High resolution map of eggplant (*Solanum melongena*) reveals extensive chromosome rearrangement in domesticated members of the Solanaceae

Sami Doğanlar · Amy Frary · Marie-Christine Daunay · Koen Huvenaars · Rolf Mank · Anne Frary

Abstract A linkage map of eggplant was constructed for an interspecific F_2 population derived from a cross between *Solanum linnaeanum* MM195 and *S. melongena* MM738. The map contains 400 AFLP® (amplified fragment length polymorphism), 348 RFLP (restriction fragment length polymorphism) and 116 COSII (conserved ortholog set) markers. The 864 mapped markers encompass 12 linkage groups, span 1,518 cM and are spaced at an average interval of 1.8 cM. Use of orthologous markers allowed confirmation of the established syntenic relationships between eggplant and tomato chromosomes and helped delineate the nature of the 33 chromosomal rearrangements and 11 transpositions distinguishing the two species. This genetic map provides a 2- to 3-fold improvement in marker density compared to previously published interspecific maps. Because the interspecific mapping population is rich in morphological variation, this greater genome saturation will be useful for QTL (quantitative trait locus) analyses. The recent release of the tomato genome sequence will provide additional opportunities for exploiting this map for comparative genomics and crop improvement.

Electronic supplementary material The online version of this article (doi:10.1007/s10681-014-1096-2) contains supplementary material, which is available to authorized users.

S. Doğanlar (✉) · A. Frary
Plant Molecular Genetics Laboratory, Department of Molecular Biology and Genetics, Izmir Institute of Technology, Gulbahce Campus, Urla, Izmir 35430, Turkey
e-mail: samidoganlar@iyte.edu.tr

A. Frary
Department of Biological Sciences, Mount Holyoke College, South Hadley, MA 01075, USA

M.-C. Daunay
INRA, Unité de Génétique et Amélioration des Fruits et Légumes, Domaine Saint Maurice, BP 94, 84143 Montfaved Cedex, France

K. Huvenaars · R. Mank
Keygene N.V., Agro Business Park 90, 6708 PW Wageningen, The Netherlands

Introduction

Eggplant (*Solanum melongena* L.) is an important vegetable in the cuisines of Asian and Mediterranean countries. Agricultural production of eggplant is centered in these regions of the world with China and India providing more than 80 % of the crop in 2010 (FAO). Despite eggplant's economic importance (FAO ranks it 25th in top commodities), development of molecular genetic resources specific to the species has been limited as compared to other solanaceous crops. Initial genome mapping efforts in the

Solanaceae focused on the species of greatest agricultual importance in the family, tomato (*S. lycopersicum*), potato (*S. tuberosum*), and pepper (*Capsicum* spp.), and high-density molecular linkage maps for all three were constructed in the 1990s (Tanksley et al. 1992; Livingstone et al. 1999).

Molecular linkage maps are essential tools for genetic and breeding research in plant species. These maps can consist of one or more types of DNA and morphological markers and are used for mapping loci controlling both qualitative and quantitative traits, for marker-assisted selection and for positional gene cloning. Comparative linkage maps can be used to study chromosome and gene evolution across species. The first molecular linkage map of the eggplant genome was generated by Nunome et al. (1998) from an intraspecific cross between the *S. melongena* cultivars EPL-1 and WCGR112-9. Since that initial effort, several more maps based on the same population and including random amplified polymorphic DNA (RAPD), AFLP markers, and/or simple sequence repeat (SSR) markers have been published (Nunome et al. 2001, 2003, 2009). The most recent of these maps (Nunome et al. 2009) contains a total of 236 SSR markers encompassing 959 cM and 14 linkage groups with an average of one marker every 4.3 cM. Because of the tendency of SSR markers to reside in heterochromatic regions, portions of the genome are not covered by this map. Therefore, in an effort to create a saturated intraspecific map, Fukuoka et al. (2012) used two separate intraspecific F_2 mapping populations (developed from 'LS1934' × 'WCGR112-8' and 'AE-P03' × 'LS1934' parental crosses) to construct an integrated map containing 952 SSR, SNP, and insertion/deletion polymorphism (InDel) markers spread at an average distance of one marker every 1.4 cM over the 12 linkage groups. Over 465 of the SNP and InDel markers were derived from unigenes orthologous to eggplant, tomato, and potato. Of these, 70 % could be positioned in the tomato genome thereby allowing macrosyntenic relationships between the genomes of eggplant and tomato to be delineated.

Barchi et al. (2010, 2012) used an intraspecific F_2 population generated from eggplant breeding lines ('305E40' × '67/3') to construct two maps based on a combination of molecular marker types. A restriction-site associated DNA (RAD) strategy was used to develop the single nucleotide polymorphism (SNP) markers that form the chief marker type of their most recent map. This map comprises 415 markers (339 of which are SNPs)

spread over 1,390 cM and divided into 13 linkage groups with an average marker spacing of 3.8 cM (Barchi et al. 2012). The inclusion of shared RFLP and COSII markers allowed the linkage map to be compared to the tomato reference linkage map as well as the interspecific eggplant map of Wu et al. (2009). This comparison confirmed the large-scale genomic rearrangements between the tomato and eggplant genomes. In addition, eleven of the Barchi et al. (2012) linkage groups could be aligned with the Fukuoka et al. (2012) map based on the position of SSR and COSII markers. The utility of the Barchi et al. (2012) map for quantitative trait locus (QTL) analysis was assessed through the localization of several loci controlling anthocyanin pigmentation.

