the plant journal

SEB Society for Experimental Biology

*The Plant Journal* (2012) **71**, 602–614 doi: 10.1111/j.1365-313X.2012.05012.x

# Structural homology in the Solanaceae: analysis of genomic regions in support of synteny studies in tomato, potato and pepper

**Sander A. Peters[1,†,*], Joachim W. Bargsten[1,2,3,†], Dóra Szinay[4], José van de Belt[4], Richard G.F. Visser[2,5], Yuling Bai[2] and Hans de Jong[4]**

[1]*Plant Research International, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands,*

[2]*Laboratory of Plant Breeding, Wageningen University and Research Centre, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands,*

[3]*Netherlands Bioinformatics Centre, Nijmegen, The Netherlands,*

[4]*Laboratory of Genetics, Wageningen University and Research Centre, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands, and*

[5]*Department of Plant Sciences, Wageningen University and Research Centre, 6700 AJ Wageningen, The Netherlands*

Received 12 January 2012; revised 16 February 2012; accepted 26 March 2012; published online 12 June 2012.
*For correspondence (e-mail sander.peters@wur.nl).
†These authors contributed equally to this paper.

## SUMMARY

**We have analysed the structural homology in euchromatin regions of tomato, potato and pepper with special attention for the long arm of chromosome 2 (2L). Molecular organization and colinear junctions were delineated using multi-color BAC FISH analysis and comparative sequence alignment. We found large-scale rearrangements including inversions and segmental translocations that were not reported in previous comparative studies. Some of the structural rearrangements are specific for the tomato clade, and differentiate tomato from potato, pepper and other Solanaceous species. Although local gene vicinity is largely preserved, there are many small-scale synteny perturbations. Gene adjacency in the aligned segments was frequently disrupted for 47% of the ortholog pairs as a result of gene and LTR retrotransposon insertions, and occasionally by single gene inversions and translocations. Our data also suggests that long distance intra-chromosomal rearrangements and local gene rearrangements have evolved frequently during speciation in the *Solanum* genus, and that small changes are more prevalent than large-scale differences. The occurrence of sonata and harbinger transposable elements and other repeats near or at junction breaks is considered in the light of repeat-mediated rearrangements and a reconstruction scenario for an ancestral 2L topology is discussed.**

**Keywords: comparative sequence analysis, fluorescent *in situ* hybridization (FISH), chromosome rearrangement, orthology, transposon, chromosome evolution.**

## INTRODUCTION

The Solanaceae or nightshade family is a large group of more than 3000 species that includes tuber or fruit-bearing vegetables [tomato (*Solanum lycopersicum*), potato (*Solanum tuberosum*), pepper (*Capsicum annuum*) and eggplant/aubergine (*Solanum melongena*)], and plants of horticultural (petunia, *Petunia hybrida*) and medicinal (tobacco, *Nicotiana tabacum*) importance (Knapp, 2002; Sesso *et al.*, 2003). The family is economically the third most important, and ranks number one in terms of vegetable crops (Foolad, 2007). The Solanaceae show remarkable adaptability to diverse climatic conditions, ranging from wet rainforests to dry and arid environments, and exhibit a huge phenotypic diversity from tiny annual herbs to large forest trees (Knapp, 2002). In contrast, cultivated *Solanum* and *Capsicum* crops have a strikingly narrowed genetic basis through domestication, resulting in loss of desirable traits, including those that confer (a)biotic stress tolerance. In addition, rapidly changing climate conditions and increasing competing claims for arable lands will increase the demand for new varieties that tolerate harsh environmental conditions,


© 2012 The Authors
The Plant Journal © 2012 Blackwell Publishing Ltd

confer resistance against pathogens and at the same time have better productivity and nutritional quality. The need to compensate for such genetic losses requires introgression of alien chromatin from wild relatives to the crops, a process referred to as introgressive hybridization. In general, the wild relatives of solanaceous crops provide a gene pool that is sufficiently rich for crop improvement. However, their use as donor in introgressive breeding is limited (Rick *et al.*, 1987; Singh, 2007; Bai and Lindhout, 2007). Crossing barriers and linkage drag are well-known phenomena that limit the use of germplasm for introgressive hydridization (Rieseberg and Willis, 2007; Bedinger *et al.*, 2010).

The transfer of alien chromatin containing the genetic information for a desirable trait depends on homeologous recombination between the donor chromosome and its corresponding counterpart in the crop, which, amongst other factors, is determined by their level of colinearity. Severe problems may occur in those cases where the donor chromosome and its homeolog differ as a result of large-scale rearrangements (inversions or translocations). Heterozygosity for such rearrangements may lead to failure of synapsis and/or illegitimate crossovers at meiosis. As a consequence, genes are unlikely to recombine, and so are transmitted as a single locus, a phenomenon known as linkage drag. To facilitate identification of compatible donor species or genotypes for error-free homeologous introgression of important agronomic traits such as (a)biotic stress tolerance, elucidating the genome organization is imperative. This is usually accomplished by analysing colinearity, synteny and linkage at both the chromosome and gene level.

The synteny concept was introduced in 1971, and pertains to the preserved co-location of homologous genes on chromosomes between species, irrespective of genetic linkage and gene order (Ehrlich *et al.*, 1997; Passarge *et al.*, 1999). Conservation of both synteny and order of homologs determine conserved linkage of genes. Both synteny and conserved linkage have been used to investigate solanaceous genome organization, gene diversification and evolutionary ancestry (Ku *et al.*, 2000; Fulton *et al.*, 2002; Wu *et al.*, 2006; Wang *et al.*, 2008). The identification of ancestral relationships between homologous genes and their distinction into orthologs and paralogs has become more feasible with the advent of high-throughput sequencing that facilitates comparisons of entire genomes. Until now, sequence comparisons for conserved syntenic segments in *Solanum* have mainly been obtained for relatively small orthologous regions. In general, the order and sequence of orthologs was found to be conserved, despite a few small-scale differences and positive gene selections (Doganlar *et al.*, 2002a; Wang *et al.*, 2008).

Analysis of chromosome structure in Solanaceae has been based on several lines of research. For example, light microscopy observations on pachytene chromosomes of $F_1$ hybrids showed normal synapsis along the chromosomes. Furthermore, linkage maps of intra- and inter-specific hybrids were found to be largely colinear (Pertuzé *et al.*, 2002; Chetelat and Ji, 2007; Moyle, 2008). Both similarity in chromosome morphology and marker colinearity supported the notion that *Solanum* species have evolved primarily by genic change rather than by large-scale chromosomal rearrangements. Nonetheless, genetic linkage analyses indicated that tomato and potato are differentiated by a series of whole-arm inversions of chromosomes 5, 9, 10, 11 and 12 (Bonierbale *et al.*, 1988; Tanksley *et al.*, 1988; Livingstone *et al.*, 1999; Doganlar *et al.*, 2002b; Pertuzé *et al.*, 2002). Furthermore, electron microscopy studies on somatic hybrids of tomato and potato (De Jong *et al.*, 1993) and $F_1$ hybrids from interspecific crosses revealed substantial changes in chromosome structure among *Solanum* species (Anderson *et al.*, 2010). Most of these structural changes, however, were found in the heterochromatin, with comparatively few genes and low recombination, and thus would have little effect on the colinearity of linkage maps (Anderson *et al.*, 2010).

Although genetic mapping studies provided valuable starting points for unraveling plant genome organization, they are inaccurate in regions where crossover recombination is suppressed or even absent, as in the distal heterochromatin and the large pericentromere regions, for example. Furthermore, there are insufficient DNA polymorphisms for simple markers that are locus-specific across species, causing markerless gaps in linkage maps that could leave chromosome rearrangements undetected. Such problems were recognized within the framework of the tomato genome sequencing project (Szinay *et al.*, 2008; Peters *et al.*, 2009). For example, integrated mapping revealed genetic intervals comprising hundreds of tomato genes in euchromatic regions with a marker coverage insufficient to support microsynteny analysis. Some studies used a sequence-based comparison for microsyntenic analysis in Solanaceae (Fulton *et al.*, 2002; Van der Hoeven *et al.*, 2002; Datema *et al.*, 2008; Wang *et al.*, 2008), but genome-wide comparative sequence analysis is still limited for *Solanum*, as comparable sequence clades have not yet emerged. Alternatively, genome-wide cross-species fluorescence *in situ* hybridization (FISH) has provided a foundation for comprehensive comparative maps, as well as rapid and reliable detection of genetic elements that are associated with traits of interest. For example, FISH has been used to analyse the organization of the short arm of chromosome 6 (6S) in tomato and potato, which contains the *Mi* resistance homolog cluster. An additional inversion was revealed on 6S (Iovene *et al.*, 2008; Tang *et al.*, 2008) that was not reported in synteny studies with molecular markers (Tanksley *et al.*, 1992; Grube *et al.*, 2000), probably due to lack of marker coverage and suppression of recombination (Liharska *et al.*, 1996; Bai *et al.*, 2004; Seah *et al.*, 2004).

Recently, the complete genome sequences of both tomato and potato have become available, and this allowed us to

© 2012 The Authors

take advantage of a combined cytogenetic-based macrosyntenic approach and a comparative sequence-based microsyntenic approach. Here, we address a number of issues including content and gene organization in *Solanum* chromosomes. In addition, we present a detailed analysis of chromosome rearrangements in tomato (*S. lycopersicum*), potato (*S. tuberosum*) and pepper (*Capsicum annuum*), with special attention to regions containing stress tolerance and disease resistance homologs, and discuss the rearrangements in the light of chromosome evolution.

