Theor Appl Genet (2010) 120:809–827
DOI 10.1007/s00122-009-1206-z

ORIGINAL PAPER

COSII genetic maps of two diploid *Nicotiana* species provide a detailed picture of synteny with tomato and insights into chromosome evolution in tetraploid *N. tabacum*

Feinan Wu · Nancy T. Eannetta · Yimin Xu · Jörg Plieske · Martin Ganal · Carlo Pozzi · Nicolas Bakaher · Steven D. Tanksley

Received: 30 June 2009 / Accepted: 17 October 2009 / Published online: 17 November 2009

Abstract Using single-copy conserved ortholog set (COSII) and simple sequence repeat (SSR) markers, we have constructed two genetic maps for diploid *Nicotiana* species, *N. tomentosiformis* and *N. acuminata*, respectively. *N. acuminata* is phylogenetically closer to *N. sylvestris* than to *N. tomentosiformis*, the latter two of which are thought to contribute the S-genome and T-genome, respectively, to the allotetraploid tobacco (*N. tabacum* L., $2n = 48$). A comparison of the two maps revealed a minimum of seven inversions and one translocation subsequent to the divergence of these two diploid species. Further, comparing the diploid maps with a dense tobacco map revealed that the tobacco genome experienced chromosomal rearrangements more frequently than its diploid relatives, supporting the notion of accelerated genome evolution in allotetraploids. Mapped COSII markers permitted the investigation of *Nicotiana*–tomato syntenic relationships. A minimum of 3 (and up to 10) inversions and 11 reciprocal translocations differentiate the tomato genome from that of the last common ancestor of *N. tomentosiformis* and *N. acuminata*. Nevertheless, the marker/gene order is well preserved in 25 conserved syntenic segments. Molecular dating based on COSII sequences suggested that tobacco was formed 1.0MYA or later. In conclusion, these COSII and SSR markers link the cultivated tobacco map to those of wild diploid *Nicotiana* species and tomato, thus providing a platform for cross-reference of genetic and genomic information among them as well as other solanaceous species including potato, eggplant, pepper and the closely allied coffee (Rubiaceae). Therefore they will facilitate genetic research in the genus *Nicotiana*.

Communicated by P. Heslop-Harrison.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-009-1206-z) contains supplementary material, which is available to authorized users.

F. Wu · N. T. Eannetta · Y. Xu · S. D. Tanksley (✉)
Department of Plant Breeding and Genetics,
Cornell University, Ithaca, NY 14853, USA
e-mail: sdt4@cornell.edu

F. Wu
e-mail: fw38@cornell.edu

S. D. Tanksley
Department of Plant Biology, Cornell University,
Ithaca, NY 14853, USA

J. Plieske · M. Ganal
TraitGenetics GmbH, Am Schwabeplan 1b,
06466 Gatersleben, Germany

C. Pozzi · N. Bakaher
Research and Development, Philip Morris International,
Quai Jeanrenaud 5, 2000 Neuchâtel, Switzerland

Introduction

Extensive comparative mapping studies have been performed in the family Solanaceae in the past three decades, including tomato, potato, eggplant and pepper (Doganlar et al. 2002a; Livingstone et al. 1999; Tanksley et al. 1992). These works have depicted syntenic relationships among these solanaceous species and provided insights into genome evolution in this large plant family, and practically, have allowed comparative analysis of qualitative and quantitative traits among them (Doganlar et al. 2002b). Recently, Wu et al. (2006) developed a set of single-copy conserved orthologous markers (COSII) and designed universal primers that could amplify orthologous fragments from solanaceous species and the close relative coffee (Rubiaceae). These COSII markers have been successfully

used to study syntenic relationships among tomato, eggplant and pepper (Wu et al. 2009a, b), and in this work we further demonstrated their utility in a more distant genus *Nicotiana* (vs. tomato and eggplant in *Solanum* and pepper in *Capsicum*).

The genus *Nicotiana* in the plant family Solanaceae comprises 70 naturally occurring species (Lewis and Nicholson 2007). Tobacco (*N. tabacum* L.) is an allotetraploid ($2n = 48$) and considered to be derived from a tetraploidization event subsequent to an interspecific hybridization between the diploid species *N. sylvestris* (S-genome, $2n = 24$) and a diploid member of section *Tomentosae* (T-genome, $2n = 24$), possibly *N. tomentosiformis* or *N. otophora* (Kenton et al. 1993; Lim et al. 2004). The genus *Nicotiana* has long been considered as a model system in evolution, ecology and biotechnology (Chaleff and Ray 1984; Clemente 2006; Daniell et al. 2001; Goodin et al. 2008; Kessler and Halitschke 2007; Kessler et al. 2008; Okamuro and Goldberg 1985). In the past, there have been efforts to develop molecular genetic maps of *Nicotiana* species (Bogani et al. 1997; Julio et al. 2006; Lin et al. 2001; Nishi et al. 2003; Ren and Timko 2001; Rossi et al. 2001). Recently, a well-covered simple sequence repeat (SSR) map has been reported for *N. tabacum* (Bindler et al. 2007). However, very little comparative mapping work has been done for this genus (Lin et al. 2001), and no clear syntenic relationships have been established between *Nicotiana* species.

It would be ideal to directly compare genetic maps of the tobacco and tomato genomes. However, genetic mapping is complicated in tetraploid tobacco, especially using orthologous markers. To overcome this problem, we attempted to develop two diploid mapping populations each representing one of the two putative diploid ancestral genomes—*N. tomentosiformis* and *N. sylvestris*. A set of COSII markers mapped on tomato (Wu et al. 2006) and SSR markers on tobacco (Bindler et al. 2007) were applied in both diploid mapping populations. These two diploid maps have allowed us not only to compare tobacco and tomato genomes through the bridge of diploid *Nicotiana* maps, but also to investigate genome evolution in tetraploid tobacco and its diploid relatives. Furthermore, orthologous sequences amplified by COSII universal primers also provided an opportunity to estimate the divergence time between tobacco and its diploid progenitors, i.e., when the tetraploidization occurred. In conclusion, we expect these tobacco–*Nicotiana*–tomato comparative maps will link tobacco and tomato, facilitate sharing of their genomic and genetic information including genome sequences (http://www.sgn.cornell.edu/about/tomato_sequencing.pl; http://tobaccogenome.org), allow further development of molecular breeding in tobacco, and advance studies of genome evolution and comparative genomics in the family Solanaceae.

Materials and methods

Genetic mapping in two *Nicotiana* genomes

To analyze the T-genome, we developed a mapping population of 55 interspecific F2 plants from the cross *N. tomentosiformis* TA3385 × *N. otophora* TA3353. For the S-genome, *N. sylvestris* is the only species in section *Sylvestres*. We surveyed five accessions from USDA germplasm collection and the Botanical and Experiment Garden, Radboud University, Nijmegen, The Netherlands (http://www.bgard.science.ru.nl/), and observed extremely low polymorphism between these accessions. Only 1 out of 24 tested COSII markers contained cleaved amplified polymorphic sequence (CAPS) (data not shown). With such a low level of polymorphism, it was impractical to develop a COSII map within *N. sylvestris*. Based on phylogeny (Knapp et al. 2004), species in the closest section *Alatae* all have a different chromosome number ($2n = 18$ vs. $2n = 24$ in *N. sylvestris*) and would make it difficult to develop fertile F1 hybrids; therefore, we searched for crossable species in the more distant section *Petunioides*. We tried interspecific crosses between *N. sylvestris* and members in section *Petunioides* as well as within section *Petunioides*; however, the vast majority of them did not produce a suitable mapping population due to one of the following reasons: failure to germinate F1 seedlings, or sterile F1 plants, or a low polymorphism level between parental plants. In the end, we developed a mapping population of 51 intraspecific F2 plants from *N. acuminata* TA3460 × *N. acuminata* var. multiflora TA3461. *N. sylvestris* and *N. acuminata* are phylogenetically closer to each other than to *N. tomentosiformis* that is basal in phylogeny of the genus *Nicotiana* (Knapp et al. 2004).