Maps of the eggplant genome have also been developed for interspecific F_2 populations from crosses between the wild relative *S. linnaeanum* (syn. *S. sodomeum*) and *S. melongena* (Doganlar et al. 2002a, Sunseri et al. 2003). Before the advent of SNP technology, such populations had the benefit of higher levels of detectable polymorphism. The first interspecific map *S. linnaeanum* 'MM195' × *S. melongena* 'MM738' (Doganlar et al. 2002a) took advantage of the close relationship between eggplant and tomato; 233 tomato restriction fragment length polymorphism (RFLP) markers were positioned in the eggplant genome, allowing for the first comparisons of genome evolution in eggplant, tomato and potato. Moreover, this map provided the foundation for the first quantitative trait mapping analyses in the crop (Doganlar et al. 2002b, Frary et al. 2003). Sunseri et al. (2003), using *S. linnaeanum* × *S. melongena* 'Buia', described a combined RAPD and AFLP marker map and reported preliminary mapping of *Verticillium* wilt resistance QTL in their interspecific F_2 population. As the markers were specific to eggplant, no comparison to the tomato linkage map was made.

The interspecific maps of the eggplant genome have been enhanced using the abundant genomic resources of the Solanaceae. Thus, 110 conserved ortholog set (COSII) markers [developed from single copy genes conserved between the Solanaceae and Rubiaceae (Wu et al. 2006)] were added to the Doganlar et al. (2002a) eggplant map (Wu et al. 2009). This augmented map improved the resolution of the previous map, bringing the average distance between framework markers down to 6.1 cM (from 7.6 cM) and allowing for a more complete characterization of the chromosomal rearrangements distinguishing the tomato and eggplant

genomes (Wu et al. 2009). Because of synteny between the tomato and eggplant genomes, the locations of an additional 522 COSII markers on the eggplant linkage map were deduced, bringing the total number of RFLP and COS markers of known position in this interspecific eggplant population to 869 (Wu et al. 2009). While in silico analysis is an efficient strategy for adding a large number of markers to a linkage map, maps produced in this way have disadvantages. Because the strategy relies on inference, marker positions are less reliable than those derived directly from genetic data. In the effort described here to generate a high-density map of the eggplant genome, we decided to implement AFLP marker technology. AFLP markers are cheaper and easier to develop and screen than RFLP markers and are generally more reliable than RAPD markers. Here we report a substantially improved version of the interspecific eggplant map of Doganlar et al. (2002a), achieved by the use of a larger F_2 population (108 vs. 58 individuals), the placement of 400 eggplant-specific AFLP markers as well as the addition of 117 previously unmapped RFLP and COSII markers.

Materials and methods

Mapping population

A mapping population of 108 F_2 individuals, including the 58 individuals used by Doganlar et al. (2002a) and 50 siblings not used in that publication, was obtained from a cross made by M.C. Daunay, Institut National de Recherche Agronomique, France. The female parent, *S. linnaeanum* Hepper & Jaeger 'MM195', is a spiny wild relative that produces small, round, green striped fruit. The male parent, *S. melongena* L. 'MM738', is a non-spiny European commercial type that bears large, oblong, purple fruit. Total genomic DNA samples of the parents and the mapping population were obtained from leaf tissue as described for tomato (Bernatzky and Tanksley 1986), stored, and used in the marker analyses.

RFLP and COSII analysis

RFLP markers were screened and genotyped on the 58 F_2 individuals of the original mapping population as

previously described (Doganlar et al. 2002a). Universal primers of COSII markers (Wu et al. 2006) were used to amplify orthologous segments from the two parents as described in Wu et al. (2009). COSII primers yielding amplicons of different size were then used to genotype the 58 F_2 individuals.

AFLP® analysis

AFLP analysis was carried out as described by Vos et al. (1995) using the enzyme combinations *Eco*RI/*Mse*I and *Pst*I/*Mse*I. A total of 10 primer combinations (PCs) were selected based on the total number of bands and the level of polymorphism observed when analyzing the two parental lines with 96 PCs. A two-step amplification strategy was followed. In a selective pre-amplification, both AFLP primers had a single selective nucleotide. A further selective amplification was carried out using primers having six selective nucleotides for *Eco*RI/*Mse*I (+3/+3), and five selective nucleotides for *Pst*I/*Mse*I (+2/+3). Gel analysis on a capillary electrophoresis system (MegaBACE) was performed according to Van Eijk et al. (2004). MegaBACE allows multiloading of three AFLP reactions in parallel, each reaction is labelled with a specific fluorophore. Only the *Eco*RI and *Pst*I-primers were end-labelled using fluorescent labels (FAM, NED and JOE). Pseudo gel images were generated and all AFLP markers were scored co-dominantly using proprietary software developed specifically for AFLP analysis at Keygene N.V. (Wageningen, Netherlands). This software allows the display and analysis of pseudo gel images and provides tools to navigate through the pixel images to size and quantify the AFLP bands with great precision. Each marker-specific band is classified with respect to its intensity using a mixture model of normal distributions, as described by Jansen et al. (2001). In this way, genotypic data was obtained from the full mapping population (108 F_2 individuals). Each polymorphic AFLP fragment was named according to the standard of using the code referring to the corresponding PC (see Table 1) followed by the estimated molecular length of the DNA fragment in nucleotides. A MegaBACE™ ET900-R size standard from Amersham Biosciences was used in each capillary to estimate the molecular weight of the fragments.
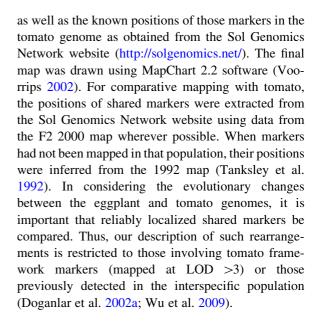
Map construction