## RESULTS

### Cytogenetic macrosynteny between tomato, potato and pepper

Nucleotide sequences and translations from annotated coding regions of 32 genes that have been implicated in stress tolerance (Jenks *et al.*, 2007) were collected from the Genbank and Refseq sequence databases (http://www.ncbi.nlm.nih.gov), and were used for BLASTN- or TBLASTX-based similarity searches. We identified 19 stress tolerance gene homologs in 19 tomato BACs, and verified the cytogenetic mapping position on tomato and potato chromosome 2, 6, 10 and 11 pachytene complements using BAC FISH. Eight BACs, of which six contain a stress tolerance homolog, display a single clear fluorescent signal on chromosome 2. This chromosome can easily be distinguished from the other chromosomes by its acocentric structure and the large nucleolar organizing region. Previously, cytogenetic studies of tomato pachytene chromosomes revealed long continuous stretches of less condensed euchromatin in both chromosome arms, flanked by highly condensed heterochromatin at the telomere ends and the centromeres (Ganal *et al.*, 1991; De Jong *et al.*, 2000; Chang *et al.*, 2008). Based on these morphological properties, the foci observed were in the euchromatic part on the long arm of tomato chromosome 2 (2L) (Figure 1). Genetically, these BACs have been mapped at an interval between 82 and 143 cM based on BLASTN hits to SGN EXPEN2000 markers (http://solgenomics.net/cview/map.pl?map_version_id=52). The genetic map order of all BACs is consistent with the linear cytogenetic mapping order, except for H146O6, which appears to be inverted (Figure 1 and S1). Using common markers, we mapped the tomato BACs to the euchromatic portion of potato 2L (see below) in an interval between 33 and 50 cM on the potato TXB genetic map (Koo *et al.*, 2008). Strikingly, the cytogenetic mapping order on potato and tomato pachytene is clearly different, suggesting multiple rearranged segments. On potato and pepper chromosome 2L, the BAC FISH mapping order appears colinear (Figure 1).

In addition to 2L, cross-species FISH analysis previously revealed major differences in chromosome organization for 6S, including the *Mi* disease resistance homolog cluster. Mapping results in the 1–10 cM interval were interpreted as a large paracentric inversion that covers the 4.5 Mb euchromatin part with breakpoints close to the top arm telomere and at the border of the pericentromeric heterochromatin. The observations also suggested rearrangements in the middle part of the short arm. However, the precise differences in chromosomal organization were not resolved at the time (Tang *et al.*, 2008; Peters *et al.*, 2009). Here we aim to further elucidate the topology of 6S. A distal inversion was detected in pepper and potato 6S compared to tomato. In addition, one BAC displayed two foci on pepper 6S, which could be due to a duplicated segment or a breakpoint in pepper, or a deletion in potato (Figure 1).

Differences in the mapping order of markers on tomato, potato and pepper genetic maps also indicate rearrangements near the *TSW* and *L* loci. These loci reside on the long arm of chromosomes 10 (10L) and 11 (11L), respectively, and have been implicated in tomato spotted wilt virus and tobacco mosaic virus resistance in *Capsicum* spp. (Jahn *et al.*, 2000; Yang *et al.*, 2009). We selected three anchored tomato BACs in the 52–62 cM interval and the 74–90 cM interval on the EXPEN2000 genetic map for FISH analysis (Figures S2 and S3). These tomato BACs each display a focus in the euchromatic portion of tomato, potato and pepper pachytene chromosomes, and the order of foci on potato and pepper is colinear. However, the order of foci is inverted on tomato 10L, indicating rearrangements in this part of the tomato genome. On tomato and potato 11L the order is colinear, but cytogenetic mapping indicates a translocated segment with a reversed orientation in 11L of pepper (Figure 1), in agreement with the comparative mapping in the 57–115 cM interval (Yang *et al.*, 2009) (Figure S3).

### Comparative sequence alignment

To further delineate differences in the chromosomal organization, we selected tomato scaffold sequences based on BLASTN hits using tomato BAC ends and genetic marker sequences (http://solgenomics.net). On 2L, the 82–143 cM interval corresponds to a sequence of approximately 7 Mb from tomato scaffold SL2.40sc03665 (http://solgenomics.net/sequencing/agp.pl), which spans H159F19 and H072B02 on the FISH map (Figure 1). Tomato BAC ends and markers anchoring BACs to the EXPEN2000 map show BLASTN matches to five potato scaffolds (DM244, DM441, DM1213, DM12 and DM4) from *Solanum tuberosum* group Phureja DM1-3 516 R44 (hereafter referred to as *S. tuberosum* DM), which have a total length of approximately 7 Mb. Both DM scaffold order and orientation on the potato 2L FISH map are consistent with the genetic marker order, and the 'accessioned golden path' (AGP) map for potato chromosome 2 (Potato Genome Sequencing Consortium, 2011). Subsequently, a comparative alignment revealed the coordinates of the junction breaks of translocated segments for scaffold DM4, DM12 and DM244. In addition, two segments of approximately 3 Mb and 980 kb originating from scaffold

© 2012 The Authors

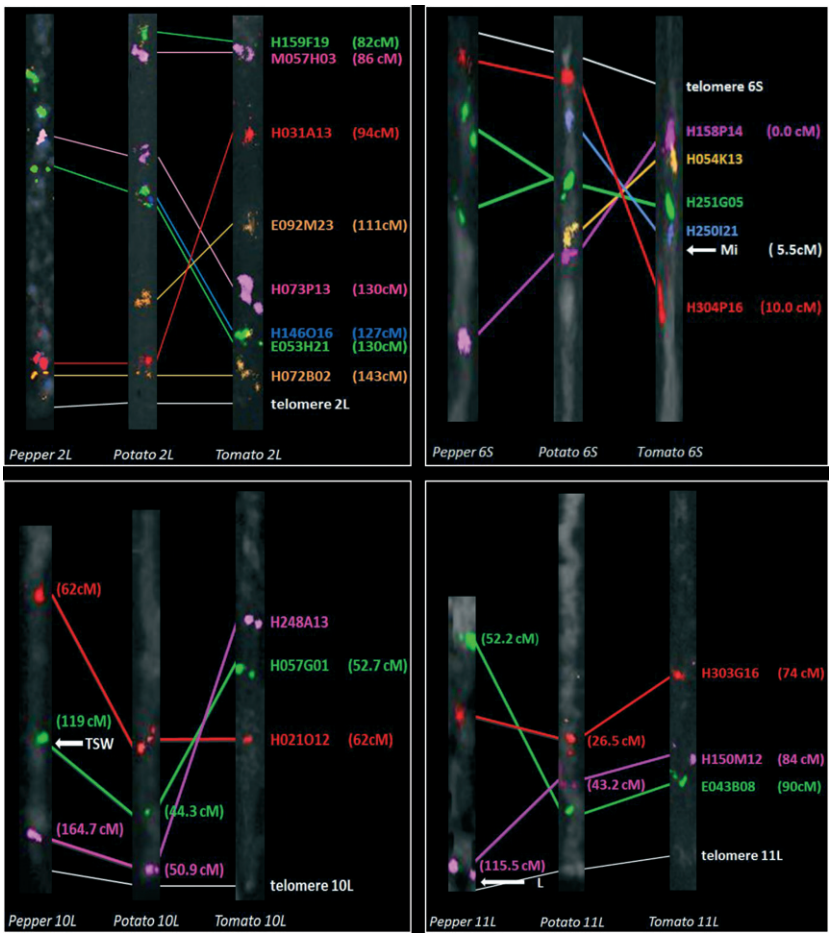



**Figure 1.** Comparative FISH analysis of tomato BACs on tomato, potato and pepper 2L, 6S, 10L and 11L. Identifiers for tomato BACs are indicated on the right of each pachytene chromosome. White arrows indicate the positions of *Mi*, *L* and *TSW* loci. Corresponding cytogenetic mapping positions of *Hin*dIII (H), *Eco*RI (E) and *Mbo*I (M) tomato BACs, and telomere positions on tomato, potato and pepper pachytenes are shown by coloured connecting lines and white connecting lines, respectively. The positions of SGN markers from the tomato EXPEN2000, potato TXB1992 and pepper AC99 genetic maps (http://solgenomics.net/cview/map.pl?map_id=11) associated with the tomato BACs are indicated (cM).

DM4, two DM12 segments of 300 and 690 kb, and a small 20 kb segment of DM1213 align in the opposite orientation (Figure 2), which suggests inversions have occurred.

The identity plot also displays multiple rearrangements between corresponding segments of tomato and potato 6S, 10L and 11L (Figure 2b–d and Table S1). We observed large inversions for practically the entire 4.5 Mb of tomato 6S, with several segments that appear translocated compared to potato 6S. In addition, the first 0.5 Mb just downstream of the tomato 6S telomere position appears almost entirely missing in potato, leaving only a small colinear fragment of 60 kb. For a 2.9 Mb segment from tomato 10L, we found two colinear fragments of 375 kb and 1.35 Mb. Two fragments of approximately 470 and 710 kb aligned in opposite orientation to fragments from a 7.5 Mb potato segment, of which 4.14 Mb did not align. A comparative alignment of a 1.9 Mb tomato 11L segment to a corresponding 2.5 Mb potato segment showed three inversions of 144, 295 and 190 kb, and deletions of approximately 230, 140 and 90 kb. Although cytogenetic mapping did not reveal the inversions in tomato 11L, the FISH mapping order is explained by the BLASTN hit positions (Figure 2d). Because genomic sequences are currently lacking for pepper, we cannot confirm the structural rearrangements between tomato and pepper 6S, 10L and 11L by sequence comparison.