The two diploid *Nicotiana* maps were designated as Tmf (*N. tomentosiformis*) and Acn (*N. acuminata*). A set of COSII markers and a small number of tomato-derived markers were selected at intervals of approximately 10 cM throughout the tomato genetic map (93% coverage of the tomato map for the Tmf map and 55% for the Acn map, calculated by excluding gaps of >20 cM or gaps on either end of a chromosome). The lower coverage of the Acn map is mainly due to its lower polymorphism level. Universal primers for the COSII markers, based on sequence alignments of orthologs from multiple solanaceous species (Wu et al. 2006), were used to amplify orthologous fragments from the mapping parents. If the COSII primers used for tomato did not provide suitable polymorphic fragments for mapping in *Nicotiana*, a second primer pair was designed in a different region of the same sequence alignment using the method described in Wu et al. (2006). Amplicon size differences between the two parents of a mapping population were used to genotype the mapping population

directly. Otherwise, the single fragment amplicons were purified and sequenced. Sequences of the two parents were then aligned and examined for CAPS (Konieczny and Ausubel 1993) using the program CAPSdesigner (http://www.sgn.cornell.edu/tools/caps_designer/caps_input.pl). Thereafter, the mapping population was genotyped via CAPS assays. In the cases where CAPS assays were not feasible, other single nucleotide polymorphisms (SNPs) were exploited to design derived CAPS (dCAPS) assays using the program dCAPS (Neff et al. 1998, 2002).

Tomato-derived single-copy markers were used for mapping only when COSII markers were not available. Primers were designed based on the tomato cDNA sequences and used to amplify orthologous fragments from the *Nicotiana* parents and the same method as for COSII markers was used to genotype these markers in the *Nicotiana* mapping populations. Information of all COSII and tomato-derived markers is available in supplementary Tables S1 and S2, and ftp://ftp.sgn.cornell.edu/COSII/Nicotiana.

A set of SSR markers, some of which were developed from tobacco genomic sequences and used to construct a published tobacco map (Bindler et al. 2007) and others from unpublished maps (developed by Philip Morris International, Switzerland), were also genotyped in these two populations using the protocol described in Bindler et al. (2007). Subsequently, the three sets of markers were combined to construct a genetic map for each population using the Mapmaker software (Lander et al. 1987). We first estimated linkage groups using the command "group 3, 0.2", and then established a framework (markers ordered at LOD $\geq$ 3) for each group using the commands "compare" and "ripple". Subsequently, genetic distances were calculated using framework markers and the Kosambi mapping function (Kosambi 1944). Finally, we positioned additional markers in the intervals between framework markers using the commands "try" and "ripple". Interval markers were assumed to be evenly distributed between the flanking framework markers for calculation of their approximate mapping positions.

Genetic mapping in tomato

The tomato map, used in the comparisons to the *Nicotiana* maps, has been previously generated based on an interspecific cross *S. lycopersicum* × *S. pennellii* (Frary et al. 2005; Fulton et al. 2002). Currently, more than 2,500 markers have been mapped in this population of which 877 are COSII markers (Wu et al. 2006). For the purpose of comparison with *Nicotiana*, we prepared a modified tomato genetic map for which the framework was based predominantly on the COSII markers. The complete tomato map is available at Solanaceae Genomics Network (http://www.sgn.cornell.edu/cview/map.pl?map_id=9&show_offsets=1&show_ruler=1) and bulk download of all COSII marker information is available at SGN FTP site (ftp://ftp.sgn.cornell.edu/COSII/).

Assessment of *Nicotiana*–tomato synteny

The degree of synteny between *Nicotiana* and tomato was assessed using synteny marker pairs (SMPs) and conserved syntenic segments (CSSs) as in previous comparative studies in Solanaceae (Wu et al. 2009a, b). A SMP is defined as a pair of synteny markers that are adjacent to each other (regardless of other non-synteny markers) on both maps. To minimize erroneous results, we searched for SMPs only within the subset of synteny markers that had been mapped and ordered in both genomes with LOD $\geq$ 2 (a relatively high confidence given the size of the two mapping populations and the number of markers). We first identified SMPs between the Tmf and tomato maps and between the Acn and tomato maps and then combined the results and removed conflicts. How we combined SMPs is described in the following example. Markers a, b, c and d are in order on the tomato map, markers a and c are adjacent on the Tmf map and thus a SMP and so are markers b and d on the Acn map. In such a case, we combined the two SMPs and re-divided them into three consecutive SMPs, a–b, b–c and c–d. An example to remove conflicts was as follows. Suppose a–b is a SMP between the Tmf and tomato maps, a and b are mapped to different linkage groups on the Acn map (regardless of mapping confidence). In this case, we removed a–b from the SMP list. Subsequently, the resultant SMPs were coalesced into CSSs defined as shared blocks of genes/markers with preserved order between genomes. In this way, the identified CSSs are conserved among all three genomes. In this analysis, markers ordered at LOD $<$ 2 were utilized only in reference to interchromosomal translocation or considered as a single gene transposition.

Molecular dating of the tobacco speciation

In a project of *Nicotiana* phylogeny (Wu and Tanksley, unpublished results), we identified six markers—C2At1g14850, C2At1g30540, C2At1g67700, C2At1g74470, C2A3g15380 and C2At4g16580, each of which amplified two distinct sequences from *N. tabacum*. We were able to ascertain that each pair of the tobacco amplicons (amplified by the primer set of a COSII marker) corresponds to the homeologs in the two distinct genomes of the allotetraploid tobacco. For these markers, we obtained a single amplified sequence from *N. sylvestris* TA3347 and the four mapping parents in this work—*N. tomentosiformis* TA3385, *N. otophora* TA3353, *N. acuminata* TA3460 and *N. acuminata* var. multiflora TA3461. These data were subsequently used to estimate an

upper limit of the time when tetraploid tobacco was formed via interspecific hybridization (between the ancestral diploid progenitors) followed by tetraploidization. Introns in these diploid *Nicotiana* amplicons were identified by aligning them with the corresponding tomato ortholog with known intron positions (Wu et al. 2006). Then, we removed the intron sequences and concatenated the remaining exon sequences of the six markers. We used MODELTEST (Posada and Crandall 1998) to choose the most suitable DNA substitution model via likelihood ratio test between null model (i.e., equal base frequency, equal transition rates and equal transversion rates, rates equal among sites and no invariable sites) and each of the other 55 complex models. Finally, we reconstructed a phylogenetic tree using maximum likelihood (ML) method in PAUP*4.0 (Swofford 2003) and the above DNA substitution model, based on the overlapping region in the sequence alignment. The divergence time between T-genome/S-genome allele and its closest diploid *Nicotiana* relative was estimated using non-parametric rate smoothing (NPRS) method, which did not rely on the assumption of a molecular clock, in the program r8s (Sanderson 1997). The calibration point was 27.7MYA for the tomato–*Nicotiana* split (Wu and Tanksley, unpublished results). Based on the Solanaceae phylogeny (Olmstead et al. 1999), this estimate is consistent with the recent estimates of tomato–petunia divergence 31.2MYA and tomato–pepper divergence 19.1MYA (Wang et al. 2008). Sequence alignments are available at ftp://ftp.sgn.cornell.edu/COSII/Nicotiana.