The genetic map was constructed using CarteBlanche software (Buntjer 2002; Buntjer et al. 2002; Mank and Van Haaren 2003; commercially available from Keygene N.V.) as described in Truong et al. (2012). The first step in the mapping process was the assignment of markers to linkage groups (the genetic equivalent of chromosomes), based on observed recombination fractions. The next step involved the maximum likelihood estimation of pairwise recombination frequencies, corresponding LOD scores and genetic distances for all marker pairs in each linkage group (Kosambi 1944). For each linkage group, genetic maps were constructed using SpeedMap (Keygene N.V.) and Seriation (Buetow and Chakravarti 1987) mapping algorithms in 100-fold each and the Branch and Bound algorithm (Thompson 1987) in fivefold. After completion, the best map was selected out of these preliminary results, based on minimal sum-of-adjacent-recombination-frequencies and maximal sum-of-adjacent-LOD-scores (Buntjer 2002; Buntjer et al. 2002; Mank and Van Haaren 2003). Finally, the genetic distances in this map were optimized using methods described by Watowich et al. (1988).

The quality of the best map found was evaluated by plotting its marker order among those of the maps generated by the preliminary analysis. Frequent positioning of markers in one of the alternative maps in orders deviating from the best map indicated that either insufficient information was present to obtain a definite solution, or that a part of the locus segregation data was conflicting. Such markers were eliminated from the map. The best map we obtained with this comparative strategy consisted of 12 major and six minor linkage groups.

To resolve the issue of the minor linkage groups, the mapping data were additionally analyzed with the "group" command of MAPMAKER/EXP 3.0 software (Lander et al. 1987). The major and minor groups were then merged into 12 linkage groups using MAPMAKER/EXP's Kosambi mapping function (Kosambi 1944) to estimate genetic distances across junctions. Comparison with the Wu et al. (2009) RFLP–COSII map confirmed the identities and alignments of each linkage group. The placement of all RFLP and COSII loci was cross-checked against the previous maps (Doganlar et al. 2002a; Wu et al. 2009)

as well as the known positions of those markers in the tomato genome as obtained from the Sol Genomics Network website (http://solgenomics.net/). The final map was drawn using MapChart 2.2 software (Voorrips 2002). For comparative mapping with tomato, the positions of shared markers were extracted from the Sol Genomics Network website using data from the F2 2000 map wherever possible. When markers had not been mapped in that population, their positions were inferred from the 1992 map (Tanksley et al. 1992). In considering the evolutionary changes between the eggplant and tomato genomes, it is important that reliably localized shared markers be compared. Thus, our description of such rearrangements is restricted to those involving tomato framework markers (mapped at LOD >3) or those previously detected in the interspecific population (Doganlar et al. 2002a; Wu et al. 2009).

Results

Generation of AFLP marker data set

Markers showing single and double dose band intensities were found confirming the F_2 nature of the eggplant population. In total 404 AFLP markers were scored. The number of markers scored per primer combination (PC) ranged from a low of 17 (for PC P12/M61) to a high of 60 (for PC E44/M50) (Table 1). On average 40.4 markers were obtained per primer combination.

Table 1 Number of markers scored per primer combination (PC)

PC	# of markers
E32/M59	42
E35/M60	39
E38/M60	52
E39/M59	37
E44/M47	35
E44/M50	60
E44/M54	33
P12/M61	17
P14/M47	37
P15/M47	52
Total	404

Table 2 Marker loci at the boundaries of gaps spanning the merged linkage groups

Linkage group	Flanking marker loci	Estimated interval distance (cM)
3	TG585–TG520	18
5	CT172–CT51	15
	T0801–TG468	24
7	C2At4g30580–P14/M47-317.49	30
8	E39/M59-171.21–CT241	25
12	T0849–P15/M47-337.90	14

Significant ($P \leq 0.01$) segregation distortion was detected for 10 % of the markers. The majority (75 %) of these skewed markers fell into three clusters within the genome: the upper half of linkage group (LG) 2, the pericentromeric region of LG3, and the distal portion of the upper arm of LG7. The regions on LG3 and LG7 are skewed toward *S. melongena* homozygotes while LG2 markers skew toward the wild parent, *S. linnaeanum*.