### Topology of tomato and potato 2L segments

To exclude the possibility that the observed colinearity breaks in the comparative sequence alignment arose from aberrantly assembled sequences, we validated the borders of colinear scaffold segments by comparative FISH analysis using BACs that span a junction break. The FISH mapping shows single clear foci for the junction break BACs on tomato 2L (Figure 3). The mapping order is in agreement with the order and BLAST position of BACs on scaffold 3665, and thus confirms correct assembly of the 7 Mb region. Remarkably, several tomato BACs that span alignment breaks displayed two FISH signals on 2L of both potato accessions G254 and RH89039 (Figure 3, lanes 1, 3, 5 and 7). In addition, the FISH map position of H138J12 is just north of E129C17 (Figure 3, lanes 3 and 4), in agreement with the inverted orientation of a 300 kb DM12 segment. H028F18 co-localizes with H015P22 on potato 2L from accession RH8903916 (Figure 3, lane 3), in contrast to two distinct foci on tomato (Figure 3, lane 4). This is in agreement with two overlapping BLASTN hits on DM4 and two hits on scaffold

© 2012 The Authors

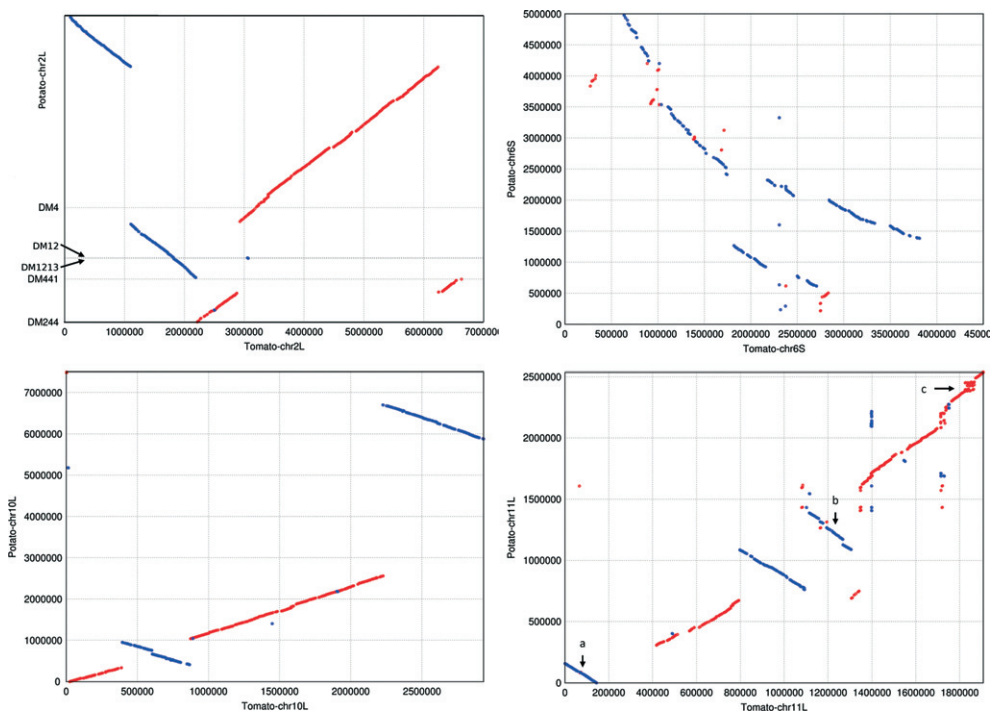



**Figure 2.** Identity plots of *S. lycopersicum* and *S. tuberosum* group Phureja DM1-3 516 R44 segments of 2L, 6S, 10L and 11L. Sequences aligned in forward and reversed orientation are represented by red and blue lines, respectively. Tomato and potato chromosome labels are indicated on the *x* and *y* axes. Segment positions aligning to BAC H303G16, H150M12 and E043B08 are marked in the bottom right plot by arrows (a) to (c), respectively.

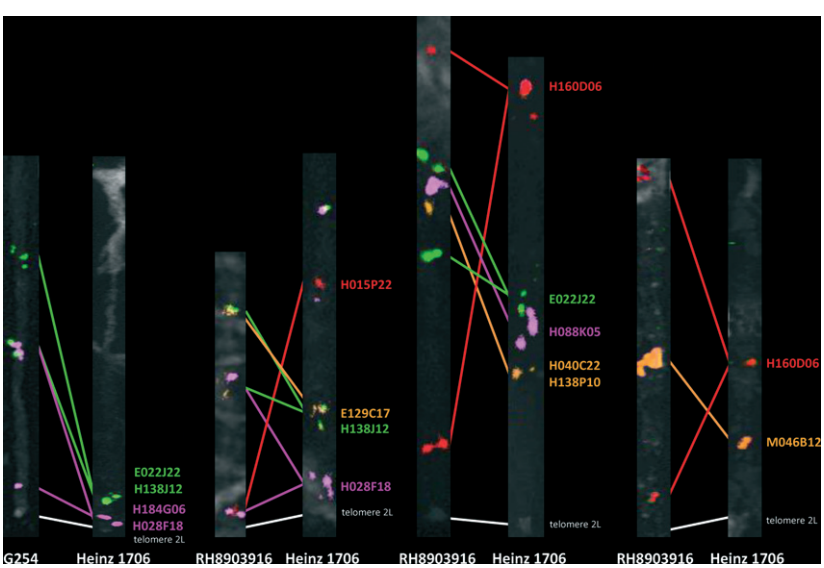



**Figure 3.** FISH analysis of colinearity breaks between tomato and potato 2L. Tomato BACs were hybridized on pachytene complements from tomato Heinz 1706 (lanes 2, 4, 6 and 8) and pachytenes from potato G254 (lane 1) and RH8903916 (lanes 3, 5 and 7). Corresponding cytogenetic mapping positions of *Hindl*II (H), *EcoRI* (E) and *MboI* (M) tomato BACs on tomato and potato pachytenes are shown by coloured connecting lines. Telomere positions on the 2L are indicated by white connecting lines.

SL2.40sc03665, which map several megabases apart. Furthermore, M046B12 and E129C17 both show a single focus on tomato and potato (Figure 3, lanes 3, 4, 7 and 8), and both H015P22 and H028F18 co-localize on potato pachytene (Figure 3, lane 3). Apparently, a large 3 Mb segment from DM4 has been translocated and has a reversed orientation in tomato. H088K05 and H040C22/H138P10 each display a single focus in a similar order on both tomato and potato pachytene (Figure 3, lanes 5 and 6). This confirms the orientation and position of DM441 and a translocated 690 kb segment from DM12. The two signals for H160D06 (Figure 3, lanes 5 and 7) are in agreement with BLAST hits to DM244 and DM4 sequences. For E022J22, we also observed one signal on tomato (Figure 3, lane 2) and two foci on potato (Figure 3, lanes 1 and 5), consistent with two BLAST hits to DM244 and DM12, which are 1 Mb apart on the physical map. The corresponding cytogenetic spacing between two foci from E022J22 appeared to be comparable to that ob-

© 2012 The Authors

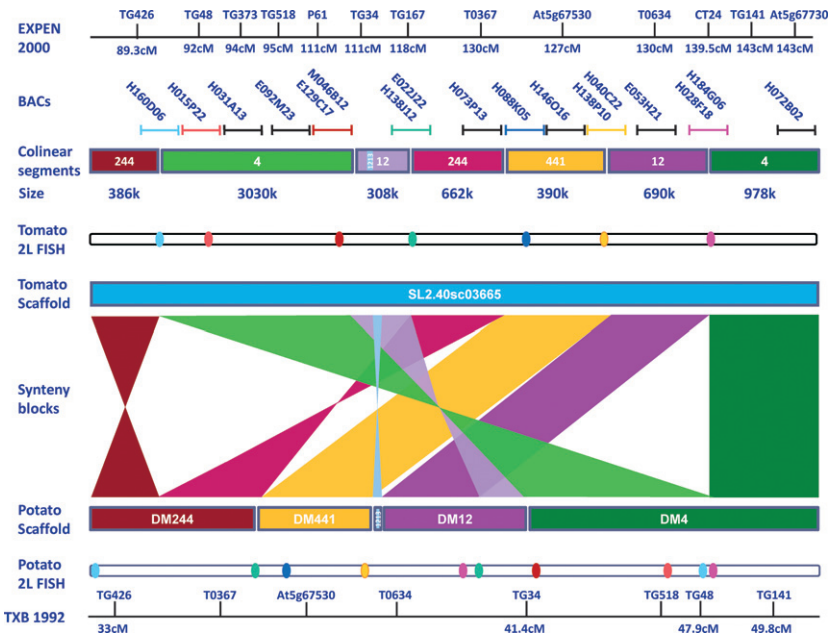



**Figure 4.** Comparative map for tomato and potato 2L. Chromosome 2 markers in tomato and potato scaffolds are indicated on the top and bottom bars with cM positions from the tomato EXPEN2000 and potato TXB1992 genetic maps. The order and relative positions of markers without a genetic position are derived from BLASTN hits to potato scaffolds. Aligned segments from potato DM scaffolds to tomato scaffold SL2.40sc03665 are represented by colored rectangles. Syntenic blocks are represented by colored polygons in the middle section, and correspond to the position and orientation of rearranged potato segments. The relative positions of tomato BACs flanking or spanning colinearity breaks are labeled with their identifier above each DM segment. The relative positions and order of the tomato BACs on the tomato and potato FISH map are indicated by corresponding colored dots. Some BACs spanning a colinearity break have one corresponding position on the tomato FISH map and two corresponding positions on the potato FISH map.

served between H160D06 and E022J22 (Figure 3, lane 5). This result adds to the notion that 660 and 386 kb segments from DM244 have an inverted orientation, of which the 660 kb segment is also translocated in tomato. Furthermore, mapping confirms that potato accessions G254 and RH890391 share a similar 2L topology with potato DM. Based on these results, we reconstructed the tomato and potato chromosomal organization (Figure 4).

