newer concepts and strategies were emerging to handle sequencing of the complex genomes, particularly the polyploids.

It became a reality to chemically—and so directly—define plant genomes, popularly called whole-genome sequencing or simply genome sequencing.

The history of plant genome sequencing will always cite the sequencing of the genome of the model plant Arabidopsis thaliana in 2000 that was followed by sequencing the genome of the crop and model plant rice in 2002. Since then, the number of sequenced genomes of higher plants has been increasing exponentially, mainly due to the development of cheaper and quicker genomic techniques and, most importantly, development of collaborative platforms such as national and international consortia involving partners from public and/or private agencies.

As I write this preface for the first volume of the new series "Compendium of Plant Genomes," a net search tells me that complete or nearly complete whole-genome sequencing of 45 crop plants, eight crop and model plants, eight model plants, 15 crop progenitors and relatives, and three basal plants is accomplished, the majority of which are in the public domain. This means that we nowadays know many of our model and crop plants chemically, i.e., directly, and we may depict them and utilize them precisely better than ever. Genome sequencing has covered all groups of crop plants. Hence, information on the precise depiction of plant genomes and the scope of their utilization are growing rapidly every day. However, the information is scattered in research articles and review papers in journals and dedicated Web pages of the consortia and databases. There is no compilation of plant genomes and the opportunity of using the information in sequence-assisted breeding or further genomic studies. This is the underlying rationale for starting this book series, with each volume dedicated to a particular plant.

Plant genome science has emerged as an important subject in academia, and the present compendium of plant genomes will be highly useful to both students and teaching faculties. Most importantly, research scientists involved in genomics research will have access to systematic deliberations on the plant genomes of their interest. Elucidation of plant genomes is of interest not only for the geneticists and breeders, but also for practitioners of an array

of plant science disciplines, such as taxonomy, evolution, cytology, physiology, pathology, entomology, nematology, crop production, biochemistry, and obviously bioinformatics. It must be mentioned that information regarding each plant genome is ever-growing. The contents of the volumes of this compendium are, therefore, focusing on the basic aspects of the genomes and their utility. They include information on the academic and/ or economic importance of the plants, description of their genomes from a molecular genetic and cytogenetic point of view, and the genomic resources developed. Detailed deliberations focus on the background history of the national and international genome initiatives, public and private partners involved, strategies and genomic resources and tools utilized, enumeration on the sequences and their assembly, repetitive sequences, gene annotation, and genome duplication. In addition, synteny with other sequences, comparison of gene families, and, most importantly, the potential of the genome sequence information for gene pool characterization through genotyping by sequencing (GBS) and genetic improvement of crop plants have been described. As expected, there is a lot of variation of these topics in the volumes based on the information available on the crop, model, or reference plants.

I must confess that as the series editor, it has been a daunting task for me to work on such a huge and broad knowledge base that spans so many diverse plant species. However, pioneering scientists with lifetime experience and expertise on the particular crops did excellent jobs editing the respective volumes. I myself have been a small science worker on plant genomes since the mid-1980s, and that provided me the opportunity to personally know several stalwarts of plant genomics from all over the globe. Most, if not all, of the volume editors are my longtime friends and colleagues. It has been highly comfortable and enriching for me to work with them on this book series. To be honest, while working on this series I have been and will remain a student first, a science worker second, and a series editor last. And I must express my gratitude to the volume editors and the chapter authors for providing me the opportunity to work with them on this compendium.

I also wish to mention here my thanks and gratitude to the Springer staff, Dr. Christina Eckey and Dr. Jutta Lindenborn in particular, for all their constant and cordial support right from the inception of the idea.

I always had to set aside additional hours to edit books beside my professional and personal commitments—hours I could and should have given to my wife, Phullara, and our kids, Sourav and Devleena. I must mention that they not only allowed me the freedom to take away those hours from them but also offered their support in the editing job itself. I am really not sure whether my dedication of this compendium to them will suffice to do justice to their sacrifices for the interest of science and the science community.

New Delhi, India                                                                                   Chittaranjan Kole

# Contents

# Introduction: The Importance of Eggplant

## Mark A. Chapman

### Abstract

In this chapter, I highlight how the eggplant, whilst being globally dwarfed by other members of the Solanaceae, notably potato and tomato, offers a number of important ecological, evolutionary and agronomic features making it unique and interesting, warranting further study. It also highlights the parallels and differences between Solanaceous crops. The eggplant genome is in the process of being finalised, and once this is available to researchers, it is likely we will see a surge of papers utilising this resource for understanding the genetic basis of these important traits.

## 1.1 Overview

The Asian eggplant (*Solanum melongena* L.), known as aubergine in Britain and France, and brinjal in Southern Asia and South Africa, is a widely grown species from the nightshade family (Solanaceae). The fruit is popular in a range of cuisines and is an important part of the diet in many countries, especially India and Bangladesh, Southeast Asia and the Middle East (Daunay and

M. A. Chapman (✉)
Biological Sciences, University of Southampton, Southampton SO17 1BJ, UK
e-mail: m.chapman@soton.ac.uk

Lester 1988). S*olanum* is a large genus of ca. 1400 species (D'Arcy 1991; Frodin 2004), several of which are poisonous to humans, most famously the nightshades (e.g. *S. dulcamara* L.). Eggplant is an Old World crop, domesticated in Asia, whereas its congeners, potato (*S. tuberosum* L.) and tomato (*S. lycopersicum* L.), are New World (South American) representatives of the genus (Daunay and Lester 1988; Weese and Bohs 2010).

The focus of this book is the Asian eggplant, *S. melongena*; however, two other *Solanum* species are known as eggplants, the Ethiopian/scarlet eggplant (*S. aethiopicum* L.) and the African/Gboma eggplant (*S. macrocarpon* L.). These two species are minor crops globally relative to Asian eggplant, but may be locally important, with the fruits and leaves of both species used for food and medicine. The similarities between the three eggplant species have previously caused taxonomic confusion; however, it is clear now they are relatively distantly related within the genus (Weese and Bohs 2010).

Several non-exclusive theories have been proposed concerning the origin of Asian eggplants, *S. melongena* (hereafter simply 'eggplant'). The general consensus (Knapp et al. 2013; Weese and Bohs 2010) is that the African/Middle Eastern species *S. incanum* L. was transported, intentionally or otherwise, into Indo-China where the true wild progenitor, *S. insanum* L., evolved, from which *S. melongena* is derived. The first domesticates are possibly

represented now by relatively small-fruited *S. ovigerum*, and the landraces and other cultivated types are derived from this group. This is described in further detail in Chap. 12. More recent molecular evidence has suggested that eggplant was domesticated more than once (Meyer et al. 2012b); however, this remains contentious (see Chap. 12). Ancient literature suggests that eggplant has been used as a food for more than 2000 years in China (Wang et al. 2008), and traces of *Solanum* have been found in Harappan cooking vessels from ca. 4500 YBP in India (Kashyap and Weber 2013).