Mapping the eggplant MM195 × MM738 population

Analysis of the genotypic data using CarteBlanche mapping software generated 12 major and six minor linkage groups comprised of more than 850 markers. Comparison with the Wu et al. (2009) and Doganlar et al. (2002a) maps helped to establish the identity of each of eggplant's 12 chromosomes among the 12 major linkage groups. Based on common markers, the six minor linkage groups (LGs) were found to correspond to the distal ends of the short arms of chromosomes 3, 5 and 8, the distal end of the long arm of chromosome 5, and the short arms of chromosomes 7 and 12.

In an effort to merge the major and minor linkage groups, the mapping data were additionally analyzed with the "group" command of MAPMAKER/EXP 3.0 software (Lander et al. 1987). MAPMAKER grouped 99 % of the markers into a total of 12 linkage groups. The marker composition of each of these clusters was consistent with the chromosomal identities of the major and minor linkage groups as established by comparison with the previously published maps. The Kosambi function (Kosambi 1944) as implemented by MAPMAKER was used to estimate the genetic

distances spanning the gaps between the major and minor linkage groups forming chromosomes 3, 5, 7, 8, and 12 (Table 2). These distances ranged from 14 to 30 cM with an average interval of 21 cM.

The linkage group assignments of a fraction (5 %) of the RFLP and COSII markers were found to be inconsistent with the previous eggplant maps and the tomato genomic marker data available from the Sol Genomics Network website (http://solgenomics.net/). Because these anomalous assignments were not supported by neighboring markers, we considered them unreliable and omitted these markers from the map with one notable exception. TG68 which mapped to the centromeric region of LG10 instead of the distal end of the short arm of LG12 was preserved for reasons outlined in the discussion. RFLP and COSII marker order in all linkage groups was colinear with the framework markers (those mapped at LOD >3) in the integrated COSII–RFLP eggplant map (Wu et al. 2009).

Description of the linkage map

Based on the results of CarteBlanche and MAP-MAKER analyses, an integrated genetic map consisting of 12 linkage groups was constructed (Online Resource 1). Ninety-nine percent (400) of the genotyped AFLP markers could be assigned to one of the linkage groups (LGs). In addition, 348 RFLP and 116 COSII markers were localized for a total of 864 DNA markers spanning 1,518 cM. The distribution of different marker types, map length per linkage group, and the occurrence of gaps larger than 10 cM are summarized in Table 3. LG10 is the longest linkage group (152 cM). LG12 is the shortest (93 cM) and also has the fewest markers (43 markers, 5 % of the total). LG11 contains the greatest total number of markers (96 loci, 11 % of the mapped markers) as well as the greatest number of the newly developed markers with 52 AFLP loci, 13 % of the total. However, in terms of marker density, LG2 is the most saturated with an average marker interval of 1.3 cM (0.8 markers per cM) and LG6 is the least saturated averaging 2.6 cM between markers (0.4 markers per cM). LG6 also carries the fewest number of AFLP loci (15 loci). Overall, the average distance between adjacent markers is 1.8 cM, with 0.6 markers per centiMorgan. The newly developed AFLP markers are, on average, 3.3 cM apart from one another;

Table 3 Number of markers mapped to each linkage group, linkage group length, and marker saturation statistics

Linkage group	Number of markers (previous map)				Length (cM)	Average interval (cM)	Marker density (per cM)	Number of gaps >10 cM
	RFLP	cosII	AFLP	Total				
1	31 (26)	9 (9)	42 (0)	82 (35)	124.2 (153.2)	1.5 (4.5)	0.66 (0.23)	0 (3)
2	28 (22)	12 (11)	36 (0)	76 (33)	100.5 (104.8)	1.3 (3.3)	0.76 (0.31)	0 (1)
3	39 (23)	21 (21)	27 (0)	87 (44)	145.7 (159.2)	1.7 (3.7)	0.60 (0.28)	1 (3)
4	34 (22)	4 (5)	32 (0)	70 (27)	121.1 (121.7)	1.8 (4.7)	0.57 (0.22)	0 (4)
5	23 (17)	4 (4)	33 (0)	60 (21)	125 (128.3)	2.1 (6.4)	0.48 (0.16)	2 (4)
6	30 (17)	6 (5)	15 (0)	51 (22)	129 (129.7)	2.6 (6.2)	0.40 (0.17)	1 (5)
7	31 (18)	10 (10)	42 (0)	83 (28)	137.1 (117.1)	1.7 (4.3)	0.61 (0.24)	1 (2)
8	23 (16)	10 (9)	28 (0)	61 (25)	133.1 (128.2)	2.2 (5.3)	0.46 (0.20)	1 (4)
9	29 (22)	7 (5)	40 (0)	76 (27)	116.3 (112.4)	1.6 (4.3)	0.65 (0.24)	1 (2)
10	34 (21)	15 (14)	30 (0)	79 (35)	151.9 (133.1)	2.0 (3.9)	0.52 (0.26)	2 (3)
11	31 (21)	13 (12)	52 (0)	96 (33)	141.1 (131.6)	1.5 (4.1)	0.68 (0.25)	0 (1)
12	15 (12)	5 (5)	23 (0)	43 (17)	93.1 (116.1)	2.2 (7.2)	0.46 (0.15)	1 (5)
Total	348 (237)	116 (110)	400 (0)	864 (347)	1,518.1 (1,535.4)			10 (37)
Average	29 (20)	10 (9)	33.0 (0)	72 (29)	126.5 (128)	1.8 (4.8)	0.57 (0.23)	0.8 (3.1)

Comparisons to the Wu et al. (2009) map appear in parentheses

however the spacing between these markers ranges widely from 0 to 53.3 cM. Overall, 57 % of these markers map within 1 cM of another AFLP. The map contains 10 gaps (an average of 0.8 gaps/linkage group). Gaps are defined as intervals devoid of markers and longer than 10 cM.