### Orthologs on tomato and potato 2L and linkage conservation

The recent tomato and potato genome sequencing efforts and available genetic map information enable analysis of the gene repertoire and relative order and orthology detection, a prerequisite to investigate the extent of conserved linked genes in repositioned homeologous segments. In this respect, Bonierbale *et al.* (1988) have demonstrated that cDNA markers are largely colinear along the tomato and potato chromosomes. In particular, the segments from tomato and potato 2L have several EST-derived genetic markers in common, and thus appear to be conserved syntenic (Figure 4). To assess the conserved linkage, we used the complete set of ITAG2 and PSCG release 3.4 annotated protein sequences (see Experimental procedures) to identify and map the main orthologs in syntenic tomato and potato 2L segments, respectively.

*Ortholog detection.* The syntenic 2L segments contain 893 predicted genes in tomato and 820 in potato, respectively. Use of InParanoid (Remm *et al.*, 2001) identified 664 ortholog groups, of which 623 comprise an ortholog pair localized in corresponding colinear segments, consistent with a conserved syntenic nature. For 25 ortholog groups, 14 tomato and 11 potato ortholog genes mapped outside homeologous segments (Table S2). Within the 623 ortholog groups, we detected 721 tomato genes and 679 potato genes, apparently indicating that gene duplications have occurred. The mean tomato and potato gene copy (paralog) numbers for the 623 ortholog groups are 1.16 and 1.09, respectively (Table S2). There are 36 ortholog groups with more than two members, of which 12 groups have multiple tomato paralogs, 16 groups have multiple potato orthologs, and eight groups have multiple tomato and potato paralogs. In two ortholog groups, tomato genes Solyc02g090350 and Solyc02g085990 have adjacent paralogs. Thus, although it appears that tomato and potato share a comparable basic set of genes overall, the order and number of gene copies is substantially different.

*Gene adjacency, orientation and unclustered genes.* We found 335 ortholog pairs that have a conserved linkage, and 288 pairs that have disrupted gene adjacency. For example, comparison of tomato genes with their potato orthologs just downstream of the F18 junction in scaffold DM12 showed that local gene vicinities are preserved but gene adjacency is disrupted. Tomato and potato orthologs were frequently interrupted by putative LTR retrotransposons (Figure S4). In total, we predicted 28 LTRs in tomato scaffold 3665 and 56 LTRs in the colinear potato DM scaffolds.

Of the identified orthologs, eight orthologous gene pairs show irregular positioning compared to the layout of the colinear segments (Figure 5 and Table S3). The orthologs corresponding to the gene models located on potato scaffold DM1213 are shifted by 1.23 Mb upstream and inverted in tomato (Figure 5, genes 'e' and 'f', and Figure 2). Here, the provisional state of the current potato scaffold order appears to be the most straightforward explanation for this aberrant

© 2012 The Authors

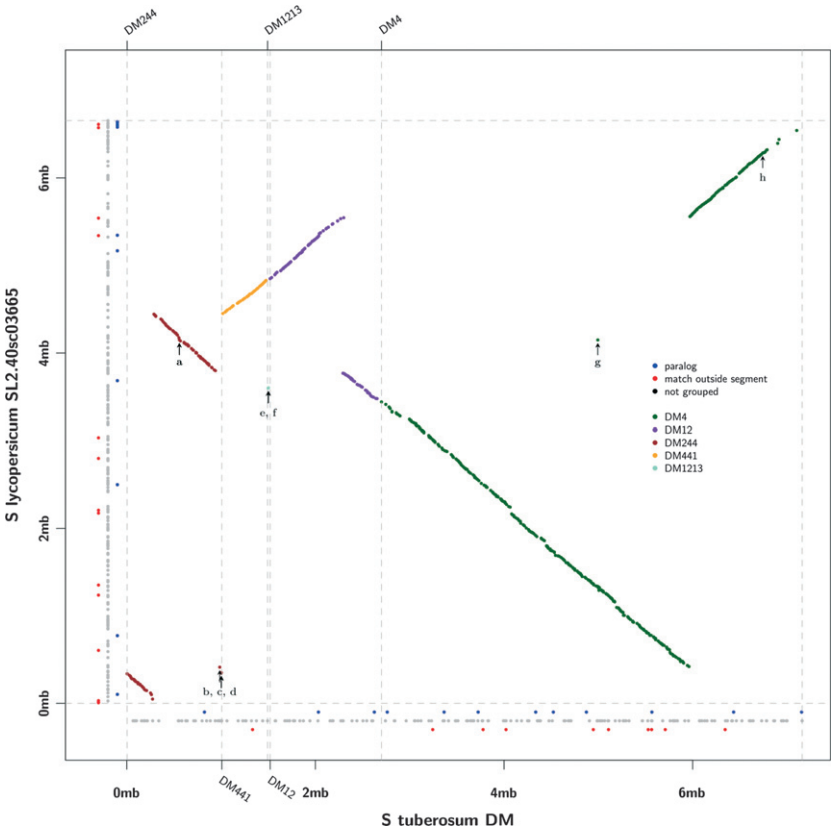



**Figure 5.** Orthologous gene matches for potato versus tomato. The first 6.7 Mb of tomato scaffold SL2.40sc03665 and potato scaffolds DM4 (green), DM12 (purple), DM244 (red), DM441 (yellow) and DM1213 (blue) are shown. The midpoint of a gene model is represented by a dot. An orthologous gene pair is indicated by a dot in the first quadrant, unclustered genes are shown at −100 000, paralogs are shown at −50 000, and clustered genes without a matching counterpart in the selected segment are shown at −500 000 along the corresponding axis. Ortholog pairs with an irregular position are indicated by arrows (a) to (h).

position, although a biological rearrangement cannot be excluded. In tomato, this domain contains two AP2-domain transcription factors. ITAG RepeatMasker annotation (see Experimental procedures) shows that these AP2 homologs are flanked by copies of LTRs that possibly originate from the same transposon. Furthermore, there are three orthologous gene pairs that remain as relicts of the proposed splitting of scaffold DM244 into separate contigs (Figure 5, genes b, c and d). In addition, we found two pairs of orthologs with an inverted orientation to the syntenic segment (Figure 5, genes a and h), and one ortholog pair that shows transposition of the tomato ortholog to a region approximately 2.8 Mb distal of its expected position (Figure 5, gene g).

A substantial proportion of the 241 and 173 putative genes, respectively, in the tomato and potato aligned segments remained unclustered (Figure 5), and probably are the result of stringent clustering cut-off values or aberrantly predicted genes in the gene annotations, or they may be species-specific. Furthermore, some unclustered genes are probably false negatives, taking into account that Inparanoid has a false-negative rate of approximately 3% when omitting the use of an outgroup (Remm *et al.*, 2001).

### A rearrangement pathway model

There is compelling evidence that recombination plays a much larger role in the evolution of plant genomes than previously appreciated. Recombination in plants is highly variable, and includes (i) meiotic recombination between homologous chromosomes, (ii) intra-strand crossing over between direct and inverted repeats, (iii) unequal crossing-over between misaligned repeats on homologous chromosomes, and (iv) illegitimate or non-homologous recombination (Gaut *et al.*, 2007). Although the mechanisms that underlie complex genome rearrangements in plants are quite diverse and not fully understood, there is accumulating evidence that links rearrangement breakpoints and repeats (Coghlan *et al.*, 2005). Thus, rearrangements in tomato and potato 2L may be explained in conformance with the type and location of repeats. We detected similarity to class I and II transposons in 20 kb bins near the synteny breaks, in particular sonata, harbiner, Long Interspersed Elements (LINEs), Short Interspersed Elements (SINEs), unknown retrotransposons and telomere-like related sequences (Figures S5 and S6). Furthermore, comparative sequence analysis revealed two copies of a 4.5 kb inverted repeat on tomato 2L flanking the junction between the 690 and 978 kb syntenic blocks, and spanning the synteny junction between the 308 and 662 kb segments. Another repeat of 267 bp is located near the 386–3030 kb segment junction and the 308–662 kb segment junction (Figure S7). We did not detect similarity to known transposable elements (TEs), and therefore these repeats are not likely to be of transposon origin. Currently, we do not have direct evidence for their involvement in the rearrange-

© 2012 The Authors

ments, but, taking into account their location near synteny breaks, we hypothesize that transposon element/repeat-mediated recombination may be explanatory for the 2L rearrangements as follows. We assume the existence of unichromosomal breakpoints, as we have not found any 7 Mb segmental duplication in other parts of the tomato genome, and according to Pevzner and Tesler (2003), that the unichromosomal breakpoints are related and inter-dependent. Furthermore, the 2L rearrangements are probably tomato lineage-specific. This notion is substantiated by several genome mapping studies (Bonierbale *et al.*, 1988; Tanksley *et al.*, 1988, 1992; Livingstone *et al.*, 1999; Thorup *et al.*, 2000; Doganlar *et al.*, 2002b; Pertuzé *et al.*, 2002; Ashrafi *et al.*, 2009) and the colinear BAC FISH maps for 2L of *S. lycopersicum*, *S. lycopersicoides*, *S. pennellii* and *S. chilense* (H.d.J., unpublished results). Therefore, we assumed the potato and pepper organization to be ancestral, and thus the rearrangements needed to transform the tomato into potato chromosome topology can be used to reconstruct an ancestral karyotype. Taking this evolution direction into account, we propose a rearrangement pathway model in which intra-strand crossing-over and ectopic recombination give rise to segmental rearrangements (Figure 6). First, two segments from DM244 become inverted by ectopic recombination mediated by copies of the same transposon elements that are positioned near the synteny breaks, or by non-homologous recombination. After the second reversal, a chromatid strand folds back, generating local pairing of the 276 bp inverted repeats, subsequent to which, inter-chromatid crossing-over occurs, leading to a third inversion. After breakage and fusion, this generates an intermediate configuration in which the 4.5 kb repeat is inverted. The 4.5 kb inverted repeat element then mediates a second intra-strand crossing-over event, resulting in a fourth inversion. Indeed, when using GRIMM (Tesler, 2002), a single and most parsimonious scenario was predicted, which consisted of a trajectory involving four recombination events (Figure S8) that resembled the conversion steps presented in the rearrangement model proposed above.