Traditionally the majority of research into crops of the Solanaceae has focussed on tomato and potato, presumably due to the relative economic importance of these three crops. The first potato genome was published in 2011 (Xu et al. 2011b) and tomato in 2012 (Sato et al. 2012). A draft genome of eggplant followed, however, this was incomplete, covering approximately 74% of the genome, and fragmented, being represented by 33,873 scaffolds (Hirakawa et al. 2014); see also Chap. 6. Eggplant (and wild relatives) nonetheless offer a range of features absent from potato and tomato, especially high tolerances to a number of pests and pathogens (Daunay 2008; Salgon et al. 2017) and tolerance of abiotic stresses (Fita et al. 2015; Keatinge et al. 2014).

Beyond the economic importance of eggplant, the similarities between the domestication of tomato and eggplant (and to a lesser degree peppers, *Capsicum annuum* L.) pose an interesting study system for investigating parallel and/or convergent evolution. Domestication of these members of the Solanaceae has involved an increase in fruit size and alteration in fruit colour and shape, and comparative mapping suggests these traits may be controlled (in part) by the same suite of genes (Doganlar et al. 2002). This is discussed briefly below and in more depth in Chap. 4.

For these reasons, the origin and evolution of eggplant pose some interesting and important questions for a range of researchers, which are beginning to be addressed utilising modern technologies.

## 1.2 Economic Importance of Eggplant

Eggplant is the third most widely grown Solanaceous vegetable after potatoes and tomatoes (Fig. 1.1a). Eggplant is especially popular in cuisines of Southeast Asia and the Mediterranean. In 2016, the area of potato and tomato harvested worldwide was greater than that of eggplant by a factor of 10 and almost three, respectively (FAO 2017). Eggplant production has increased steadily since FAOSTAT data was collected (1961) to about 1.79 M ha (51.3 M tonnes) worldwide (2016 data). Over 80% of eggplants are produced in China and India with only five other countries (Egypt, Indonesia, Turkey, Iran and the Philippines) growing more than 1% of the world's total production (Fig. 1.1b).

Whilst eggplant is not known for being high in the majority of health-related micronutrients, it is very low fat and low calorie. The Solanaceae as a whole, however, are a rich source of nutritionally and pharmaceutically useful compounds, partly
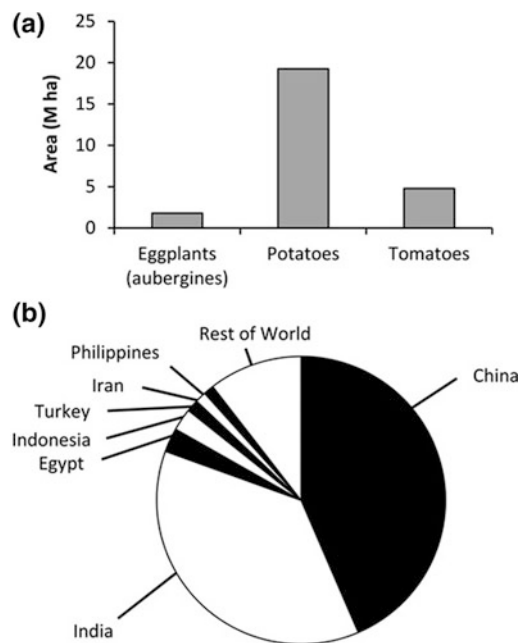


**Fig. 1.1** Area of eggplant harvested in 2016. **a** Compared to congeners potato and tomato, **b** subdivided by country (only countries growing >1% of the world's eggplant are named). Data is from FAO (2017)

explaining the large number of species in the Solanaceae used for food or medicine (Hawkes 1999). In eggplant, a number of phytonutrients are present, especially hydroxycinnamic acid (HCA) conjugates, potentially involved in consumer health, fruit taste and texture (see Meyer et al. 2015, Chap. 3 and references therein). HCA conjugates are free radical scavengers and hence may play a role in mediating oxidative stress (Ma et al. 2011). A number of compounds differ in their abundance between wild and domesticated eggplant, with the domesticates containing overall a lower total abundance of HCA conjugates (Meyer et al. 2015). The quantity of total and individual HCA conjugates varies widely amongst accessions of eggplant (Stommel and Whitaker 2003). Interestingly, literature surveys and interviews demonstrate that eggplant is used as a medicine for different illnesses (especially for gastrointestinal, immune system and cardiovascular ailments) in different parts of the world (Meyer et al. 2014). This could have influenced the trajectory of eggplant domestication and improvement in different parts of the world.

Eggplant is also valued as a rootstock for tomatoes because of resistance to certain diseases and nematodes. For example, data from Vietnam suggested that grafted tomatoes (onto a range of rootstocks, including eggplant) out-yield non-grafted tomatoes by around one third (Genova et al. 2013). Eggplant rootstocks have also been shown to confer waterlogging tolerance to tomatoes (Bahadur et al. 2015), and eggplant wild relatives are also used as rootstocks because of resistance to pathogens, for example verticillium wilt (Bletsos et al. 2003).

## 1.3 Academic Importance of Eggplant

### 1.3.1 Eggplant as a Model for Parallel Evolution

As mentioned above, the domestication of multiple members of the Solanaceae has been used as a model to study convergent evolution. During domestication, human selection on fruit colour, taste, shape and size has been pervasive across many crops (Meyer et al. 2012a), and in grass crops, some traits are controlled by orthologous genes, for example loss of shattering (i.e. a loss of natural seed dispersal) in cereal crops (Lin et al. 2012). If the same genes are involved in the domestication of multiple crops, then knowledge of the genetic basis of these traits in one crop can be transferred across crops.

Tomato and eggplant exhibit a number of conspicuous similarities in their domestication syndromes, but similarities with potato and pepper in the suite of traits that were selected by humans are also evident, especially fruit shape and size in pepper, and colour, albeit tuber and flower colour, in potato. When the first genetic maps of eggplant were produced, it was noted, based on comparing genetic maps from other Solanaceae species, that a number of quantitative trait loci (QTL) for eggplant domestication traits were found in the same genomic regions as those in tomato, pepper and potato (Doganlar et al. 2002). These findings are detailed in Chap. 4.