Chromosome rearrangements distinguishing eggplant and tomato

Comparison of the eggplant and tomato linkage maps allowed syntenic and rearranged regions to be identified (Online Resource 2). E1 and T1 are syntenic and marker order is conserved along much of the chromosomes. However, a few rearrangements have occurred at the ends. Wu et al. (2009) and Wu and Tanksley (2010) reported an inversion at the top of the chromosome however more comprehensive marker data suggests two apparent translocations in the region. Moreover, the current map reveals the existence of two inversions toward the bottom of the chromosome, only one of which was shown by Wu et al. (2009).

E2 and T2 are also syntenic but differentiated by at least three inversions as described by Wu et al. (2009) and Wu and Tanksley (2010). The uppermost rearrangement is more complex than reported by Wu. An

inverted segment of collinear markers has exchanged positions with an adjacent portion of the chromosome which maintains its orientation. This arrangement can be explained by an initial inversion of the whole region followed by a secondary smaller inversion. Alternatively, translocation of the inverted segment could explain the organization of chromosome 2 at this location.

E3 is syntenic with T3, T5 top, and T5 centromere. E3 and T3 are distinguished by at least two transpositions (of markers TG585 and C2At5g62390) and two inversions. There is a small inverted region at the top of the chromosome and an apparent inverted translocation near the centromere. These differences are consistent with those seen in the COSII map (Wu et al. 2009; Wu and Tanksley 2010) with the exception of the observation of a small inversion at the bottom of the chromosome shown in the COSII map (Wu et al. 2009). The centromeric region of E3 contains two segments syntenic with the top and centromeric regions of T5. The small T5 centromeric segment is nested within the other T5 segment on E3. Wu et al. (2009) described this as an inversion however we believe it is more properly described as a nested translocation.

E4 is syntenic with the bottom of T4 and the top of T10. The upper arm of E4 consists of markers found on

T10. In contrast to the multiple inversions reported previously (Wu et al. 2009; Wu and Tanksley 2010) marker order is highly conserved with the exception of a apparent translocation involving markers mapped to the top of E4 and the centromere of T10. Except for the inclusion of two T11 markers near the top of the region, the lower arm of E4 is syntenic with T4. While there is some indication that small rearrangements in marker order may have occurred, evidence for these events is insufficient given that they involve markers mapped with low (<3) LOD scores in tomato.

E5 is syntenic with T5 and T12. The upper arm of E5 is entirely collinear with the lower arm of T5. The lower arm is largely collinear with the lower arm of T12 except for an apparent transposition of marker T0814. These results are consistent with those seen in the COSII map (Wu et al. 2009; Wu and Tanksley 2010).

E6 and T6 are syntenic. Three inversions differentiate the chromosomes, one of these has not been previously described. A single marker (C2At2g39690) transposition was also detected.

E7 is syntenic with T7. Two small inversions are observed, one at the very top of the chromosome and one below the centromere. Wu et al. (2009) cited insufficient evidence of the uppermost inversion first reported by Doganlar et al. (2002a), however, the current map supports its presence. Moreover, a new translocation was identified in the centromeric region. As in the COSII map, a single T12 marker (C2At5g48300) mapped to E7.

E8 and T8 are syntenic and largely collinear. As reported by Wu et al. (2009) E8 carries two markers found on other tomato chromosomes: C2At1g10580 from T2 and C2At5g41270 from T4. A small inversion at the centromere was newly detected.

E9 and T9 are syntenic and distinguished by a large inverted region as reported previously (Doganlar et al. 2002a; Wu et al. 2009; Wu and Tanksley 2010). Because this region has also shifted in position, we have classified it as an inverted translocation. In addition, GP39 has undergone a transposition.

E10 is syntenic with T10, T5 and T12. The upper portion of E10 consists of markers found on the upper arm of T5. Marker order is conserved in this region. Below this region of E10, T12 markers are seen in an inverted orientation as previously reported (Wu et al. 2009; Wu and Tanksley 2010). Another region of T5 syntenic markers is found below the centromere of

E10. The positioning of these markers suggests that two separate rearrangements occurred: both inverted translocations. The end of E10 is syntenic with the lower arm of T10 but distinguished by a large inversion encompassing most of the region as described in the previous maps (Doganlar et al. 2002a; Wu et al. 2009; Wu and Tanksley 2010). One marker on E10 (C2At1g64770) maps to T3.