## DISCUSSION

### Unravelling structural differences in *Solanum* and *Capsicum*

In recent decades, marker-based mapping studies provided broad knowledge of the structural and molecular organization of various plant genomes, including tomato and potato. It is apparent that plant genomes share extensive conserved linkage despite their diversity in size and complexity (Bonierbale *et al.*, 1988; Tanksley *et al.*, 1992; Bennetzen, 2000a). These studies also indicated large-scale differences in genomic organization, most of which were found in heterochromatic domains. Nevertheless, comparative maps have limitations that arise from low marker density and low placement accuracy that complicate local resolution of chromosomal organization. Tang *et al.* (2008) and Peters *et al.* (2009) used BAC FISH painting to unravel complex rearrangements in euchromatin regions and elucidated discrepancies between genetic and physical maps. In this paper, we provide more detailed mapping at the sequence level in order to identify the coordinates of colinear segments and subsequent selection of BAC targets to validate the colinearity breaks. We have shown that large-scale structural differences in *Solanum* and *Capsicum* are not only confined to heterochromatin portions, but frequently occurred in the euchromatic portion of 2L, 6S, 10L and 11L. Thus, although genetic maps and comparative sequence analysis each have limitations in power and resolution, their combined usage is indispensable for unsurpassed

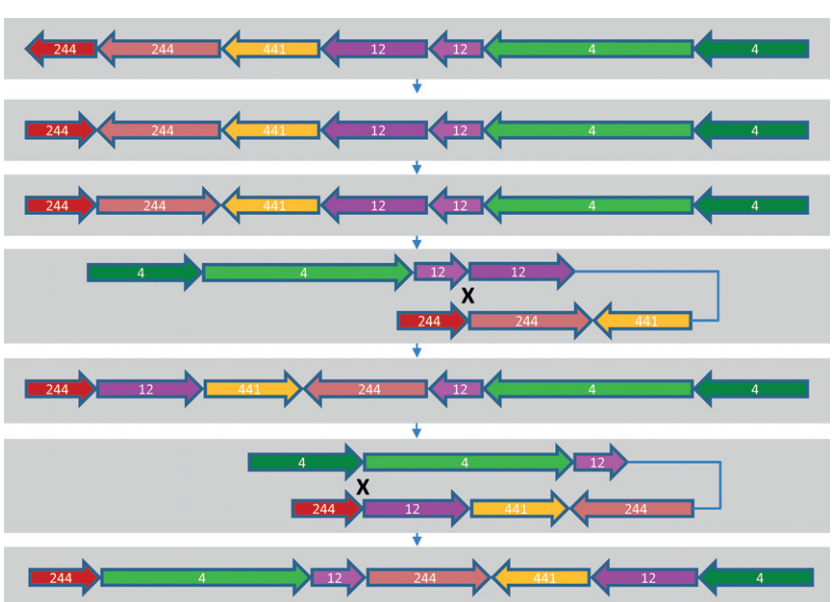



**Figure 6.** Proposed rearrangement scenario involving four structural conversions given the initial potato order (upper panel) and final tomato order (bottom panel) of 2L segments. Positions of intra-strand crossing-overs between segments are marked 'X'. Arrows indicate the relative order and orientation of 2L segments.

© 2012 The Authors

knowledge of chromosome organization at the structural and sequence level.

**Large-scale rearrangements**

Comparative sequence alignment indicated the location of colinearity breaks between tomato and potato DM 2L. Using BACs that span junction breaks, FISH revealed double signals marking the borders of colinear segments in potato accessions G254 and RH890316. In total, six large-scale rearrangements were discovered on a 7 Mb euchromatic region of 2L between tomato cv. Heinz 1706 and potato G254 and RH8903916. Although chromosome 2 markers TG34 and TG48 in the 89–143 cM intervals show an inverted order (Tanksley *et al.*, 1992), these structural differences between tomato and potato were not apparent from linkage maps.

Linkage maps between tomato and several wild relatives show a similar order of 2L genetic markers (Tanksley *et al.*, 1992; Fulton *et al.*, 1997; Pertuzé *et al.*, 2002; Ashrafi *et al.*, 2009; http://solgenomics.net), but they appear to be inverted for pepper, eggplant/aubergine (Livingstone *et al.*, 1999; Thorup *et al.*, 2000; Doganlar *et al.*, 2002b; http://solgenomics.net) and potato (this paper). Furthermore, potato, pepper and eggplant share extensive marker colinearity in the homeologous 2L segments. Taken together, these results suggest that the rearrangements in 2L are specific for the tomato clade, and thus occurred after the split from the common ancestor of tomato and potato.

A tomato-specific rearrangement was also reported for chromosome 10, for which the tomato, potato and pepper homeologs apparently differ by a paracentric inversion (Tanksley *et al.*, 1988, 1992; Livingstone *et al.*, 1999). Comparative linkage map studies in sister species of tomato, *S. lycopersicoides* and *S. sitiens* (Pertuzé *et al.*, 2002) showed a 10L configuration that was similar to that of potato and pepper (*Capsicum*), and this provided further evidence for the notion that the paracentric inversion was already fixed in the common ancestor for the tomato lineage. The FISH mapping presented here confirms two inversions in tomato 10L compared to the potato and pepper chromosome organization, and is consistent with the tomato lineage-specific nature of the rearrangements.

The structural organization appears less similar in tomato and potato 11L. The comparative sequence alignment indicates multiple breakpoints corresponding to three inversions and three deletions. Although an accurate sequence alignment for pepper 11L is currently lacking, cytogenetic mapping revealed at least one translocated segment with a reversed orientation. By comparing tomato and pepper genetic maps, a minimum of 22–32 breakages of tomato chromosomes would be necessary to transform the order and position of tomato genes to that observed in pepper (Tanksley *et al.*, 1988; Livingstone *et al.*, 1999). We have not investigated all putative chromosome rearrangements, and it is conceivable the number of rearrangements is an underestimate and a consequence of a low density of genetic markers. This notion is supported by our observation that the previously reported hidden inversion for the 4.5 Mb short arm of chromosome 6 (Tang *et al.*, 2008) actually appears more complex. A comparative sequence alignment revealed multiple reversals, translocated segments and a deletion in potato 6S. From the present results and those of previous studies, we propose a genome landscape in which evolution on the structural level for the majority of the 12 chromosomes in Solanaceae was far more dynamic than currently appreciated.

**Chromosomal rearrangements and reproductive isolation**

The role of the large structural rearrangements in 2L with respect to inter-generic and intra-generic reproductive isolation and speciation of Solanaceae, remains unclear. Rearrangements can impede proper pairing of homeologous chromosomes and reduce recombination, and may also cause decreased fitness or even sterility (Noor *et al.*, 2001; Livingstone and Rieseberg, 2004; Rieseberg and Willis, 2007; Bedinger *et al.*, 2010). For example, inter-specific hybrids between *S. lycopersicum* and *S. pennellii* are highly fertile, and light microscopy analysis demonstrated near normal levels of meiotic pairing and crossing-over (Tanksley *et al.*, 1992). However, high-resolution electron microscopic analysis of chromosomes at pachytene showed frequent unusual synaptic configurations (Anderson *et al.*, 2010). The latter observation might seem in contradiction with fertility and the unaffected recombination between tomato and *S. pennellii*. However, it is possible that relatively small structural differences between homeologous chromosomes may be tolerated, but relatively large rearrangements may disrupt meiotic synapsis or recombination and may result in infertility. For example, inter-specific hybrids between *S. lycopersicoides* and *S. sitiens* are fertile and show recombination rates similar to tomato (Pertuzé *et al.*, 2002). In contrast, inter-generic hybrids of *S. lycopersicum* with *S. lycopersicoides* or *S. sitiens* are sterile and show genome-wide suppressed recombination (Chetelat *et al.*, 1997). Moreover, recombination is completely abolished for 10L, which may be explained by a large paracentric inversion between the L-type (*Lycopersicon* spp.) and S-type (*S. lycopersicoides* and *S. sitiens*) genomes (Pertuzé *et al.*, 2002).

Although we have not investigated the molecular nature of inter-specific barriers between tomato and potato, the rearrangements presented here may very well impede proper synapsis. In line with this are results obtained from somatic hybrids of tomato and potato, in which irregular synapsis was also frequently observed (De Jong *et al.*, 1993). The latter may perhaps involve absent, repositioned or protected blocks of genes that suppress homeologous pairing (Bedinger *et al.*, 2010). In this respect, a phenomenon known as transmission ratio distortion is often observed in inter-specific crosses between tomato and wild relatives.