### 1.3.2 Eggplant Wild Relatives for Crop Improvement

Until recently, the potential for eggplant wild relatives to improve cultivated eggplants has been discussed but little progress has been made. This is disappointing because crop wild relatives (CWRs) are often major sources of alleles for biotic and abiotic tolerances (Dempewolf et al. 2017). One major hindrance to the breeding of wild species' alleles into cultivated eggplant has been the lack of a genome sequence (Gramazio et al. 2018). This absence prevents the development of genome-anchored markers which are necessary for efficient trait transfer, e.g. through marker-assisted selection (MAS; Morrell et al. 2011). Nonetheless, success in introgressing Fusarium wilt resistance from *S. aethiopicum* and Verticillium wilt resistance from *S. linnaeanum* Hepper & P.-M.L. Jaeger into eggplant has been carried out (Toppino et al. 2008; Liu et al. 2015).

More recently, however, a large number of mapping populations from crosses between eggplant and its CWRs have been generated (Kaushik et al. 2016). This extensive work, utilising multiple cultivated eggplants and ten wild species, has seen the development of dozens of backcross populations with potential for introgressing traits of interest from wild relatives into eggplant. This, and other work (e.g. Ranil et al. 2017), also highlights the ease with which closely related *S. incanum* and *S. insanum* can be crossed with eggplant, making them important candidates for aiding eggplant improvement through introgression. Members of the secondary gene pool can be crossed with varying success to eggplant, and crosses with more distantly related species are only successful through embryo rescue (Kouassi et al. 2016). Interestingly, the cross-compatibility of different eggplant accessions with the wild species is variable, suggesting that introgression from the wild species might be easier in some eggplant varieties than in others. Extensive morphological phenotyping of the parents and F1 progeny has allowed the identification of wild species which exhibit the best potential for eggplant improvement (Kaushik et al. 2016).

Gaining a better understanding of the genetic basis of adaptive phenotypes in eggplant CWRs has been a prominent goal in eggplant breeding for years, and with the advent of high-throughput sequencing (HTS), a number of wild species have been subjected to analysis with the purpose of generating molecular markers necessary for downstream genetic mapping, diversity analysis and candidate gene analysis. Through transcriptome sequencing, i.e. sequencing the expressed portion of the genome from one or more tissue(s) at one or more time-point(s), only a subset of the genome is sequenced, decreasing the cost considerably (Ozsolak and Milos 2010). In addition, the output of a transcriptome investigation is made up primarily of coding sequences, thus enriching for regions of the genome that may code for traits of interests (although it is acknowledged that the genetic basis of some traits is not attributable to the coding portion of the genome; Wilusz et al. 2009).

Simple sequence repeat (SSR) markers (aka microsatellites) can be identified from a single transcriptome sequence (e.g. Chapman 2015; White et al. 2016), acknowledging that some proportion of these markers will not be variable when tested across multiple individuals. Single nucleotide polymorphism (SNP) markers, however, require the comparison of at least two individuals to identify polymorphisms before designing SNP assays. Through the comparison of transcriptomes from four *Solanum* species, marker databases have been developed for the wild relatives *Solanum incanum* and *S. aethiopicum* (Gramazio et al. 2016). Transcriptome sequences for *S. torvum* Sw. (Yang et al. 2014) and *S. aculeatissimum* Jacq. (Zhou et al. 2016) are available (see also below), and a comparative transcriptomic investigation could be employed to identify molecular markers.

### 1.3.3 Transcriptomics in Eggplant

Despite the absence of a complete genome sequence, eggplant studies have harnessed the power of HTS for a range of studies. HTS-generated molecular markers (specifically using RADseq and genotyping-by-sequencing approaches) have been employed to aid in genetic map construction and to understand genetic diversity within eggplant accessions and between eggplant and its CWRs; however, these are covered in detail in later chapters and will not be discussed here. Instead, here I focus on two set of studies which have employed transcriptomics to (1) examine the gene expression response to pathogen infection, and (2) understand the genetic basis of anthocyanin accumulation. These investigations highlight the potential for eggplant to serve as a model species for a number of adaptive traits.

#### 1.3.3.1 Eggplant and Its Relatives as a Model for Understanding Pathogen Infection

Whilst eggplant is susceptible to several pathogens, a number of the eggplant CWRs offer

promise for resistance to these bacteria and fungi, which reduce crop yields (of eggplants and other crops) substantially if left unchecked. Efforts to understand the genetic basis of pathogen resistance are therefore of paramount importance (Piquerez et al. 2014).

*Solanum torvum* is an eggplant CWR and is often used as a rootstock because of strong resistance to several soil-borne pathogens, notably the root-knot nematode *Meloidogyne incognita* (Gousset et al. 2005). In order to understand more about the genes involved in root-knot nematode resistance, gene expression analysis was carried out on *S. torvum* plants with and without infection (Bagnaresi et al. 2013), revealing almost 400 genes which were differentially expressed (DE) under pathogen infection. Gene ontology analysis showed that many of these genes were related to known pathogen response genes, especially chitinases which are often upregulated in plants exposed to nematode infection (e.g. Qtu et al. 1997).

Another eggplant CWR with pathogen resistance is *Solanum aculeatissimum*, resistant to verticillium wilt which negatively affects eggplant worldwide, as well as tomato, potato and cotton (Klosterman et al. 2009). Gene expression analysis was carried out comparing verticillium-treated and verticillium-untreated *S. aculeatissimum* seedlings and revealed thousands of genes which were DE (Zhou et al. 2016). Genetic pathways relating to the biosynthesis of secondary metabolites (especially hormone signal transduction and phenylpropanoid biosynthesis) were found to be enriched in the lists of DE genes, indicating that the mechanisms of verticillium wilt resistance in *S. aculeatissimum* could be very similar to the pathways in cotton and tomato (Gayoso et al. 2010; Xu et al. 2011a).

Both of these case studies suggest that the mechanisms of pathogen resistance in *Solanum* species may be conserved across crops, with the potential for cross-crop transfer of knowledge to expedite breeding and crop improvement.

### 1.3.3.2 Eggplant as a Model to Understanding Anthocyanin Accumulation in Plants

Anthocyanins are red, blue and purple pigments found in some groups of plants. Anthocyanins have potent antioxidant properties in vitro (e.g. De Rosso et al. 2008); however, evidence for a positive role in humans remains elusive (Lotito and Frei 2006). Nevertheless, understanding the biosynthesis of these pigments is of interest because of the visual attractiveness, as highlighted by the popularity of recently developed blue tomatoes.