E11 is syntenic with T11 and T4. The upper arm of E11 is entirely comprised of markers found on T4. Within this region, an inverted translocated region and a simple inversion are apparent as compared to the three inversions on the COSII map (Wu et al. 2009; Wu and Tanksley 2010). The bottom arm of E11 consists of markers found on the top of T11. Marker order is conserved as previously reported (Doganlar et al. 2002a; Wu et al. 2009; Wu and Tanksley 2010).

E12 is syntenic with T12 and T11. The top of E12 contains markers from a portion of the upper arm of T12. Within this region, a small inversion is apparent as seen in the COSII map (Wu et al. 2009; Wu and Tanksley 2010). The rest of E12 is syntenic with the lower arm of T11. While the existence of a large inversion at the bottom of E12 was verified, this region was also revealed to contain a translocation. We could not confirm the second inversion seen in the COSII map (Wu et al. 2009; Wu and Tanksley 2010).

Discussion

AFLP map development in eggplant has been somewhat limited heretofore. Sunseri et al. (2003) identified 210 AFLP markers that segregated within a *S. linnaeanum* × *S. melongena* 'Buia' F_2 population. A linkage map was constructed consisting of 74 % of these markers (156 AFLP markers) as well as 117 RAPD markers. Faced with the challenge of uncovering a sufficient number of polymorphic markers in an intraspecific F_2 mapping population generated from eggplant breeding lines ('305E40' × '67/3'), Barchi et al. (2010) also adopted the AFLP strategy. The resulting map included more than 200 AFLP markers; however extensive marker clustering limited its utility. The map has since been supplanted by one based on RAD-derived SNPs (Barchi et al. 2012). In our study, 404 markers were scored and 400 (99 %) of these were positioned on the linkage map (Online Resource 1). The high yield of successfully mapped AFLP markers

in our study is no doubt due to advances in AFLP technology that made it possible to score these markers in a co-dominant fashion.

Segregation distortion was detected for just three regions of the genome, the upper portion of LG2, the pericentromeric region of LG3 and the distal end of the upper arm of LG7. The skew in favor of *S. melongena* alleles on LGS 3 and 7 and *S. linnaeanum* alleles on LG2 is in accordance with what was previously reported for this population (Doganlar et al. 2002a). As compared to the previous map constructed from this population (Doganlar et al. 2002a), a smaller proportion of the mapped markers (10 vs. 16 %) shows skewed segregation. However, Barchi et al. (2012) reported an even lower incidence of segregation distortion (4 %) for their intraspecific mapping population. Segregation distortion is not uncommon in interspecific populations and might be a consequence of selection for structural or genetic factors influencing viability of spores, gametophytes, and/or young sporophytes (Zamir and Tadmor 1986).

The primary goal of this research was to improve the level of marker saturation in the *S. linnaeanum* MM195 × *S. melongena* MM738 interspecific eggplant map. Comparison with the previously updated map for this population (Wu et al. 2009) shows a 2- to 3-fold increase in the number of markers on each linkage group and a 1 % reduction in the overall map length (1,518 vs. 1,535 cM) (Table 3). The majority of the newly positioned markers are AFLP markers (400 markers) specific to eggplant, however, a substantial number of tomato-derived RFLP markers (192) were also localized for the first time in the interspecific mapping population. The average interval between markers in the Wu et al. (2009) map was 4.8 cM (6.1 cM if only framework markers ordered at LOD >3 are considered). In the current map, the average interval distance has fallen to 1.8 cM. Thus, a more than two-fold improvement in marker density (from 0.23 to 0.57 markers per cM) is seen.

The improved saturation of the interspecific map is especially obvious when gaps (genomic regions free of markers) are considered. While the Wu et al. (2009) map contained 37 gaps longer than 10 cM, with nine of these gaps longer than 20 cM. The number of >10 cM gaps is reduced here to ten on the integrated map, with only three extending for more than 20 cM. Six (60 %) of the gaps larger than 10 cM are associated with the mergers on linkage groups 3, 5,

7, 8, and 12. Distance estimates between the marker loci flanking these junctions reveal these to be the largest gaps (averaging 21 cM long) in the current map (Table 2). The map of Wu et al. (2009) also shows large gaps and a paucity of markers in these locations of the genome. The total distance between shared markers that flank the large gaps on LGs 3, 5, 7, 8, and 12 accounts for 9 % of the overall length of the Wu et al. map and 10 % of the current map (data not shown). Thus, the extent of the genome encompassed by the gaps is similar. Such gaps could indicate particularly high rates of recombination in those regions. That recombination rates can vary along the length of a chromosome is supported by physical mapping analysis. Using fluorescence in situ hybridization, Koo et al. (2008) determined that, of four gaps (≥10 cM long) on the genetic linkage map of tomato chromosome 2, only one (24 cM in length) represented a real physical gap (∼3.93 Mb).

It is not surprising that the newly developed eggplant AFLP markers show a tendency to cluster. Such a trend has been previously observed in eggplant (Barchi et al. 2010) and other plant species (Alonso-Blanco et al. 1998; Qi et al. 1998; Pradhan et al. 2003). Although the average genetic distance between the AFLP markers in the current map is 3.3 cM, the level of variability in AFLP marker spacing is considerable. Sixteen percent of the AFLPs co-segregate with another AFLP marker and an additional 41 % of the AFLP markers fall within 1 cM of another AFLP marker (data not shown). Examination of the linkage map reveals some obvious areas of marker clustering, particularly around centromeres (Online Resource 1). Such clustering of markers is indicative of suppressed recombination, an established characteristic of heterochromatic regions in *Solanum* (Tanksley et al. 1992). In addition, DNA sequence divergence between the two parental species would tend to reduce recombination rates. Because our map integrates several marker types, genome coverage does not significantly suffer from marker clustering.