Results

COSII marker polymorphism

Over 400 COSII markers were tested in the two Tmf mapping parents—*N. tomentosiformis* and *N. otophora*. In 34 cases, the two parents had a detectable amplicon size difference (>30 bp) on agarose gels (supplementary Table S1); in 228 other cases, it was feasible to design CAPS or dCAPS assays based on SNPs detected in the amplicon sequences of the two parents. A subset of 186 COSII markers in the latter category, which had a minimum of 200 bp sequenced exon and/or intron, was subjected to further sequence analysis (supplementary Table S3; sequence alignments available at ftp://ftp.sgn.cornell.edu/COSII/Nicotiana). The intron positions of the COSII markers had been predicted previously based on comparison with the Arabidopsis orthologs (Wu et al. 2006). The analysis of these amplicon sequences (154 introns and 49 exons) further confirmed the conserved intron positions between the family Solanaceae and Arabidopsis. Not surprisingly, the average SNP frequency was significantly higher in introns (58 bp/SNP) than exons (103 bp/SNP). INDELs were identified in 90 introns (58%) but only 2 exons (4%, supplementary Table S3). In addition, the 20 most polymorphic markers were distributed across the entire map except for Tmf9 (supplementary Table S3).

In the Acn population, we observed a similar situation. About 300 COSII markers were tested in the two mapping parents—*N. acuminata* and *N. acuminata* var. multiflora. In 14 cases, the two parents had a detectable amplicon size difference (>30 bp) on agarose gels (supplementary Table S2); in 119 other cases, it was feasible to design CAPS or dCAPS assays based on the SNPs detected in the amplicon sequences of the two parents. A subset of 103 COSII markers in the latter category, which had a minimum of 200 bp sequenced exon and/or intron, was subjected to further analysis (supplementary Table S4; sequence alignments available at ftp://ftp.sgn.cornell.edu/COSII/Nicotiana). The analysis of these amplicon sequences (88 introns and 20 exons) also confirmed the conserved intron positions between the family Solanaceae and Arabidopsis. Similar to the Tmf population, the average SNP frequency was significantly higher in introns (75 bp/SNP) than exons (115 bp/SNP). INDELs were identified in 52 introns (59%) but none of the exons (supplementary Table S4). The 20 most polymorphic markers (3 in common with those of the Tmf population) were distributed across the entire map except for Acn4 and Acn11 (supplementary Table S4); therefore, the combined set of 37 markers from the Tmf and Acn populations could potentially be useful for assessing genetic diversity in *Nicotiana*.

Construction of diploid *Nicotiana* genetic maps

Of the COSII markers surveyed in the Tmf population, 262 produced usable polymorphisms (including amplicon size difference, CAPS and dCAPS) for genetic mapping. This marker set, together with six tomato-derived markers, provided a good coverage (93%, see "Materials and methods") of the tomato genetic map and hereafter will be referred to as Tmf synteny markers. Subsequently, Tmf synteny markers were combined with 221 tobacco SSR markers to generate a relatively dense genetic map, comprising 489 markers and 12 linkage groups (named Tmf1–12 based on synteny with tomato chromosomes T1–12; Fig. 1). These linkage groups ranged in map distance from 59 cM (Tmf2) to 112 cM (Tmf3), and the entire genetic map totals 1,071 cM with an average density of one framework marker every 4.1 and 2.2 cM for all markers (Table 1).

Of all the COSII markers surveyed in the Acn population, 133 produced usable polymorphisms (including amplicon size differences, CAPS and dCAPS) for genetic mapping. This marker set, together with one tomato-derived marker, covered 55% (lower coverage than the

Tmf map due to the lower polymorphism lever and thus the smaller number of markers) of the tomato genetic map and hereafter will be referred to as Acn synteny markers. Subsequently, Acn synteny markers were combined with 174 tobacco SSR markers to generate a genetic map. The Acn map is comprised of 308 markers and 12 linkage groups (named Acn1–12 based on synteny with tomato chromosomes T1–12; Fig. 1). These linkage groups ranged in map distance from 65 cM (Acn9) to 106 cM (Acn2), and the entire genetic map totals 1,033 cM with an average density of one framework marker every 5.0 and 3.4 cM for all markers (Table 1).

In general, the Tmf linkage groups were longer than the orthologous Acn ones, which may be attributed to its greater tomato genome coverage. However, the entire map length was not significantly different between the two maps due to the fact that the Tmf map had a much shorter map distance in Tmf2 (47 cM shorter) and Tmf8 (27 cM shorter). Tmf2 was shorter due to recombination reduction (possibly caused by an inversion between the two mapping parents) while Tmf8, according to comparison with Acn8, may be extended by mapping more markers.

Genome evolution within the genus *Nicotiana*

The following pairwise comparisons were performed in order to study genome evolution within the genus *Nicotiana*. First, we compared the two diploid maps in order to identify putative chromosomal rearrangements since the divergence of these diploid species. This analysis relied on the common COSII markers since there were only four SSR markers shared between the two maps. Secondly, we compared the Tmf and Acn maps with a dense tobacco map of 24 linkage groups Tob1–24 (Bindler et al. 2007; unpublished results) via common SSR markers. This analysis allowed the identification of two sets of homeologous linkage groups—12 for T-genome (the 12 chromosomes corresponding to the Tmf map) and 12 for S-genome (the other 12 chromosomes corresponding to the Acn map). If a rearrangement existed between the Tmf and the Tob maps but not between Tmf and Acn in the same map region, we considered such rearrangement likely occurred in the tobacco T-genome after the tetraploidization event. If a rearrangement existed between the Acn and the Tob maps but not between Tmf and Acn in the same map region, we considered such rearrangement occurred either in the tobacco S-genome after the tetraploidization or in the lineage leading to *N. sylvestris* before the tetraploidization. Rearrangements observed between the Tmf and Acn maps were further analyzed using tomato as an outgroup and will be discussed in next section.

For the Tmf–Acn comparison, we declared a disruption in synteny between the two genomes when the following two criteria were both met. First, a structural difference was inferred only if two or more linked markers in at least one genome confirmed the rearrangement (an inversion involving one synteny marker at the end of a linkage group and another within the linkage group was also accepted); second, for inversions, the involved markers should be ordered at LOD $\geq$ 2 on both maps (a relatively high confidence given the population sizes of 55 and 51, respectively). This method is less likely to declare false-positive rearrangements, although we cannot rule out the possibility of some rearrangements not being detected or other falsely identified rearrangements. For Tmf–Tob and Acn–Tob comparisons, we reported the structural difference where at least two markers were mapped on the Tmf/Acn map at LOD $\geq$ 2. In the following paragraphs, we will describe these comparisons chromosome by chromosome (Fig. 2; supplementary Fig. S1).

Tmf1/Tob19 versus Acn1/Tob7

Tmf1 differs from Acn1 by one major inversion corresponding to about half of the two linkage groups. Acn1 may potentially be extended from the lower end by mapping the markers from the last 30 cM of Tmf1. Tob19 corresponds to Tmf1 but possibly experienced a small inversion (PT50028 and PT50645). On the other hand, Tob7 is orthologous to Acn1 and the gene order has been preserved since their divergence.

Tmf2/Tob17 versus Acn2/Tob3

Tmf2 has an inversion relative to Acn2. Tob17 shares the same gene order with Tmf2. Tob3 is orthologous to Acn2 with one inversion that occurred either in Tob3 or its orthologous chromosome in the *N. sylvestris* genome prior to the tobacco speciation.

Tmf3/Tob4 versus Acn3/Tob6

Tmf3 and Acn3 differ by two inversions. The upper inversion likely occurred in the Tmf genome because in the same map region Tmf3 also has an inversion relative to Tob4 (both involve the same loci—co-segregating PT61454 and C2At1g74470 and co-segregating PT60273 and C2At3g48610). Similarly, the lower inversion (between markers C2At3g14200 and C2At3g18270) likely occurred in the Acn genome because Acn3 has an inversion relative to Tob6 in the same map region (between markers PT50200 and PT54982). Besides the above two inversions, Tob6 (between markers PT50200 and PT50885) or its orthologous chromosome in the *N. sylvestris* genome (prior to the tobacco speciation) possibly experienced an inversion.