Eggplants exhibit a significant variation in fruit colours and hence pose a model for understanding anthocyanin biosynthesis. Using a cultivar in which the purple pigmentation is induced by exposure to light, recent research has identified genes which may play a role in initiating the biosynthesis of anthocyanins even before any purple pigmentation is evident (Li et al. 2017, 2018). Understandably a number of genes that were differentially expressed in the light-exposed fruit were from the anthocyanin biosynthesis pathway (Li et al. 2017); however, other genes not known to be involved in the regulation of anthocyanins were unearthed (Li et al. 2018), posing new targets for breeding both in eggplant and in other species.

Recently, successful modification of fruit colour in *S. aethiopicum* has been carried out through the transgenic expression of an eggplant gene, *SmMYB1* (Zhang et al. 2016).

### 1.3.3.3 Eggplant as a Model for Understanding the Wider Effects of Genetic Modification

The transgenic improvement of crops through transfer of genes from different organisms has promise for feeding the future growing population under climate change (James 2003), yet public perception is mixed, and in some cases

strongly negative. Part of this negativity comes from a lack of published knowledge about the consequences to other organisms of growing a genetically modified (GM) crop. These non-target effects, for example the effects on soil microbes (Liu et al. 2005) and food webs (Groot and Dicke 2002), remain understudied.

Several GM eggplants possessing a *Cry1Ac* toxin from *Bacillus thuringiensis* have been developed (collectively 'Bt brinjal') and were released in India in 2009, but soon a moratorium caused their use to be prohibited. This was based largely on public outcry and the insistence of anti-GM groups who called for more research (Herring 2015). In 2013, in Bangladesh, four varieties of Bt brinjal were released after seven years of trials. Putting the controversy to one side, these Bt brinjals allow the assessment of the potential for a GM crop to be effective as well as to assess their non-target effects, as follows:

Firstly, Bt brinjal (including lines developed with different Cry1 alleles) has been shown to be effective against eggplant shoot and fruit borer (ESFB) both under controlled (glasshouse) conditions (Rai et al. 2013) and in the field (Hautea et al. 2016). Infestation of Bt brinjal with ESFB was shown to be almost zero in field trials in the Philippines (Hautea et al. 2016).

Second, non-target effects have been investigated for both soil microbes and non-target arthropods. Singh et al. (2013) demonstrated that soil microbe abundance was lower in plots of Bt brinjal than non-Bt counterparts. In addition, different species of bacteria were present in soils of the two treatments (Bt brinjal and non-Bt brinjal) (Singh et al. 2013) indicating the potential for microbial non-target effects. Conversely, in field trials, there was no significant difference in the non-target arthropod communities of Bt and non-Bt brinjal indicating that the Bt brinjal is selective in its control of the ESFB (Navasero et al. 2016).

Finally, it is clear that in natural environments, there is a high potential for crop eggplant to hybridise with wild eggplants (Davidar et al. 2015; see also Chap. 12), which could be an avenue for transgene escape (Chapman and Burke 2006). The consequences for the escape of the Bt transgene into wild relatives are not known and require further research.

## 1.4 Conclusions

Despite eggplant being less economically important than its congeners, it serves as an important model for a number of agronomic and evolutionary processes; hence, the development of a genome sequence represents an important step forward in these fields of research. The following chapters discuss our current knowledge of eggplant as a crop in its own right as well as a model for understanding genome evolution, domestication and speciation.

## Literature Cited

Bagnaresi P, Sala T, Irdani T, Scotto C, Lamontanara A, Beretta M, Rotino GL, Sestili S, Cattivelli L, Sabatini E (2013) *Solanum torvum* responses to the root-knot nematode *Meloidogyne incognita*. BMC Genom 14:540. https://doi.org/10.1186/1471-2164-14-540

Bahadur A, Rai N, Kumar R, Tiwari SK, Singh AK, Rai AK, Singh U, Patel PK, Tiwari V, Rai AB, Singh M, Singh BB (2015) Grafting tomato on eggplant as a potential tool to improve waterlogging tolerance in hybrid tomato. Veg Sci 42:82–87

Bletsos F, Thanassoulopoulos C, Roupakias D (2003) Effect of grafting on growth, yield, and verticillium wilt of eggplant. HortScience 38:183–186

Chapman MA (2015) Transcriptome sequencing and marker development for four underutilized legumes. Appl Plant Sci 3(2):1400111

Chapman MA, Burke JM (2006) Letting the gene out of the bottle: the population genetics of genetically modified crops. New Phytol 170(3):429–443

D'Arcy WG (1991) The solanaceae since 1976, with a review of its biogeography. In: Hawkes JG, Lester RN, Nee M, Estrada N (eds) Solanaceae III: taxonomy, chemistry, evolution. Royal Botanic Gardens, Kew, pp 75–137

Daunay MC (2008) Eggplant. In: Prohens J, Nuez F (eds) Vegetables II: fabaceae, liliaceae, solanaceae, and umbelliferae. Springer, New York, pp 163–220

Daunay MC, Lester RN (1988) The usefulness of taxonomy for Solanaceae breeders, with special reference to the genus *Solanum* and to *Solanum melongena* L. (eggplant). Capsicum Newslett 7:70–79

Davidar P, Snow AA, Rajkumar M, Pasquet R, Daunay MC, Mutegi E (2015) The potential for crop to wild hybridization in eggplant (*Solanum melongena*; Solanaceae) in southern India. Am J Bot 102(1):129–139. https://doi.org/10.3732/ajb.1400404

De Rosso VV, Morán Vieyra FE, Mercadante AZ, Borsarelli CD (2008) Singlet oxygen quenching by anthocyanin's flavylium cations. Free Radical Res 42(10):885–891. https://doi.org/10.1080/10715760802506349

Dempewolf H, Baute G, Anderson J, Kilian B, Smith C, Guarino L (2017) Past and future use of wild relatives in crop breeding. Crop Sci 57(3):1070–1082. https://doi.org/10.2135/cropsci2016.10.0885

Doganlar S, Frary A, Daunay MC, Lester RN, Tanksley SD (2002) Conservation of gene function in the Solanaceae as revealed by comparative mapping of domestication traits in eggplant. Genetics 161(4):1713–1726

FAO (2017) FAOSTAT database collections (http://faostat.fao.org). Food and Agriculture Organization of the United Nations, Rome, Italy. Sept 2017

Fita A, Fioruci F, Plazas M, Rodriguez-Burruezo A, Prohens J (2015) Drought tolerance among accessions of eggplant and related species. Bull Univ Agric Sci Vet Med Cluj Napoca Hortic 72:461–462