The current genetic map was aligned with the molecular linkage maps previously derived from the *S. linnaeanum* Hepper & Jaeger 'MM195' × *S. melongena* L. 'MM738' F$_2$ population (Doganlar et al. 2002a; Wu et al. 2009) to check its reliability. In addition, the placement of newly mapped RFLP and COSII loci was cross-checked against the known positions of these markers in the tomato genome [as

Table 4 Summary of chromosomal rearrangements distinguishing the linkage groups of eggplant and tomato

Chromosome	Number of rearrangements relative to tomato			
	Inversions	Translocations	Inverted translocations	Transpositions
1	2	2 (0)	0	0
2	2	0	1 (0)	0
3	1	2 (2)	1 (0)	2 (0)
4	0	2 (1)	0	2 (2)
5	0	1 (1)	0	1 (0)
6	3	0	0	1 (0)
7	2	1 (0)	0	1 (1)
8	1	0	0	2 (2)
9	0	0	1 (0)	1 (0)
10	1	2 (2)	2 (0)	1 (1)
11	1	1 (1)	1 (0)	0
12	1	2 (1)	0	0
Total	14	13 (8)	6 (0)	11 (6)

Numbers in parentheses indicate interchromosomal events

obtained from the Sol Genomics Network website (http://solgenomics.net/)]. Unexpected linkage group assignments were observed for 27 solitary markers scattered throughout the genome. The lack of evidence supporting these anomalous positions (in the form of corresponding positions for neighboring markers) led us to eliminate all but one of these markers from the map. In the one discrepancy that was preserved, TG68 mapped adjacent to CT99 on LG10 instead of the extreme distal end of LG12 (see Wu et al. 2009 map). Comparing the marker composition of eggplant chromosome 10 (E10) to the map of the tomato genome reveals that this alternative position for TG68 is plausible given the number of translocations that distinguish the eggplant and tomato genomes. The segment of E10 encompassing CT99 is homologous to tomato chromosome 12 (T12), thus our data suggest that one of the translocation events on E10 may have involved a larger segment of T12 than previously detected. In all other instances, the current map supports the syntenic relationships between the eggplant and tomato genomes established by Doganlar et al. (2002a) and confirmed by Wu et al. (2009).

The high-resolution map developed in this study gives additional insight into chromosome evolution in the Solanaceae. Previous mapping of the interspecific eggplant population revealed that at least five translocations and 24 inversions differentiate the linkage maps of eggplant and tomato (Doganlar et al. 2002a; Wu et al. 2009). In an analysis of chromosome evolution across the Solanaceae, Wu and Tanksley

(2010) deduced that a minimum of 16 inversions and three translocations have occurred in the eggplant genome since its divergence from tomato and potato ~15.5 million years ago. They also report a total of 20 inversions and at least three translocations distinguishing tomato and eggplant. Our analysis reveals a slightly different picture: 33 rearrangements were identified between the eggplant and tomato linkage maps (Table 4; Online Resource 2). Of these differences, only fourteen are now classified as inversions. Thus, a greater proportion of translocation events have occurred than previously thought: 13 simple translocations and six inverted translocations. Eight of the 19 translocations were interchromosomal. Only two chromosomes, 6 and 8, showed no evidence of translocation events. In addition, eleven transpositions of single markers were detected with six of these being interchromosomal.

In summary, we have constructed a high-density linkage map of the eggplant genome that integrates AFLP markers specific to eggplant with RFLP markers derived from tomato as well as COSII markers widely conserved in dicots. The map comprises 864 markers spread across eggplant's twelve chromosomes at an average genetic interval of 1.8 cM and represents a considerable improvement of the maps previously developed for this interspecific population (Doganlar et al. 2002a; Wu et al. 2009). While an integrated high density map of the eggplant genome has recently been produced from two intraspecific populations (Fukuoka et al. 2012), the interspecific

population used here offers a greater degree of morphological polymorphism that can form the basis for analyzing a wider range of quantitative traits in eggplant. The recent release of the tomato genome sequence (Tomato Genome Consortium 2012) provides a springboard for comparative genomics within the Solanaceae. The fact that more than half of the markers on the current map have also been localized in the tomato genome will provide greater opportunities for leveraging tomato genomic resources for the betterment of eggplant genetics.

Conflict of interest The authors declare that they have no conflict of interest.