© 2012 The Authors

Possibly, selection against particular allelic combinations that are associated with 'transmission ratio distortion' loci underlies hybrid incompatibility (Moyle and Graham, 2006). For example, strong reproductive barriers have been observed between *S. lycopersicum* and the tomato-like nightshades *S. ochrantum* and *S. juglandifolium*, which have been placed phylogenetically and morphologically in an intermediate position between tomato and potato (Spooner *et al.*, 2005). Remarkably, we physically mapped markers associated with distortion loci *trd2.2*, *sd10.1*, *sd10.2*, and *trd11.1* from *Solanum* species (Pertuzé *et al.*, 2002; Albrecht and Chetelat, 2009) to the rearranged segments in 2L, 10L and 11L discussed here (Figures S1–S3), suggesting a relationship between hybrid incompatibility, transmission ratio distortion and large rearrangements.

**Small-scale rearrangements and changes in gene repertoire**

Overall, tomato and potato have a 7% difference in gene copy number in rearranged 2L segments (Table S2). Studies on gene repertoire and gene order in tomato and pepper indicated a change in locus number for approximately 12% of the loci, accompanied by an extensively modified linear order of genes and many chromosome rearrangements (Tanksley *et al.*, 1988). These differences in copy number are in line with the slightly larger genome size of tomato compared to potato and the expanded pepper genome compared to tomato.

Our observation that small rearrangements are more frequent than large-scale differences seems consistent with earlier observations made for other plant genomes (Bennetzen, 2000a). In particular, micro-colinearity and conserved linkage between orthologs is apparent, but we nevertheless found many small exceptions. Observations in yeast and *Drosophila*, for example, show that direct repeats of LTR transposon copies may act in reciprocal recombination, giving rise to gene loss. Reciprocal recombination between inverted repeats from LTR retrotransposons may result in gene inversions, while recombination between repetitive elements on different chromosomes may lead to reciprocal translocation (Bennetzen, 2000b; Gray, 2000). In some cases, we suspect the micro-colinearity in 2L has been disrupted by TEs. For example, two translocated AP2-like transcription factors in tomato 2L with a disrupted conserved linkage are flanked by LTRs from a single retrotransposon, suggesting its involvement in the relocation and inversion of these genes. Strikingly, the inversion of AP2-like transcription factors in tomato maps to near the *fw2.2* and *fs2.2* loci controlling fruit weight and bell-shaped fruit morphology in the heirloom tomato cv. Yellow Stuffer and garden pepper (Grandillo *et al.*, 1999; Van der Knaap and Tanksley, 2003). The inverted context of genes functioning in signalling pathways has been reported to affect traits and cause phenotypic changes in *Drosophila* (Hoffman *et al.*, 2004). However, we currently have no functional or phenotypic evidence indicating an inversion-induced difference in gene interaction or regulation in tomato.

**Analyses of chromosomal rearrangement junctions and rearrangement scenario**

Currently, it is unclear what might have caused the large rearrangements in 2L. TEs are known to induce large inversions, deletions and translocations, mediated directly via their transposition mechanisms, or indirectly via homologous recombination, or by ectopic or non-allelic homologous recombination (Bennetzen, 2000b). For example, in maize (*Zea mays*), pairs of TEs spaced beyond 100 kb are efficient chromosome breakers, generating deletions and inversions via alternative tranposition (Huang and Dooner, 2008; Zhang *et al.*, 2009). Recombination in plants is not limited to homologous chromosomes only. For example, inter-chromatid recombination can result in deletions and inversions, and such recombination between different genomic regions can lead to large chromosomal rearrangements (Gaut *et al.*, 2007). Furthermore, it is important to realize that meiotic as well as mitotic rearrangements in plants can be passed to progeny. In Arabidopsis, elevated somatic recombination rates have been observed and found to be positively correlated with DNA damage and stress, suggesting that genomic flux caused by recombination plays an important role in environmental stress adaption. Direct evidence for genomic flux involving large-scale rearrangements caused by repeat-mediated recombination has not been reported for plants. However, indirect evidence has been found in Brassicaceae species, in which rearrangements are to a large extent located near repetitive sequences (Ziolkovski *et al.*, 2003; Lysak *et al.*, 2006).

In general, any chromosomal rearrangement involves a breakage and a subsequent repair of the chromosome ends or fusion to another chromosome end. Of the different types of rearrangements in plant genomes, inversions probably occur the most frequently, and can range size from a few kb up to hundreds of genes in length (Coghlan *et al.*, 2005). Our results are in agreement with this notion, showing that inversion is the predominant rearrangement type in tomato 2L, 6S, 10L and 11L. Previously, Livingstone *et al.* (1999) inferred the most recent ancestral genome using lineage-specific rearrangements, the phylogeny of tomato, potato and pepper (Spooner *et al.*, 1993), and comparative maps. We reasoned that, as the potato and pepper 2L organization appear similar, with pepper being considered as an outgroup, and because the rearrangements are tomato lineage-specific, the potato/pepper topology may be considered ancestral. In the proposed model, the inversion rearrangement was preferred above other types, and the reconstruction was directed from potato/pepper towards the tomato 2L organization. Indeed, when calculating a most parsimonious rearrangement scenario, only four reversion steps were needed to transform the potato 2L organization

© 2012 The Authors

into the tomato 2L organization. However, we currently cannot exclude other rearrangement pathways. Tandem repeats and (retro)transposons are present also outside the 2L junction breaks, and as we have no indications about the relative order of the proposed inversion steps, the proposed model should be considered as a working hypothesis.

The efficiency and success of introgressive hybridization breeding on the basis of DNA-based selection depends, among others, on adequate identification of chromosome organization. The implication of technological advances with respect to next-generation sequencing technology for the use of extant germplasm resources is that large numbers of complex genomes can be sequenced relatively fast and cheaply. This will undoubtedly speed up identification of compatible genomes for introgression breeding, the rearrangement phylogeny within the Solanaceae, and reconstruction of the ancestral *Solanum* karyotype. The present project is a first step in mining of structural genetic diversity and the development of genome-based breeding tools.

## EXPERIMENTAL PROCEDURES

### Scaffold selection

*S. lycopersicum* Heinz 1706 (build 2.40) and *S. tuberosum* DM scaffolds were obtained from ftp://ftp.solgenomics.net/tomato_genome/wgs/assembly/ and http://potatogenomics.plantbiology.msu.edu/, respectively. Tomato and potato scaffolds were aligned against genetic and physical maps from the Sol Genomics Network (http://solgenomics.net) to verify the chromosomal location. Genome sequences were aligned using MUMmer 3.22 (http://mummer.sourceforge.net/) with the tomato scaffold as a reference. Subsequently, the coords file output from generated by the MUMmer script 'NUCmer' was converted into ClustalW format and used as an input for the generic synteny browser GBrowse_syn (http://gmod.org/wiki/Synteny) for visualization and further analysis. Tomato Heinz 1706 BACs and potato BAC sequences from the RH clone library were aligned to scaffolds using BLASTN.

### BAC selection, growth and DNA preparation

For solanaceous species-derived stress tolerance genes, a genomic location was determined by best BLASTN hits against tomato scaffolds from the Sol Genomics Network (http://solgenomics.net). Translated coding regions from non-solanaceous species-derived stress tolerance genes were used in a TBLASTX screen against tomato unigene sequences. Unigenes were then used for a BLASTN screen against tomato scaffolds to determine sequence coordinates. Subsequently, BAC clones containing stress tolerance gene homologs were identified by BLASTN of a scaffold sequence interval against a tomato BAC end sequence database, and linked to the EXPEN2000 genetic map by screening the genomic intervals against the genetic marker database. Subsequently, BAC clones of tomato cv. Heinz 1706 *Hin*dIII, *Eco*RI and *Mbo*I libraries were grown, and BAC DNA was isolated as described previously (Peters *et al.*, 2009).

### Chromosome preparations

All FISH experiments were performed on tomato cv. Heinz 1706 ($2n = 2x = 24$). Pachytene preparations from young anthers containing pollen mother cells and spreads of extended DNA fibers from young leaves were made as described by Zhong *et al.* (1996b) and Budiman *et al.* (2004).

### Fluorescence *in situ* hybridization (FISH)

Two-color and multi-color FISH of BAC clones to pachytene chromosomes were performed as described by Zhong *et al.* (1996b). Slides were examined under an Axioplan 2 imaging photomicroscope (Zeiss, Jena, Germany) equipped with epifluorescence illumination, and small band filter sets for DAPI (4′,6-diamino-2-phenylindole) and for FITC (fluorescein-5-isothiocyanate), Cy3 (cyanine 3), Cy5 (cyanine 5), DEAC (7-diethylaminocoumarin-3-carboxylic acid) and Cy3.5 (cyanine 3.5) fluorescence. Capturing of selected images and image processing were performed as previously described (Szinay *et al.*, 2008).

### Sequence annotation

Interspersed repeats were identified through similarity searches to the *Magnoliaphyta* section of the RepBase repeat database (release 2008-08-01) (Jurka *et al.*, 2005) and the Institute for Genomic Research *Lycopersicon* repeats version 3.1 and Solanaceae repeats version 3.1 (currently available via http://plantrepeats.plantbiology.msu.edu/) using RepeatMasker 3.2.5 (http://www.repeatmasker.org/) and cross_match 0.990319 (http://www.phrap.org/). In addition, LTR retrotransposons were predicted *ab initio* using LTR Finder (Xu and Wang, 2007) (http://tlife.fudan.edu.cn/ltr_finder). *Ab initio* gene prediction on the repeat masked sequences was performed using Genscan (Burge and Karlin, 1997) using *Arabidopsis thaliana* gene models. Alignments of tomato and potato ESTs and genetic markers were generated using blastn 2.2.17 (Altschul *et al.*, 1997). Tomato and potato EST sequences and marker sequences were downloaded from the Sol Genomics Network (http://solgenomics.net).