(Tmf4 + Tmf8)/(Tob12 + Tob23) versus (Acn4/Tob16 + Acn8/Tob1)

Tmf4 and Acn4 differ by one inversion and Tmf8 and Acn8 have the same gene order. Tmf4 combines the majority of Tob12 and a small segment of Tob23, while the rest of Tob12 and Tob23 form Tmf8, indicating a reciprocal translocation in the tobacco T-genome subsequent to polyploidization. Two inversions differentiate Tob16 and Acn4 and both likely occurred in the tobacco S-genome or the *N. sylvestris* genome prior to the tetraploidization event. Tob1 matches Acn8 and both have the same gene order.

Tmf5/Tob13 versus Acn5/Tob11

This chromosome appears to be mostly conserved during evolution of these species. Except for a possible small inversion between Tmf5 and Acn5, gene order and gene content are in good agreement among Tob13, Tmf5, Acn5 and Tob11.

Tmf6/(Tob21 + Tob22) versus Acn6/Tob8

Gene order and gene content are well preserved between Tmf6 and Acn6. Tmf6 combines Tob21 and Tob22, suggesting a putative chromosome breakage event in the tobacco genome. Tob22 differs from Tmf6 by at least two inversions. Tob8 is orthologous to Acn6 yet with one putative inversion in the tobacco S-genome or the *N. sylvestris* genome prior to the tetraploidization event.

Tmf7/Tob2 versus Acn7/Tob10

This chromosome is well conserved without inversions or translocations.

Tmf9/Tob14 versus Acn9/Tob7

Gene order is conserved in this chromosome too. It is worth mentioning that both Acn1 and Acn9 are orthologous to Tob7. Given the opposite event described earlier—Tob21 and Tob22 correspond to Tmf6, a chromosome fusion event may have occurred in the tobacco S-genome. The fusion is not likely to have occurred in the *N. sylvestris* genome since it will change the chromosome number. Thus, the current 48 chromosomes in the tobacco genome are possibly not a simple combination of the two diploid progenitors, but a consequence of a few chromosome fusion and breakage events.

(Tmf10/Tob24 + Tmf12/Tob15) versus (Acn10/Tob5 + Acn12/Tob20)

These two chromosomes apparently experienced a reciprocal translocation, which occurred prior to the tobacco speciation in the lineage leading to either *N. tomentosiformis*/*N. otophora* or the last common ancestor of *N. acuminata* and *N. sylvestris* genomes. Further, Tmf10 and Acn10 differ by one inversion. Gene content and gene order are generally preserved between tobacco and its diploid orthologous counterparts except for an inversion in Tob24 with regard to Tmf10.

Fig. 1 Two diploid *Nicotiana* genetic maps based on *N. tomentosiformis* (the Tmf map) and *N. acuminata* (the Acn map). The *Nicotiana* linkage groups are designated as Tmf1–12 and Acn1–12 based on synteny with the tomato chromosomes T1–12 (Frary et al. 2005; Fulton et al. 2002). COSII marker names are shortened, e.g., 2g01110 is for COSII marker C2At2g01110 (Wu et al. 2006); SSR marker names each have a prefix "PT" (Bindler et al. 2007; unpublished results); others are tomato-derived markers. Markers in *bold* are framework markers (LOD $\geq$ 3); markers in *bold* and *italic* are internal markers with 2 $\leq$ LOD <3; others are interval markers with LOD < 2; co-segregating markers are denoted by a *vertical bar* beside the marker names. A *tick mark* connects a framework marker to its actual map position or an interval marker to its estimated map position. Markers mapped beyond either of the end framework markers are not connected by *tick marks*. "~Tx" following the name of a marker indicates its chromosome location on the tomato map where "x" specifies the linkage group; "~Tobx" following the name of a marker indicates its chromosome location on the *N. tabacum* map

Tmf11/Tob09 versus Acn11/Tob18

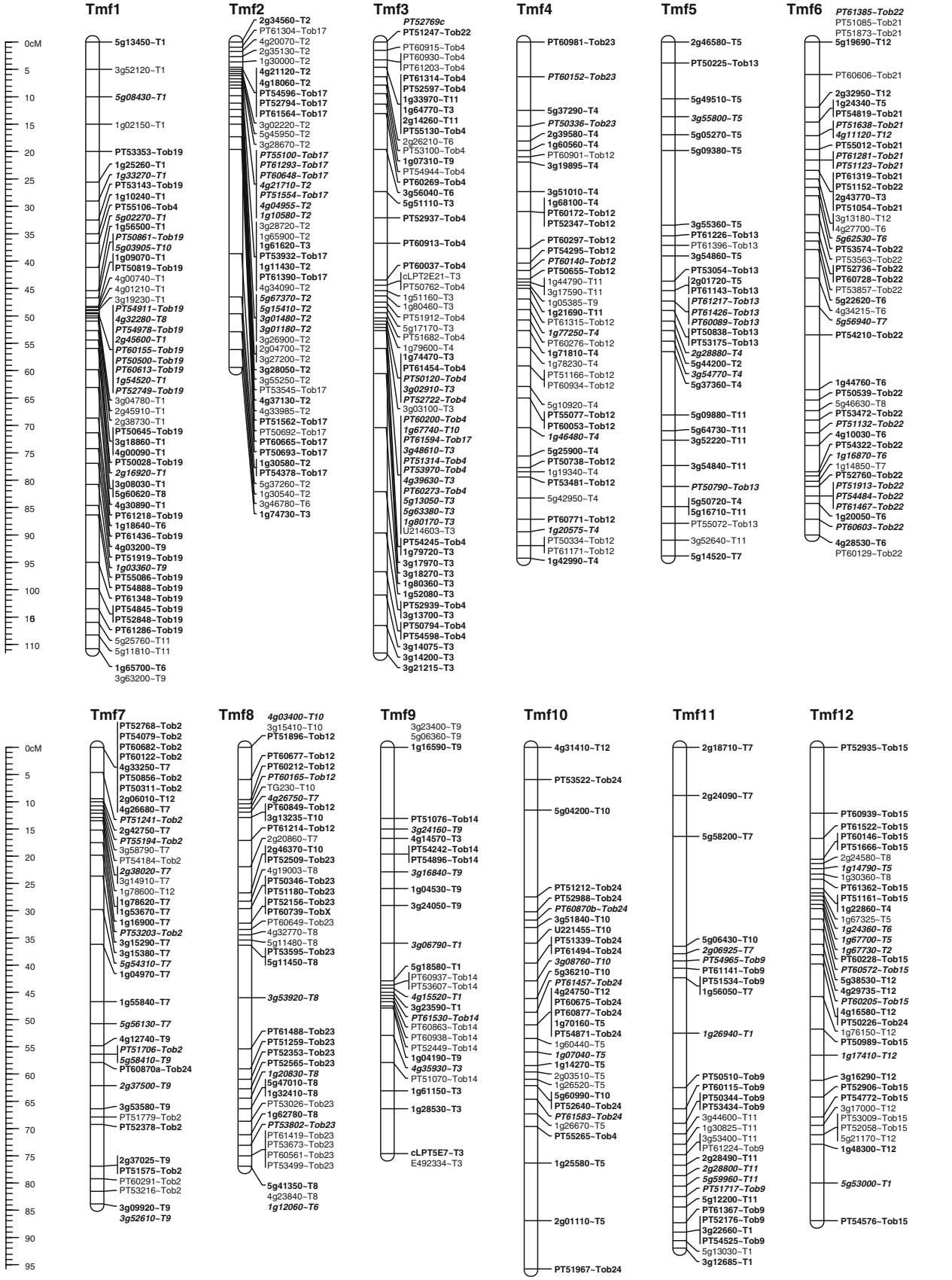
Gene order and gene content are well preserved in this chromosome among the diploid species and the tetraploid homeologs.