References

Alonso-Blanco C, Peeters AJ, Koornneef M, Lister C, Dean C, van den Bosch N, Pot J, Kuiper MT (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. Plant J 14:259–271

Barchi L, Lanteri S, Portis E, Stagel A, Vale 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, Vale G, Volante A, Pulcini L, Ciriaci T, Acciarri N, Barbierato V, Toppino L, Rotino GL (2012) A RAD tag derived marker based eggplant linkage map and location of QTLs determining anthocyanin pigmentation. PLoS ONE 7:e43740

Bernatzky R, Tanksley S (1986) Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics 112:887–898

Buetow KH, Chakravarti A (1987) Multipoint gene mapping using seriation I general methods. Am J Hum Genet 41:180–188

Buntjer J (2002) Carte Blanche: software for high-density linkage mapping. Plant Animal & Microbe Genomes Conference X. San Diego, CA USA

Buntjer J, Van Der Meulen M, Witsenboer H, Turk S, Van Schaik R (2002) Construction of an integrated map of tomato using novel high-density mapping software Carte Blanche. Plant Animal & Microbe Genomes Conference X. San Diego, CA USA

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

FAO Statistics (2012) http://faostatfao.org. Accessed 15 Oct 2012

Frary A, Doganlar S, Daunay M-C, Tanksley SD (2003) QTL analysis of morphological traits in eggplant and implications for conservation of gene function during evolution of Solanaceous species. Theor Appl Genet 107:359–370

Fukuoka H, Miyatake K, Nunome T, Negoro S, Shirasawa K, Isobe S, Asamizu E, Yamaguchi H, Ohyama A (2012) Development of gene-based markers and construction of an integrated linkage map in eggplant by using Solanum orthologous (SOL) gene sets. Theor Appl Genet 125:47–56

Jansen R, Geerlings H, Van Oeveren AJ, Van Schaik RC (2001) A comment on codominant scoring of AFLP markers. Genetics 158:925–926

Koo D-K, Jo S-H, Bang J-W, Park H-M, Lee S, Choi D (2008) Integration of cytogenetic and genetic linkage maps unveils the physical architecture of tomato chromosome 2. Genetics 179:1211–1220

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newberg LA (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Livingstone KD, Lackney VK, Blauth JR, van Wijk R, Jahn MK (1999) Genome mapping in *Capsicum* and the evolution of genome structure in the Solanaceae. Genetics 152:1183–1202

Mank R, Van Haaren M (2003) CarteBlanche: novel high-density mapping software. Plant & Animal Genomes Conference XI. San Diego, CA USA

Nunome T, Yoshida T, Hirai M (1998) Genetic linkage map of eggplant. In: Proceedings of the 10th eucarpia meeting on genetics and breeding of *Capsicum* and eggplant. Avignon, pp 239–242

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. Breed Sci 51:19–26

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

Nunome T, Negoro S, Kono I, Kanamori H, Miyataki K, Yamaguchi H, Ohyama A, Fukuoka H (2009) Development of SSR markers derived from SSR-enriched genomic library of eggplant (*Solanum melongena* L.). Theor Appl Genet 119:1143–1153

Pradhan AK, Gupta V, Mukhopadhyay A, Arumugam N, Sodhi YS, Pental D (2003) A high-density linkage map in *Brassica juncea* (Indian mustard) using AFLP and RFLP markers. Theor Appl Genet 106:607–614

Qi X, Stam P, Lindhout P (1998) Use of locus-specific AFLP markers to construct a high-density linkage map in barley. Theor Appl Genet 96:376–384

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. In: Hammerschlag FA Saxena P [eds] XXVI international horticultural congress: biotechnology in horticultural crop improvement. Acta Horticulturae 625:107–115

Tanksley SD, Ganal MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Roder S, Wing RA, Wu W, Young ND (1992) High-density linkage maps of the tomato and potato genomes. Genetics 132:1141–1160

Thompson EA (1987) Crossover counts and likelihood in multipoint linkage analysis. IMA J Math Appl Med Biol 4:93–108

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

Truong HT, Ramos M, Yalcin F, De Ruiter M, Van Der Poel HJA, Huvenaars KHJ, Hogers RCJ, Van Enckevort LJG, Janssen A, Van Orsouw NJ, Van Eijk MJT (2012) Sequence-based genotyping for marker discovery and co-dominant scoring in germplasm and populations. PLoS ONE 7:e37565

Van Eijk M, Broekhof J, Van Der Poel H, Hogers R, Schneiders H, Kamerbeek J, Verstege E, Van Aart J, Geerlings H, Buntjer J, Van Oeveren J, Vos P (2004) SNPWave™: a flexible multiplexed SNP genotyping technology. Nucleic Acids Res 32:e47

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77–78

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Friters A, Pot J, Paleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Watowich SJ, Meyer ES, Hagstrom R, Josephs R (1988) A stable rapidly converging conjugate gradient method for energy minimization. J Comput Chem 9:650–661

Wu F, Tanksley SD (2010) Chromosomal evolution in the plant family Solanaceae. BMC Genom 11:182–193

Wu FN, Mueller LA, Crouzillat D, Petiard V, Tanksley SD (2006) Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative evolutionary an systematic studies: a test case in the euasterid plant clade. Genetics 174:1407–1420

Wu F, Eannetta N, Xu Y, Tanksley S (2009) A detailed synteny map of the eggplant genome based on conserved ortholog set II (COSII) markers. Theor Appl Genet 118:927–935

Zamir D, Tadmor Y (1986) Unequal segregation of nuclear genes in plants. Bot Gaz 147:355–358