Frodin DG (2004) History and concepts of big plant genera. Taxon 53:753–776

Gayoso C, Pomar F, Novo-Uzal E, Merino F, de Ilarduya OM (2010) The Ve-mediated resistance response of the tomato to *Verticillium dahliae* involves $H_2O_2$, peroxidase and lignins and drives PAL gene expression. BMC Plant Biol 10:232. https://doi.org/10.1186/1471-2229-10-232

Genova C, Schreinemachers P, Afari-Sefa V (2013) An impact assessment of AVRDC's tomato grafting in Vietnam. AVRDC—The World Vegetable Center, Shanhua

Gousset C, Collonnier C, Mulya K, Mariska I, Rotino GL, Besse P, Servaes A, Sihachakr D (2005) *Solanum torvum*, as a useful source of resistance against bacterial and fungal diseases for improvement of eggplant (*S. melongena* L.). Plant Sci 168(2):319–327. https://doi.org/10.1016/j.plantsci.2004.07.034

Gramazio P, Blanca J, Ziarsolo P, Herraiz FJ, Plazas M, Prohens J, Vilanova S (2016) 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 Genom 17:300. https://doi.org/10.1186/s12864-016-2631-4

Gramazio P, Prohens J, Plazas M, Mangino G, Herraiz FJ, Garcia-Fortea E, Vilanova S (2018) Genomic tools for the enhancement of vegetable crops: a case in eggplant. Not Bot Horti Agrobo Cluj Napoca 46(1):1–13. https://doi.org/10.15835/nbha46110936

Groot AT, Dicke M (2002) Insect-resistant transgenic plants in a multi-trophic context. Plant J 31:387–406

Hautea DM, Taylo LD, Masanga APL, Sison MLJ, Narciso JO, Quilloy RB, Hautea RA, Shotkoski FA, Shelton AM (2016) Field performance of Bt eggplants (*Solanum melongena* L.) in the Philippines: Cry1Ac expression and control of the eggplant fruit and shoot borer (*Leucinodes orbonalis* Guenee). PLOS ONE 11(6):e0157498. https://doi.org/10.1371/journal.pone.0157498

Hawkes JG (1999) The economic importance of the family Solanaceae. In: Nee M, Symon DE, Lester RN, Jessop JP (eds) Solanaceae IV: advance in biology and utilization. Royal Botanic Gardens, Kew, pp 1–9

Herring RJ (2015) State science, risk and agricultural biotechnology: Bt cotton to Bt Brinjal in India. J Peasant Stud 42(1):159–186. https://doi.org/10.1080/03066150.2014.951835

Hirakawa H, Shirasawa K, Miyatake K, Nunome T, Negoro S, Ohyama A, Yamaguchi H, Sato S, Isobe S, Tabata S, Fukuoka H (2014) Draft genome sequence of eggplant (*Solanum melongena* L.): the representative *Solanum* species indigenous to the Old World. DNA Res 21(6):649–660. https://doi.org/10.1093/dnares/dsu027

James C (2003) Global review of commercialized transgenic crops. Curr Sci 84(3):303–309

Kashyap A, Weber S (2013) Starch grain analysis and experiments provide insights into Harappan cooking practices. In: Abraham SA, Gullapalli P, Raczek TP, Rizvi UZ (eds) Connections and complexity: new approaches to the archaeology of South Asia. Left Coast Press, Taylor and Francis, Oxford, pp 177–193

Kaushik P, Prohens J, Vilanova S, Gramazio P, Plazas M (2016) Phenotyping of eggplant wild relatives and interspecific hybrids with conventional and phenomics descriptors provides insight for their potential utilization in breeding. Front Plant Sci 7:677. https://doi.org/10.3389/fpls.2016.00677

Keatinge JDH, Lin LJ, Ebert AW, Chen WY, Hughes JdA, Luther GC, Wang JF, Ravishankar M (2014) Overcoming biotic and abiotic stresses in the Solanaceae through grafting: current status and future perspectives. Biol Agric Hortic 30(4):272–287. https://doi.org/10.1080/01448765.2014.964317

Klosterman SJ, Atallah ZK, Vallad GE, Subbarao KV (2009) Diversity, pathogenicity, and management of verticillium species. Annu Rev Phytopathol 47(1):39–62. https://doi.org/10.1146/annurev-phyto-080508-081748

Knapp S, Vorontsova MS, Prohens J (2013) Wild relatives of the eggplant (*Solanum melongena* L.: Solanaceae): new understanding of species names in a complex group. PLoS ONE 8(2):e57039. https://doi.org/10.1371/journal.pone.0057039

Kouassi B, Prohens J, Gramazio P, Kouassi AB, Vilanova S, Galan-Avila A, Herraiz FJ, Kouassi A, Segui-Simarro JM, Plazas M (2016) Development of backcross generations and new interspecific hybrid combinations for introgression breeding in eggplant (*Solanum melongena*). Sci Hortic 213:199–207. https://doi.org/10.1016/j.scienta.2016.10.039

Li J, Ren L, Gao Z, Jiang MM, Liu Y, Zhou L, He YJ, Chen HY (2017) Combined transcriptomic and

proteomic analysis constructs a new model for light-induced anthocyanin biosynthesis in eggplant (*Solanum melongena* L.). Plant Cell Environ 40 (12):3069–3087. https://doi.org/10.1111/pce.13074

Li J, He YJ, Zhou L, Liu Y, Jiang MM, Ren L, Chen HY (2018) Transcriptome profiling of genes related to light-induced anthocyanin biosynthesis in eggplant (*Solanum melongena* L.) before purple color becomes evident. BMC Genom 19:324. https://doi.org/10.1186/s12864-018-4587-z

Lin Z, Li X, Shannon LM, Yeh C-T, Wang ML, Bai G, Peng Z, Li J, Trick HN, Clemente TE, Doebley J, Schnable PS, Tuinstra MR, Tesso TT, White F, Yu J (2012) Parallel domestication of the *Shattering1* genes in cereals. Nat Genet 44(6):720–724. https://doi.org/10.1038/ng.2281

Liu B, Zeng Q, Yan FM, Xu HG, Xu CR (2005) Effects of transgenic plants on soil microorganisms. Plant Soil 271(1–2):1–13

Liu J, Zheng Z, Zhou X, Feng C, Zhuang Y (2015) Improving the resistance of eggplant (*Solanum melongena*) to Verticillium wilt using wild species *Solanum linnaeanum*. Euphytica 201(3):463–469. https://doi.org/10.1007/s10681-014-1234-x