Syntenic relationships between *Nicotiana* and tomato genomes

The syntenic relationships of tomato and the two diploid *Nicotiana* genomes were deduced based on 268 Tmf synteny markers and 134 Acn synteny markers, and were depicted with a Tmf/Acn linkage group and its orthologous tomato chromosome(s) or segment(s) side by side (a schematic depiction in Fig. 3 and a detailed close-up picture in supplementary Fig. S2). We used the same criteria as that for Tmf–Acn comparison to declare a disruption in synteny. In the following paragraphs, we describe how the Tmf/Acn genome differs with respect to the tomato genome that is used as a standard of reference. For convenience, the results are presented in the order of *Nicotiana* linkage groups (supplementary Fig. S2).

Tmf1/Acn1 versus T1, T9 and T11

Tmf1/Acn1 combines three segments from T1, T9 and T11, which indicates at least two translocations between the *Nicotiana* and tomato genomes. By combining the


Tmf1
Tmf2
Tmf3
Tmf4
Tmf5
Tmf6
Tmf7
Tmf8
Tmf9
Tmf10
Tmf11
Tmf12
0cM

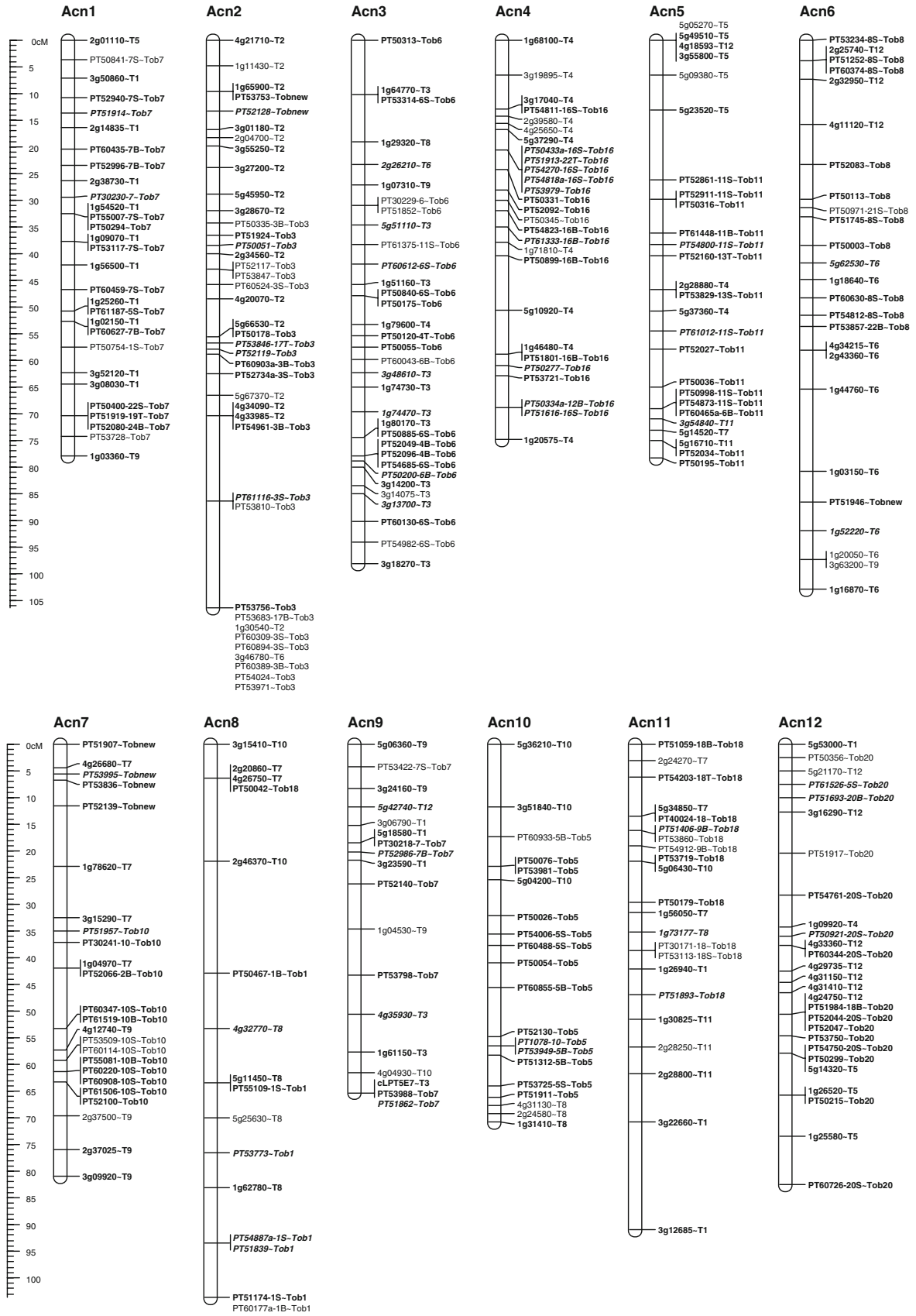


Fig. 1 continued

Table 1 Statistics of Tmf and Acn maps and their comparisons with tobacco and tomato maps

Tmf linkage group	Tmf1	Tmf2	Tmf3	Tmf4	Tmf5	Tmf6	Tmf7	Tmf8	Tmf9	Tmf10	Tmf11	Tmf12	Sum
Map distance (cM)	111	59	112	94	94	90	84	77	75	96	92	87	1,071
Number of framework markers	28	18	29	21	22	22	24	22	13	22	18	19	258
Number of markers	55	49	60	41	32	46	41	42	28	31	29	35	489
Number of COSII markers	33	33	29	20	20	18	24	18	16	16	17	18	262
Orthologous tobacco linkage group	Tob19	Tob17	Tob4	Tob12 + 23	Tob13	Tob21 + 22	Tob2	Tob12 + 23	Tob14	Tob24	Tob9	Tob15	–
Acn linkage group	Acn1	Acn2	Acn3	Acn4	Acn5	Acn6	Acn7	Acn8	Acn9	Acn10	Acn11	Acn12	Sum
Map distance (cM)	78	106	98	75	78	103	81	104	65	71	91	83	1,033
Number of framework markers	24	19	20	13	22	20	19	10	10	15	13	21	206
Number of markers	29	40	33	27	27	25	24	16	18	20	22	27	308
Number of COSII markers	12	17	15	10	11	13	8	8	10	6	11	12	133
Orthologous tobacco linkage group	Tob7	Tob3	Tob6	Tob16	Tob11	Tob8	Tob10	Tob1	Tob7	Tob5	Tob18	Tob20	–
Number of SMPs	5	8	1	8	5	2	3	4	5	3	0	1	45
Number of CSSs	3	2	1	2	3	2	2	3	3	3	0	1	25
Min./max. length of CSS (cM)[a]	1/26	1/6	6/6	31/45	13/19	8/13	8/19	5/19	6/16	12/29	0/0	21/21	1/45
Mean length of CSS (cM)	17	4	6	38	16	11	14	12	12	20	0	21	15

SMP synteny marker pair, *CSS* conserved syntenic segment

[a] cM length is based on the Tmf map

information from Tmf–Acn comparison, in the common ancestor of the Tmf and Acn genomes, this chromosome has a similar structure to Acn and differs from T1 by one inversion, and then Tmf1 had an additional inversion after it diverged from its last common ancestor with Acn1. No inversions were found between Tmf1/Acn1 and the other two tomato segments from T9 and T11. Besides translocations and inversions, six markers from T5, T6, T8 and T10 mapped to Tmf1/Acn1, which were possibly due to single gene transpositions or other mechanisms (see "Discussion").