### Orthology analysis

For ortholog detection, tomato ITAG annotation 2.31 (http://solgenomics.net/gbrowse/bin/gbrowse/ITAG1_genomic/) and potato PGSC annotation 3.4 (Potato Genome Sequencing Consortium, 2011) were used to obtain the sequences of the predicted proteins. In each case, only the longest transcripts were taken into consideration. Reciprocal best hits were obtained using BLASTP (Altschul *et al.*, 1997) and results were clustered into orthologous groups by InParanoid version 4.1 (Remm *et al.*, 2001) with a score cut-off of 40. Further processing of annotation data to fit the coordinates on the selected segment was perfomed using custom Perl and R scripts in combination with BioPerl (http://www.bioperl.org/wiki/Main_Page).

The coordinates and identity of ortholog group members, together with the mapped positions of BAC sequences, LTR retrotransposons and unigenes and coordinates for mapped annotated proteins were subsequently stored in a gff format for visualization of the local gene repertoire with SynBrowse (Pan *et al.*, 2005).

The predicted orthologs were clustered into ortholog groups and subsequently classified into genes with conserved or disrupted linkage based on the linear order of orthologs.

## ACKNOWLEDGEMENTS

This work is supported by Technological Top Institute Green Genetics grant number 2CC037RP, and financial aid from Rijk Zwaan, Syngenta AG and Monsanto.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

© 2012 The Authors

**Figure S1.** Integrated map of tomato and potato chromosome 2L segments.
**Figure S2.** Comparative map of tomato, potato and pepper chromosome 10L.
**Figure S3.** Comparative map of tomato, potato and pepper chromosome 11L.
**Figure S4.** Tomato synteny browser snapshot of a 150 kb region at the F18 junction.
**Figure S5.** Distribution of class I transposon-related repeats in chromosome 2L.
**Figure S6.** Distribution of class II transposon-related repeats in chromosome 2L.
**Figure S7.** Identity plot of synteny junctions.
**Figure S8.** Parsimonious trajectory for rearrangements between potato and tomato chromosome 2L.
**Table S1.** Large-scale rearrangements between tomato, potato and pepper.
**Table S2.** Predicted ortholog pairs and paralogs in tomato and potato chromosome 2L.
**Table S3.** Outliers identified by colinearity analysis of orthologous gene pairs.
Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

## REFERENCES


**Albrecht, E. and Chetelat, R.T.** (2009) Comparative genetic linkage map of *Solanum* sect. Juglandifolia: evidence of chromosomal rearrangements and overall synteny with tomatoes and related nightshades. *Theor. Appl. Genet.* **118**, 831–847.

**Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J.** (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402.

**Anderson, L.K., Covey, P.A., Larsen, L.R. and Stack, S.M.** (2010) Structural differences in chromosomes distinguish species in the tomato clade. *Cytogenet. Genome Res.* **29**, 24–34.

**Ashrafi, H., Kinkade, M. and Foolad, M.R.** (2009) A new genetic linkage map of tomato based on a *Solanum lycopersicum* x *S. pimpinellifolium* RIL population displaying locations of candidate pathogen response genes. *Genome*, **52**, 935–956.

**Bai, Y. and Lindhout, P.** (2007) Domestication and breeding of tomatoes: what have we gained and what can we gain in the future? *Ann. Bot.* **100**, 1085–1094.

**Bai, Y., van der Hulster, R., Huang, C.C., Wei, L., Stam, P. and Lindhout, P.** (2004) Mapping *Ol-4*, a gene conferring resistance to *Oidium neolycopersici* and originating from *Lycopersicon peruvianum* LA2172, requires multi-allelic, single-locus markers. *Theor. Appl. Genet.* **109**, 1215–1223.

**Bedinger, P.A., Chetelat, R.T., McClure, B. *et al.*** (2010) Interspecific reproductive barriers in the tomato clade: opportunities to decipher mechanisms of reproductive isolation. *Sex. Plant Reprod.* **24**, 171–187.

**Bennetzen, J.L.** (2000a) Comparative sequence analysis of plant nuclear genomes: microcolinearity and its many exceptions. *Plant Cell*, **12**, 1021–1029.

**Bennetzen, J.L.** (2000b) Transposable element contributions to plant gene and genome evolution. *Plant Mol. Biol.* **42**, 251–269.

**Bonierbale, M.W., Plaisted, R. and Tanksley, S.D.** (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. *Genetics*, **120**, 1095–1103.

**Budiman, M.A., Chang, S.-B., Lee, S., Yang, T.J., Zhang, H.-B., de Jong, H. and Wing, R.A.** (2004) Localization of *jointless-2* gene in the centromeric region of tomato chromosome 12 based on high resolution genetic and physical mapping. *Theor. Appl. Genet.* **108**, 190–196.

**Burge, C. and Karlin, S.** (1997) Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268**, 78–94.

**Chang, S.-B., Yang, T.-J., Datema, E. *et al.*** (2008) FISH mapping and molecular organization of the major repetitive sequences of tomato. *Chromosome Res.* **16**, 919–933.

**Chetelat, R.T. and Ji, Y.** (2007) Cytogenetics and evolution. In *Genetic Improvement of Solanaceous Crops* (Razdan, M.K. and Mattoo, A.K., eds). Enfield, NH: Science Publishers, pp. 77–112.

**Chetelat, R.T., Cisneros, P., Stamova, L. and Rick, C.M.** (1997) A male-fertile *Lycopersicon esculentum* x *Solanum lycopersicoides* hybrid enables direct backcrossing to tomato at the diploid level. *Euphytica*, **95**, 99–108.

**Coghlan, A., Eichler, E.E., Oliver, S.G., Paterson, A.H. and Stein, L.** (2005) Chromosome evolution in eukaryotes: a multi-kingdom perspective. *Trends Genet.* **21**, 673–682.

**Datema, E., Mueller, L.A., Buels, R., Giovannoni, J.J., Visser, R.G.F., Stiekema, W.J. and Van Ham, R.C.H.J.** (2008) Comparative BAC end sequence analysis of potato and tomato reveals overrepresentation of specific gene families in potato. *BMC Plant Biol.* **8**, 34.

**De Jong, J.H., Wolters, A.M.A., Kok, J.M., Verhaar, H. and Van Eden, J.** (1993) Chromosome pairing and potential for intergeneric recombination in some hypotetraploid somatic hybrids of *Lycopersicum esculentum* (+) *Solanum tuberosum*. *Genome*, **36**, 1032–1041.

**De Jong, J.H., Zhong, X.-B., Fransz, P.F., Wennekes-van Eden, J., Jacobsen, E. and Zabel, P.** (2000) High resolution FISH reveals the molecular and chromosomal organisation of repetitive sequences of individual tomato chromosomes. In *Chromosomes Today*, volume 13 (Olmo, E. and Redi, C.A., eds). Basel, Switzerland: Birkhäuser Verlag, pp. 267–275.

**Doganlar, S., Frary, A., Daunay, M. -C., Lester, R.N. and Tanksley, S.D.** (2002a) Conservation of gene function in the *Solanaceae* as revealed by comparative mapping of domestication traits in eggplant. *Genetics*, **161**, 1713–1726.

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

**Ehrlich, J., Sankoff, D. and Nadeau, J.H.** (1997) Synteny conservation and chromosome rearrangements during mammalian evolution. *Genetics*, **147**, 289–296.

**Foolad, M.R.** (2007) Current status of breeding tomatoes for salt and drought tolerance. In *Advances in Molecular Breeding toward Drought and Salt Tolerant Crops* (Jenks, M.A., Hasegawa, P.M. and Jain, S.M., eds). Dordrecht, The Netherlands: Springer, pp. 669–700.

**Fulton, T.M., Nelson, J.C. and Tanksley, S.D.** (1997) Introgression and DNA marker analysis of *Lycopersicum peruvianum*, a wild relative of the cultivated tomato, into *Lycopersicum esculentum*, followed through three successive backcross generations. *Theor. Appl. Genet.* **95**, 895–902.

**Fulton, T.M.R., van der Hoeven, R., Eannetta, N.T. and Tanksley, S.D.** (2002) Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. *Plant Cell*, **14**, 1457–1467.

**Ganal, M.W., Lapitan, N.L.V. and Tanksley, S.D.** (1991) Macrostructure of the tomato telomeres. *Plant Cell*, **3**, 87–94.

**Gaut, B.S., Wright, S.I., Rizzon, C., Dvorak, J. and Anderson, L.K.** (2007) Recombination: an underappreciated factor in the evolution of plant genomes. *Nature*, **8**, 77–84.

**Grandillo, S., Ku, H.M. and Tanksley, S.D.** (1999) Identifying the loci responsible for natural variation in fruit size and shape in tomato. *Theor. Appl. Genet.* **99**, 978–987.

**Gray, Y.H.M.** (2000) It takes two transposons to tango. *Trends Genet.* **16**, 461–468.

**Grube, R.C., Radwanskii, E.R. and Jahn, M.** (2000) Comparative genetics of disease resistance within the *Solanaceae*. *Genetics*, **155**, 873–887.

**Hoffman, A.A., Sgrò, C.M. and Weeks, A.R.** (2004) Chromosomal inversion polymorphisms and adaption. *Trends Ecol. Evol.* **19**, 482–488.

**Huang, J.T. and Dooner, H.K.** (2008) Macrotransposition and other complex chromosomal restructuring in maize by closely linked transposons in direct orientation. *Plant Cell*, **20**, 2019–2031.