Lotito SB, Frei B (2006) Consumption of flavonoid-rich foods and increased plasma antioxidant capacity in humans: Cause, consequence, or epiphenomenon? Free Radical Biol Med 41(12):1727–1746. https://doi.org/10.1016/j.freeradbiomed.2006.04.033

Ma C, Dastmalchi K, Whitaker BD, Kennelly EJ (2011) Two new antioxidant malonated caffeoylquinic acid isomers in fruits of wild eggplant relatives. J Agric Food Chem 59(17):9645–9651. https://doi.org/10.1021/jf202028y

Meyer RS, DuVal AE, Jensen HR (2012a) Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops. New Phytol 196(1):29–48. https://doi.org/10.1111/j.1469-8137.2012.04253.x

Meyer RS, Karol KG, Little DP, Nee MH, Litt A (2012b) Phylogeographic relationships among Asian eggplants and new perspectives on eggplant domestication. Mol Phylogen Evol 63:685–701

Meyer RS, Bamshad M, Fuller DQ, Litt A (2014) Comparing medicinal uses of eggplant and related solanaceae in china, india, and the Philippines suggests the independent development of uses, cultural diffusion, and recent species substitutions. Econ Bot 68(2):137–152. https://doi.org/10.1007/s12231-014-9267-6

Meyer RS, Whitaker BD, Little DP, Wu S-B, Kennelly EJ, Long C-L, Litt A (2015) Parallel reductions in phenolic constituents resulting from the domestication of eggplant. Phytochemistry 115:194–206. https://doi.org/10.1016/j.phytochem.2015.02.006

Morrell PL, Buckler ES, Ross-Ibarra J (2011) Crop genomics: advances and applications. Nat Rev Genet 13:85. https://doi.org/10.1038/nrg3097

Navasero MV, Candano RN, Hautea DM, Hautea RA, Shotkoski FA, Shelton AM (2016) Assessing potential impact of Bt eggplants on non-target arthropods in the Philippines. PLoS ONE 11(10):e0165190. https://doi.org/10.1371/journal.pone.0165190

Ozsolak F, Milos PM (2010) RNA sequencing: advances, challenges and opportunities. Nat Rev Genet 12:87. https://doi.org/10.1038/nrg2934

Piquerez SJM, Harvey SE, Beynon JL, Ntoukakis V (2014) Improving crop disease resistance: lessons from research on *Arabidopsis* and tomato. Front Plant Sci 5:671. https://doi.org/10.3389/fpls.2014.00671

Qtu J, Hallmann J, Kokalis-Burelle N, Weaver DB, Rodriguez-Kabana R, Tuzun S (1997) Activity and differential induction of chitinase isozymes in soybean cultivars resistant or susceptible to root-knot nematodes. J Nematol 29(4):523–530

Rai NP, Rai GK, Kumar S, Kumari N, Singh M (2013) Shoot and fruit borer resistant transgenic eggplant (*Solanum melongena* L.) expressing cry1Aa3 gene: development and bioassay. Crop Protect 53:37–45. https://doi.org/10.1016/j.cropro.2013.06.005

Ranil RHG, Prohens J, Aubriot X, Niran HML, Plazas M, Fonseka RM, Vilanova S, Fonseka HH, Gramazio P, Knapp S (2017) *Solanum insanum* L. (subgenus Leptostemonum Bitter, Solanaceae), the neglected wild progenitor of eggplant (*S. melongena* L.): a review of taxonomy, characteristics and uses aimed at its enhancement for improved eggplant breeding. Genet Resour Crop Evol 64(7):1707–1722. https://doi.org/10.1007/s10722-016-0467-z

Salgon S, Jourda C, Sauvage C, Daunay M-C, 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. Front Plant Sci 8:828

Sato S, Tabata S, Hirakawa H, Asamizu E, Shirasawa K, Isobe S, Kaneko T, Nakamura Y, Shibata D, Aoki K, Egholm M, Knight J, Bogden R, Li C, Shuang Y, Xu X, Pan S, Cheng S, Liu X, Ren Y, Wang J, Albiero A, Dal Pero F, Todesco S, Van Eck J, Buels RM, Bombarely A, Gosselin JR, Huang M, Leto JA, Menda N, Strickler S, Mao L, Gao S, Tecle IY, York T, Zheng Y, Vrebalov JT, Lee J, Zhong S, Mueller LA, Stiekema WJ, Ribeca P, Alioto T, Yang W, Huang S, Du Y, Zhang Z, Gao J, Guo Y, Wang X, Li Y, He J, Li C, Cheng Z, Zuo J, Ren J, Zhao J, Yan L, Jiang H, Wang B, Li H, Li Z, Fu F, Chen B, Han B, Feng Q, Fan D, Wang Y, Ling H, Xue Y, Ware D, McCombie WR, Lippman ZB, Chia J-M, Jiang K, Pasternak S, Gelley L, Kramer M, Anderson LK, Chang S-B, Royer SM, Shearer LA, Stack SM, Rose JKC, Xu Y, Eannetta N, Matas AJ, McQuinn R, Tanksley SD, Camara F, Guigo R, Rombauts S, Fawcett J, Van de Peer Y, Zamir D, Liang C, Spannagl M, Gundlach H, Bruggmann R, Mayer K, Jia Z, Zhang J, Ye Z, Bishop GJ, Butcher S, Lopez-Cobollo R, Buchan D, Filippis I,