Tmf2/Acn2 versus T2

This chromosome is conserved in gene content with no evidence of a translocation event differentiating *Nicotiana* and tomato genomes. Tmf2 has a significant recombination reduction, which caused the vast majority of markers to cluster in a 15 cM region. Marker orders appear to be conserved between Tmf2 and T2, while Acn2 differs from T2 and Tmf2 by one inversion; therefore, T2 and Tmf2 represent an ancestral condition and the inversion is derived in Acn2. C2At1g61620 and C2At1g74730 from T3 mapped to Tmf2 and C2At3g46780 from T6 to Tmf2/Acn2.

Tmf3/Acn3 versus T3

Tmf3/Acn3 is orthologous to lower T3 and a segment of T6. Given the information from Tmf–Acn comparison, the upper inversion (within C2At3g48610 and C2At1g74470) between Tmf3 and T3 likely occurred in the Tmf lineage, while the lower inversion (within C2At3g14200 and C2At3g14075) may have occurred either in the Tmf lineage or in the common ancestor of Tmf and Acn genomes, which is uncertain due to the low LOD score of C2At3g14075 on the Acn map. The inversion of C2At3g18270 and C2At3g14200 between Acn3 and T3 is likely to have occurred in the Acn lineage. Several markers from T4, T6, T8, T9, T10 and T11 mapped to Tmf3 and/or Acn3.

Tmf4/Acn4 versus T4 and T11

Tmf4/Acn4 matches T4 except that a segment in the middle of T4 was replaced by a T11 segment, indicating an exchange between the two chromosomes. Gene content and gene order are fairly well conserved except for an inversion between Tmf4 and T4 which must have occurred in the Tmf lineage; therefore, T4 and Acn4 likely represent the ancestral condition. C2At1g05385 from T9 mapped to Tmf4.

Tmf5/Acn5 versus T5, T4 and T11

Tmf5/Acn5 combines the lower T5, a small segment of T4 and the upper T11. Tmf5 has an inversion relative to T5. Otherwise, gene order is conserved among the three genomes. This inversion cannot be timed due to co-segregation of the markers on the Acn5. The inversion between Tmf5 and Acn5 (see Tmf–Acn comparison) cannot be seen here because the two markers come from different tomato chromosomes—C2At5g16710 from T11 while C2At5g14520 from T7. C2At5g44200 from T2 mapped to Tmf5, C2At4g18593 from T12 to Acn5, and C2At5g14520 from T7 to both Tmf5 and Acn5.

Tmf6/Acn6 versus T12 and T6

Tmf6/Acn6 combines the lower T6 and the upper T12, although some markers from the T12 segment mapped to other *Nicotiana* linkage groups. Tmf6 has one small inversion near the end with respect to T6, but the same markers were not all mapped on Acn6 so that it was not possible to determine the timing of this event. Putative gene transpositions are quite common in this chromosome in that several markers from T3, T5, T7, T8 and T9 mapped to either Tmf6 or Acn6.

Tmf7/Acn7 versus T7 and T9

Tmf7/Acn7 combines the lower T7 and the upper T9. Both gene content and gene order are well conserved among the three genomes except for C2At2g06010 and C2At1g78600 from T12 mapped to Tmf7.

Tmf8/Acn8 versus T10 and T8

Tmf8/Acn8 combines upper T10 and lower T8. One inversion differentiates Tmf8 from T10 but the involved markers could not be mapped on the Acn map. Both Tmf8 and Acn8 differ from T8 by one inversion, and they are likely due to the same event although more markers are needed to confirm this conjecture. C2At4g26750 and C2At2g20860 from T7 mapped to Tmf8 and Acn8, and C2At1g12060 from T6 mapped to Tmf8.

Tmf9/Acn9 versus T9, T3 and T1

Tmf9/Acn9 combines the lower T9, the upper T3 and a small segment of T1. Tmf9 has an inversion relative to T9 but the involved markers could not be mapped on the Acn map. Otherwise, Tmf9/Acn9 shares the same gene order with upper T3 and the T1 segment. C2At4g04930 from T10 and C2At5g42740 from T12 mapped to Acn9.

Fig. 2 Comparative maps of diploid and tetraploid *Nicotiana* genomes (see close-up in supplementary Fig. S1). Nomenclature of the diploid *Nicotiana* linkage groups follows Fig. 1, and the tetraploid *N. tabacum* linkage groups are designated as Tob1–24 (Bindler et al. 2007). A number of *N. tabacum* linkage groups are placed in reversed order (indicated by a suffix "r", e.g., Tob19r) for a better depiction of synteny. Orthologous markers are connected by a *line*, in which a *broken line* indicates that either or both of the markers were mapped at LOD < 2 and thus were not used for deduction of an inversion

Tmf10/Acn10 and Tmf12/Acn12 versus T10, T5, T8 and T12

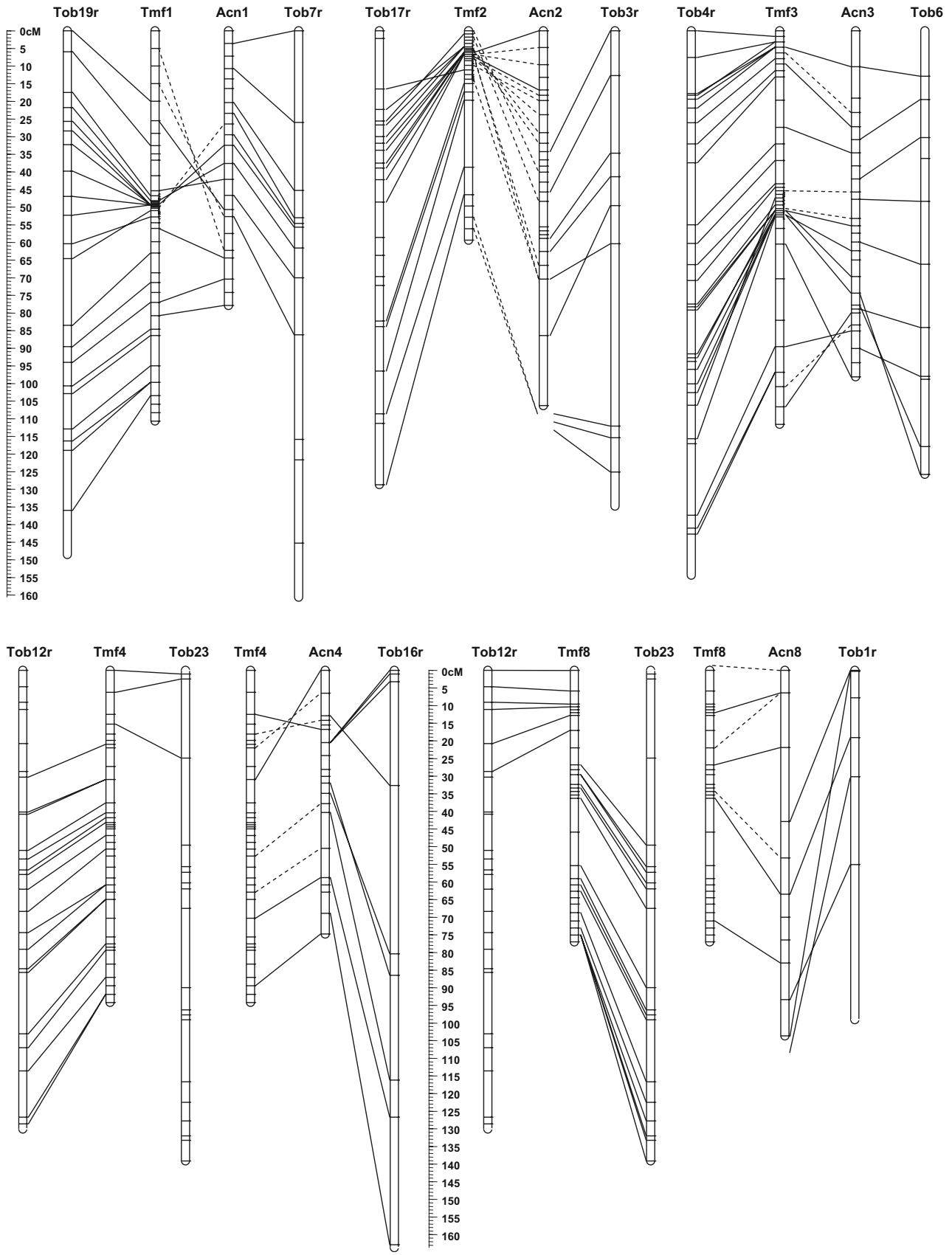
As discussed earlier in Tmf–Acn–Tob comparison, one translocation causes the exchange of part of *Nicotiana* linkage groups 10 and 12. As a result, the upper Tmf10 and upper Acn10 are orthologous to the majority of T10, the lower Tmf10 and lower Acn12 are orthologous to upper T5, the upper Tmf12 and lower Acn10 are orthologous to T8, and the lower Tmf12 and upper Acn12 are orthologous to lower T12. In these four orthologous regions, multiple putative inversions were observed. The inversion in upper Tmf10 has been discussed in the Tmf–Acn comparison. Both the lower Tmf10 and lower Acn12 differ from T5 by one inversion that is likely caused by the same event. Gene order is conserved between the upper Tmf12, lower Acn10 and T8. One inversion exists between lower Tmf12 and T12 but the markers could not be mapped on the Acn map. Putative gene transpositions are also quite frequent in these two chromosomes, which move markers from T12 to Tmf10, from T1, T2, T4, T5 and T6 to Tmf12, and from T1 and T4 to Acn12.