**Iovene, M., Grzebelus, E., Carputo, D., Jiang, J. and Simon, P.W.** (2008) Major cytogenetic landmarks and karyotype analysis in *Daucus carota* and other *Apiaceae*. *Am. J. Bot.* **95**, 793–804.

**Jahn, M., Paran, I., Hoffmann, K., Radwanski, E.R., Livingstone, K.D., Grube, R.C., Aftergoot, E., Lapidot, M. and Moyer, J.** (2000) Genetic mapping of the *Tsw* locus for resistance to the *Tospovirus Tomato spotted wilt virus* in


© 2012 The Authors

*Capsicum* spp. and its relationship to the *Sw-5* gene for resistance to the same pathogen in tomato. *Mol. Plant–Microbe Interact.* **13**, 673–682.

**Jenks, M.A., Hasegawa, P.M. and Jain, S.M.** (2007) *Advances in Molecular Breeding Toward Drought and Salt Tolerant Crops*. Dordrecht, The Netherlands: Springer.

**Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O. and Walichiewicz, J.** (2005) Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* **110**, 462–467.

**Knapp, S.** (2002) Tobacco to tomatoes: a phylogenetic perspective on fruit diversity in the *Solanaceae*. *J. Exp. Bot.* **53**, 2001–2022.

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

**Ku, H. -M., Vision, T., Liu, J. and Tanksley, S.D.** (2000) Comparing sequenced segments of the tomato and *Arabidopsis* genomes: large-scale duplication followed by selective gene loss creates a network of synteny. *Proc. Natl Acad. Sci. USA*, **97**, 9121–9126.

**Liharska, T.B., Koornneef, M., Van Wordragen, M., Van Kammen, A. and Zabel, P.** (1996) Tomato chromosome 6: effect of alien chromosomal segments on recombinant frequencies. *Genome*, **39**, 485–491.

**Livingstone, K. and Rieseberg, L.** (2004) Chromosomal evolution and speciation: a recombination-based approach. *New Phytol.* **161**, 107–112.

**Livingstone, K.D., Lackney, V.K., Blauth, J.R., van Wijk, R. and Jahn, M.K.** (1999) Genome mapping in *Capsicum* and the evolution of genome structure in the *Solanaceae*. *Genetics*, **152**, 1183–1202.

**Lysak, M.A., Berr, A., Pacinka, A., Schmidt, R., McBreen, K. and Schubert, I.** (2006) Mechanisms of chromosome reduction in *Arabidopsis thaliana* and related *Brassicaceae* species. *Proc. Natl Acad. Sci. USA*, **103**, 5224–5229.

**Moyle, L.C.** (2008) Ecological and evolutionary genomics in the wild tomatoes (*Solanum* sect. Lycopersicon). *Evolution*, **62**, 2995–3013.

**Moyle, L.C. and Graham, E. B.** (2006) Genome-wide associations between hybrid sterility QTL and marker transmission ratio distortion. *Mol. Biol. Evol.* **23**, 973–980.

**Noor, M.A.F., Grams, K.L., Bertucci, L.A. and Reiland, J.** (2001) Chromosomal inversions and the reproductive isolation of species. *Proc. Natl Acad. Sci. USA*, **98**, 12084–12088.

**Pan, X., Stein, L. and Brendel, V.** (2005) SynBrowse: a synteny browser for comparative sequence analysis. *Bioinformatics*, **21**, 17.

**Passarge, E., Horsthemke, B. and Farber, R.A.** (1999) Incorrect use of the term synteny. *Nat. Genet.* **23**, 387.

**Pertuzé, R.A., Ji, Y. and Chetelat, R.T.** (2002) Comparative linkage map of the *Solanum lycopersicoides* and *S. sitiens* genomes and their differentiation from tomato. *Genome*, **45**, 1003–1012.

**Peters, S.A., Datema, E., Szinay, D. *et al.*** (2009) *Solanum lycopersicum* cv. Heinz 1706 chromosome 6: distribution and abundance of genes and retrotransposable elements. *Plant J.* **58**, 857–869.

**Pevzner, P. and Tesler, G.** (2003) Genome rearrangements in mammalian evolution: lessons from human and mouse genomes. *Genome Res.* **13**, 37–45.

**Potato Genome Sequencing Consortium** (2011) Genome sequence and analysis of the tuber crop potato. *Nature*, **475**, 189–197.

**Remm, M., Storm, C.E.V. and Sonnhammer, E.L.L.** (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J. Mol. Biol.* **314**, 1041–1052.

**Rick, C.M., DeVerna, J.W., Chetelat, R.T. and Stevens, M.A.** (1987) Potential contributions of wide crosses to improvement of processing tomatoes. *Acta Hortic.* **200**, 45–55.

**Rieseberg, L.H. and Willis, J.H.** (2007) Plant speciation. *Science*, **317**, 910–914.

**Seah, S.J., Rossi, Y.M., Gleason, C.A. and Williamson, V.M.** (2004) The nematode-resistance gene, *Mi-1*, is associated with an inverted chromosomal segment in susceptible compared to resistant tomato. *Theor. Appl. Genet.* **108**, 1635–1642.

**Sesso, H.D., Liu, S., Gaziano, J.M. and Buring, J.E.** (2003) Dietry lycopene, tomato-based food products and cardiovascular disease in woman. *J. Nutr.* **133**, 2336–2341.

**Singh, R.J.** (2007) *Genetic Resources, Chromosome Engineering, and Crop Improvement: Vegetable Crops* Volume 3. Boca Raton, USA: CRC Press.

**Spooner, D.M., Anderson, G.J. and Jansen, R.K.** (1993) Chloroplast DNA evidence for the interrelationships of tomatoes, potatoes, and pepinos (*Solanaceae*). *Am. J. Bot.* **80**, 676–688.

**Spooner, D.M., Peralta, I.E. and Knapp, S.** (2005) Comparison of AFLPs with other markers for phylogenetic inference in wild tomatoes [*Solanum* L. section Lycopersicon (Mill.) Wettst.]. *Taxon*, **54**, 43–61.

**Szinay, D., Chang, S. -B., Khrustaleva, L., Peters, S., Schijlen, E., Bai, Y., Stiekema, W.J., van Ham, R.C.H.J., de Jong, H. and Klein Lankhorst, R.M.** (2008) High-resolution chromosome mapping of BACs using multi-colour FISH and pooled-BAC FISH as a backbone for sequencing tomato chromosome 6. *Plant J.* **56**, 627–637.

**Tang, X., Szinay, D., Lang, C. *et al.*** (2008) Cross-species bacterial artificial chromosome–fluorescence *in situ* hybridization painiting of the tomato and potato chromosome 6 reveals undescribed chromosomal rearrangements. *Genetics*, **180**, 1319–1328.

**Tanksley, A.D., Bernatzkey, R., Lapitan, N.L. and Prince, J.P.** (1988) Conservation of gene repertoire but not gene order in pepper and tomato. *Proc. Natl Acad. Sci. USA*, **85**, 6419–6423.

**Tanksley, S.D., Ganal, M.W., Prince, J.P. *et al.*** (1992) High density molecular linkage maps of the tomato and potato genomes. *Genetics*, **132**, 1141–1160.

**Tesler, G.** (2002) GRIMM: genome rearrangements web server. *Bioinformatics*, **18**, 492–493.

**Thorup, T.A., Tanyolac, B., Livingstone, K.D., Popovsky, S., Paran, I. and Jahn, M.** (2000) Candidate gene analysis of organ pigmentation loci in the *Solanaceae*. *Proc. Natl Acad. Sci. USA*, **97**, 11192–11197.

**Van der Hoeven, R., Ronning, C., Giovannoni, J., Martin, G. and Tanksley, S.** (2002) Deductions about the number, organization and evolution of genes in the tomato genome based on analysis of a large expressed sequence tag collection and selective genomic sequencing. *Plant Cell*, **14**, 1441–1456.

**Van der Knaap, E. and Tanksley, S.D.** (2003) The making of a bell pepper-shaped tomato fruit: identification of loci controlling fruit morphology in Yellow Stuffer tomato. *Theor. Appl. Genet.* **107**, 139–147.

**Wang, Y., Diehl, A., Wu, F., Vrebalov, J., Giovannoni, J., Siepel, A. and Tanksley, S.D.** (2008) Sequencing and comparative analysis of a conserved syntenic segment in the *Solanaceae*. *Genetics*, **180**, 391–408.

**Wu, F., Mueller, L.A., Crouzillat, D., Pétiard, V. and Tanksley, S.D.** (2006) Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: a test case in the Euasterid plant clade. *Genetics*, **174**, 1407–1420.

**Xu, Z. and Wang, H.** (2007) LTR-FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265–W268.

**Yang, H.-B., Liu, W.Y., Kang, W. -H., Jahn, M. and Kang, B. -C.** (2009) Development of SNP markers linked to the *L* locus in *Capsicum* spp. by a comparative genetic analysis. *Mol. Breed.* **24**, 433–446.

**Zhang, J., Yu, C., Pulletikurti, V., Lamb, J., Danilova, T., Weber, D.F., Birchler, J. and Peterson, T.** (2009) Alternative Ac/Ds transposition induces major chromosomal rearrangements in maize. *Genes Dev.* **23**, 755–765.

**Zhong, X.B., de Jong, J.H. and Zabel, P.** (1996b) Preparation of tomato meiotic pachytene and mitotic metaphase chromosomes suitable for fluorescence *in situ* hybridization (FISH). *Chromosome Res.* **4**, 24–28.

**Ziolkovski, P.A., Blanc, G. and Sadowski, J.** (2003) Structural divergence of chromosomal segments that arose from successive duplication events in the *Arabidopsis* genome. *Nucleic Acids Res.* **31**, 1339–1350.

© 2012 The Authors