Abbott J, Dixit R, Singh M, Singh A, Pal JK, Pandit A, Singh PK, Mahato AK, Dogra V, Gaikwad K, Sharma TR, Mohapatra T, Singh NK, Causse M, Rothan C, Schiex T, Noirot C, Bellec A, Klopp C, Delalande C, Berges H, Mariette J, Frasse P, Vautrin S, Zouine M, Latche A, Rousseau C, Regad F, Pech J-C, Philippot M, Bouzayen M, Pericard P, Osorio S, Fernandez del Carmen A, Monforte A, Granell A, Fernandez-Munoz R, Conte M, Lichtenstein G, Carrari F, De Bellis G, Fuligni F, Peano C, Grandillo S, Termolino P, Pietrella M, Fantini E, Falcone G, Fiore A, Giuliano G, Lopez L, Facella P, Perrotta G, Daddiego L, Bryan G, Orozco M, Pastor X, Torrents D, van Schriek KNVMGM, Feron RMC, van Oeveren J, de Heer P, daPonte L, Jacobs-Oomen S, Cariaso M, Prins M, van Eijk MJT, Janssen A, van Haaren MJJ, Jo S-H, Kim J, Kwon S-Y, Kim S, Koo D-H, Lee S, Hur C-G, Clouser C, Rico A, Hallab A, Gebhardt C, Klee K, Joecker A, Warfsmann J, Goebel U, Kawamura S, Yano K, Sherman JD, Fukuoka H, Negoro S, Bhutty S, Chowdhury P, Chattopadhyay D, Datema E, Smit S, Schijlen EWM, van de Belt J, van Haarst JC, Peters SA, van Staveren MJ, Henkens MHC, Mooyman PJW, Hesselink T, van Ham RCHJ, Jiang G, Droege M, Choi D, Kang B-C, Kim BD, Park M, Kim S, Yeom S-I, Lee Y-H, Choi Y-D, Li G, Gao J, Liu Y, Huang S, Fernandez-Pedrosa V, Collado C, Zuniga S, Wang G, Cade R, Dietrich RA, Rogers J, Knapp S, Fei Z, White RA, Thannhauser TW, Giovannoni JJ, Angel Botella M, Gilbert L, Gonzalez R, Goicoechea JL, Yu Y, Kudrna D, Collura K, Wissotski M, Wing R, Schoof H, Meyers BC, Gurazada AB, Green PJ, Mathur S, Vyas S, Solanke AU, Kumar R, Gupta V, Sharma AK, Khurana P, Khurana JP, Tyagi AK, Dalmay T, Mohorianu I, Walts B, Chamala S, Barbazuk WB, Li J, Guo H, Lee T-H, Wang Y, Zhang D, Paterson AH, Wang X, Tang H, Barone A, Chiusano ML, Ercolano MR, D'Agostino N, Di Filippo M, Traini A, Sanseverino W, Frusciante L, Seymour GB, Elharam M, Fu Y, Hua A, Kenton S, Lewis J, Lin S, Najar F, Lai H, Qin B, Qu C, Shi R, White D, White J, Xing Y, Yang K, Yi J, Yao Z, Zhou L, Roe BA, Vezzi A, D'Angelo M, Zimbello R, Schiavon R, Caniato E, Rigobello C, Campagna D, Vitulo N, Valle G, Nelson DR, De Paoli E, Szinay D, de Jong HH, Bai Y, Visser RGF, Lankhorst RMK, Beasley H, McLaren K, Nicholson C, Riddle C, Gianese G, Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485(7400):635–641. https://doi.org/10.1038/nature11119

Singh AK, Rai GK, Singh M, Dubey SK (2013) Bacterial community structure in the rhizosphere of a Cry1Ac Bt-Brinjal crop and comparison to its non-transgenic counterpart in the tropical soil. Microb Ecol 66(4):927–939. https://doi.org/10.1007/s00248-013-0287-z

Stommel JR, Whitaker BD (2003) Phenolic acid composition of eggplant fruit in a germplasm core subset. J Am Soc Hort Sci 128:704–710

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 of associated PCR-based markers. Mol Breed 22(2):237–250. https://doi.org/10.1007/s11032-008-9170-x

Wang JX, Gao TG, Knapp S (2008) Ancient Chinese literature reveals pathways of eggplant domestication. Ann Bot 102(6):891–897. https://doi.org/10.1093/aob/mcn179

Weese TL, Bohs L (2010) Eggplant origins: out of Africa, into the Orient. Taxon 59:49–56

White OW, Doo B, Carine MA, Chapman MA (2016) Transcriptome sequencing and simple sequence repeat marker development for three Macaronesian endemic plant species. Appl Plant Sci 4(8):1600050. https://doi.org/10.3732/apps.1600050

Wilusz JE, Sunwoo H, Spector DL (2009) Long noncoding RNAs: functional surprises from the RNA world. Genes Dev 23(13):1494–1504. https://doi.org/10.1101/gad.1800909

Xu L, Zhu L, Tu L, Liu L, Yuan D, Jin L, Long L, Zhang X (2011a) Lignin metabolism has a central role in the resistance of cotton to the wilt fungus *Verticillium dahliae* as revealed by RNA-seq-dependent transcriptional analysis and histochemistry. J Exp Bot 62(15):5607–5621. https://doi.org/10.1093/jxb/err245

Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, Zhang G, Yang S, Li R, Wang J, Orjeda G, Guzman F, Torres M, Lozano R, Ponce O, Martinez D, De la Cruz G, Chakrabarti SK, Patil VU, Skryabin KG, Kuznetsov BB, Ravin NV, Kolganova TV, Beletsky AV, Mardanov AV, Di Genova A, Bolser DM, Martin DMA, Li G, Yang Y, Kuang H, Hu Q, Xiong X, Bishop GJ, Sagredo B, Mejia N, Zagorski W, Gromadka R, Gawor J, Szczesny P, Huang S, Zhang Z, Liang C, He J, Li Y, He Y, Xu J, Zhang Y, Xie B, Du Y, Qu D, Bonierbale M, Ghislain M, del Rosario Herrera M, Giuliano G, Pietrella M, Perrotta G, Facella P, O'Brien K, Feingold SE, Barreiro LE, Massa GA, Diambra L, Whitty BR, Vaillancourt B, Lin H, Massa A, Geoffroy M, Lundback S, DellaPenna D, Buell CR, Sharma SK, Marshall DF, Waugh R, Bryan GJ, Destefanis M, Nagy I, Milbourne D, Thomson SJ, Fiers M, Jacobs JME, Nielsen KL, Sonderkaer M, Iovene M, Torres GA, Jiang J, Veilleux RE, Bachem CWB, de Boer J, Borm T, Kloosterman B, van Eck H, Datema E, Hekkert BTL, Goverse A, van Ham RCHJ, Visser RGF, Potato Genome Sequencing Consortium (2011b) Genome sequence and analysis of the tuber crop potato. Nature 475(7355):189–194. https://doi.org/10.1038/nature10158

Yang X, Cheng Y-F, Deng C, Ma Y, Wang Z-W, Chen X-H, Xue L-B (2014) Comparative transcriptome analysis of eggplant (*Solanum melongena* L.) and turkey berry (*Solanum torvum* Sw.): phylogenomics and disease resistance analysis. BMC Genom 15 (1):412. https://doi.org/10.1186/1471-2164-15-412

Zhang YJ, Chu GH, Hu ZL, Gao Q, Cui BL, Tian SB, Wang B, Chen GP (2016) Genetically engineered anthocyanin pathway for high health-promoting pigment production in eggplant. Mol Breed 36(5):54. https://doi.org/10.1007/s11032-016-0454-2

Zhou XH, Bao SY, 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 Mol Biol Rep 34(6):1193–1203. https://doi.org/10.1007/s11105-016-0998-7