Tmf11/Acn11 versus T7, T11 and T1

Tmf11/Acn11 combines upper T7, lower T11 and a small segment of T1. At least one inversion differentiates Tmf11 and T11 but the markers could not be mapped on the Acn map. Otherwise, no inversions were found in these orthologous regions. C2At5g06430 from T10 was mapped on both Tmf11 and Acn11.

The use of synteny to predict the position of additional COSII markers on the *Nicotiana* maps

The detailed synteny between *Nicotiana* and tomato genomes permitted identification of the conserved genome regions among these species—45 SMPs coalesced into 25 CSSs (supplementary Fig. S2; see "Materials and methods" for the definition and method). The CSSs range in size from 1 to 45 cM with an average size of 15 cM (cM value based on the Tmf map, Table 1). They cover from 5% (Tmf3) to 81% (Tmf4) of the different Tmf linkage groups (except for Tmf11 with no CSSs identified) and totaled


Tob19r
Tmf1
Acn1
Tob7r
Tob17r
Tmf2
Acn2
Tob3r
Tob4r
Tmf3
Acn3
Tob6
0cM
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
135
140
145
150
155
160
Tob12r
Tmf4
Tob23
Tmf4
Acn4
Tob16r
Tob12r
Tmf8
Tob23
Tmf8
Acn8
Tob1r
0cM
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
135
140
145
150
155
160