# Eggplant (*Solanum melongena* L.): Taxonomy and Relationships

Sandra Knapp, Xavier Aubriot and Jaime Prohens

**Abstract**

*Solanum melongena* L. (brinjal eggplant) is a member of a small monophyletic group (Eggplant clade) of mainly andromonoecious species in the large and diverse Leptostemonum clade of *Solanum* (previously referred to as subgenus *Leptostemonum* Bitter). The Leptostemonum clade (also known as the spiny solanums) is the most diverse monophyletic group in the species-rich genus *Solanum* and contains more than 500 species occurring on all continents except Antarctica. In this chapter, we summarise the current state of knowledge of the taxonomy and phylogeny of *Solanum*, the Leptostemonum clade and that of the monophyletic group of Old World taxa to which *S. melongena* belongs. We provide a species list with distributions of the

currently recognised members of the Eggplant clade and discuss character evolution and biogeography in the group in the context of phylogeny.

## 2.1 Introduction

The brinjal eggplant (*Solanum melongena* L.) is one of approximately 1300 species in the extremely species-rich genus *Solanum* L. in the nightshade family Solanaceae. The family comprises 101 genera, including many economic and horticultural importance such as *Nicotiana* L. (the tobaccos, see Knapp et al. 2004) and *Petunia* L. (Stehmann et al. 2000). Generic diversity in the family is concentrated in the Americas, but there have been several instances of long-distance dispersal giving rise to genera and/or groups that are endemic to the Old World (Dupin et al. 2017). Generic limits in the family are under active investigation, and new genera have been included (e.g. *Nolana* L.f. and *Sclerophylax* Miers, traditionally recognised as separate families; see Olmstead et al. 2008) and segregated based on new understanding from molecular phylogenetics (e.g. *Trompettia* Dupin & S.S.Smith; Dupin and Smith 2018). With the inclusion of previously segregated genera such as *Lycopersicon* Mill., *Cyphomandra* Sendtn. and *Normania* Lowe, *Solanum* is resolved as strongly monophyletic and as sister to the genus *Jaltomata* Schltdl. (Särkinen et al. 2013).

S. Knapp (✉)
Department of Life Sciences, Natural History Museum, Cromwell Road, London SW7 5BD, UK
e-mail: s.knapp@nhm.ac.uk

X. Aubriot
Laboratoire Écologie, Systématique et Évolution, UMR 8079, Université Paris-Sud/CNRS/AgroParisTech, Orsay, France

J. Prohens
Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, Universitat Politècnica de València, Camino de Vera 14, 46022 Valencia, Spain

*Solanum* comprises around half of the species diversity of the family and is one of only a handful of flowering plant genera with more than 1000 species (Frodin 2004). Not only because of its large size, *Solanum* is also important for containing species of great economic importance for humans, such as potato (*S. tuberosum* L.), tomato (*S. lycopersicum* L.) and of course the eggplant, plus a host of minor fruit and leaf crops cultivated locally worldwide (see Anderson 1977; Whalen et al. 1981; Särkinen et al. 2018). Species of *Solanum* occur on all continents except Antarctica, in a wide variety of habitats from tropical rainforests to the driest deserts and have a wide range of life forms, from annual herbs to rainforest trees. The traditional view has been that the large majority of species of *Solanum* occurred in the New World, mostly in South America (e.g. see D'Arcy 1972) but recent work in Africa (Vorontsova and Knapp 2016), Asia (Aubriot et al. 2016) and Australia and New Guinea (e.g. Bean 2002, 2004, 2010, 2011, 2014, 2016; Bean and Albrecht 2008) has revealed hitherto poorly understood diversity in the Old World, especially in the spiny solanums (see below).

The genus was divided into two main groups by authors in the nineteenth and early twentieth centuries (e.g. Dunal 1852; Seithe 1962), simply described as the spiny and non-spiny solanums. These broad groups were defined on the presence or absence of prickles and anther shape (see Vorontsova and Knapp 2016 for a more complete discussion). Within those broad groups, *Solanum* was divided into a number of sections (see D'Arcy 1972, who listed many sections, subsections and series), defined largely on macro-morphological characteristics. Phylogenetic work using DNA sequences showed that most of these sectional groupings were not monophyletic (Bohs 2005); the genus can be divided into 13 major clades (Bohs 2005; Särkinen et al. 2013; Weese and Bohs 2007). Some of these (e.g. the Potato clade, including tomatoes and their relatives and a number of smaller groups such as section *Pteroidea* Dunal and the Regmandra clade; see Tepe et al. 2016) are well-supported and monophyletic, while the

relationships of others are less clear (e.g. species like *S. clandestinum* Bohs and *S. mapiriense* Bohs; see Särkinen et al. 2013). Relationships between the clades are relatively stable, but a polytomy at the base of Clade 2 of Särkinen et al. (2013) means the sister group of the largest and most species-rich clade of *Solanum*—the Leptostemonum clade or the spiny solanums—is not yet clear (Särkinen et al. 2013).

## 2.2 The Leptostemonum Clade

The prickly solanums are the largest monophyletic group within the genus *Solanum* (Bohs 2005; Särkinen et al. 2013; Stern et al. 2011). They were traditionally referred to as subgenus *Leptostemonum* Bitter (Bitter 1919), or as "chorus subgenerum" *Stellatipilum* Seithe (Seithe 1962), highlighting the two characters whose combination defined the group—the presence of stellate trichomes and long attenuate anthers. Neither of these is unique to the Leptostemonum clade. Stellate trichomes are found in the Brevantherum clade (Stern et al. 2013; Giacomin and Stehmann 2014) and attenuate anthers in two small groups, *S. nemorense* Dunal and relatives (Bohs 2005) and the *S. wendlandii* clade (Clark et al. 2016); all of these groups are part of the polytomy at the base of Clade 2 of Särkinen et al. (2013).

The Leptostemonum clade is strongly monophyletic (Stern et al. 2011) and comprises at current estimates some 560 accepted species distributed on all continents except Antarctica (S. Knapp, unpublished). Approximately half of these occur in the New World and half in the Old World (see Aubriot et al. 2016). Using DNA sequence data to delimit monophyletic groups within the Leptostemonum clade revealed that the Old World species were a single, monophyletic clade (with a few exceptions; Levin et al. 2006; Stern et al. 2011) rather than being related to diverse groups of New World taxa as had previously been thought (e.g. D'Arcy 1972; Whalen 1984). This large Old World group was derived within the spiny solanums and was sister to a small group of taxa that exhibit an