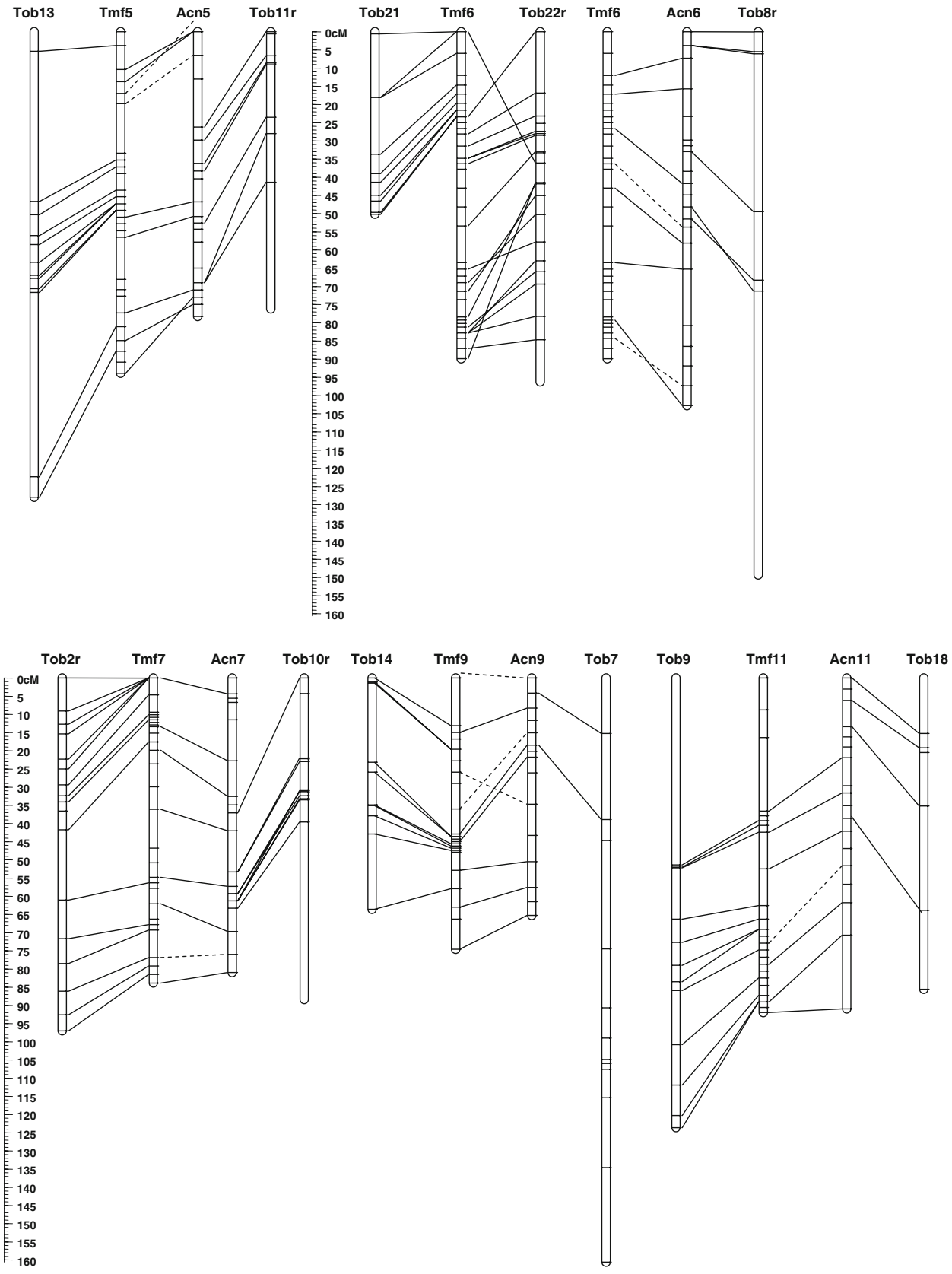


Fig. 2 continued

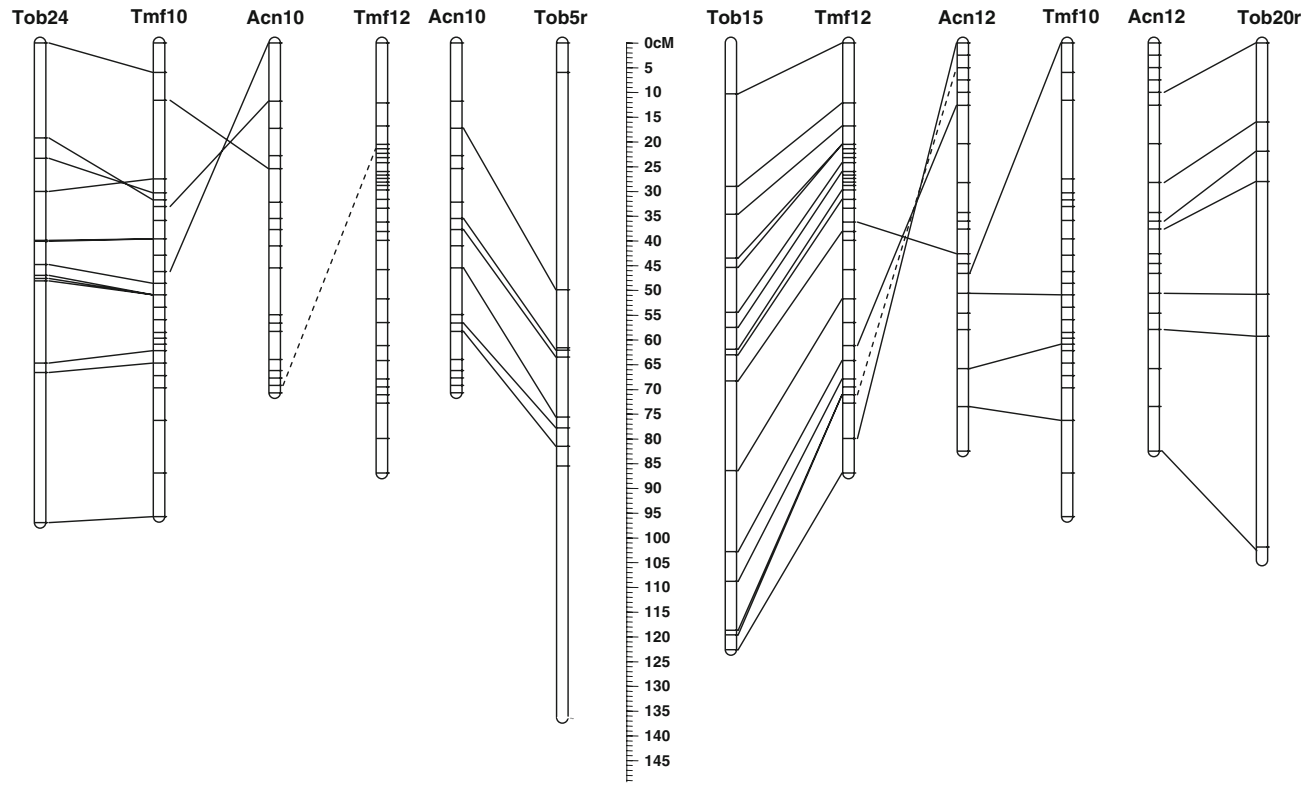


Fig. 2 continued

368 cM corresponding to 34% of the Tmf map. These SMPs and CSSs can be used to infer relatively precise map position of additional COSII markers on the *Nicotiana* maps, thereby facilitating mapping studies in *Nicotiana* and permitting comparisons of QTL studies between *Nicotiana* and tomato. This method has been proved successful in previous work on eggplant and pepper (Wu et al. 2009a, b). In *Nicotiana*, we applied the same method and displayed the results on the Tmf map (since this map has more COSII markers than the Acn map, supplementary Fig. S3). By searching the tomato COSII markers that are located within these SMPs, we inferred the position of an additional set of 143 COSII markers on the *Nicotiana* map, which were referred to as "inferred *Nicotiana* COSII markers".

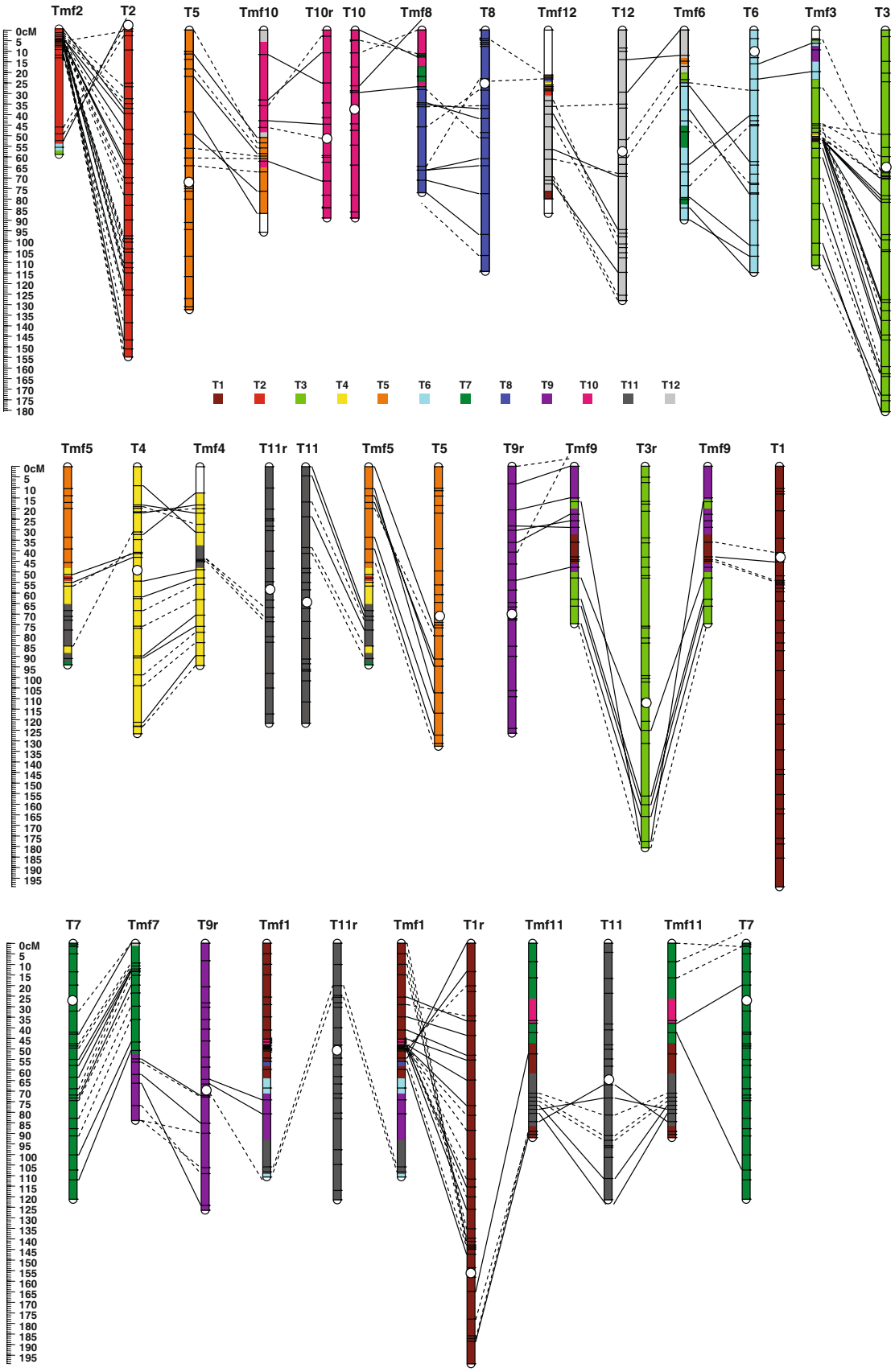
Discussion

Molecular dating of the tetraploidization event leading to tobacco

It has previously been proposed that *N. tomentosiformis* (or *N. otophora*) and *N. sylvestris* correspond to the two diploid species that led to the allotetraploid tobacco (Kenton et al. 1993; Lim et al. 2004). As described in "Materials and methods", we reconstructed a ML tree (Fig. 4a) using concatenated exons of six COSII markers, which depicted the phylogenetic relationships among the two homeologous copies from tetraploid tobacco, and the corresponding orthologs from the diploid species closely related to *N. tomentosiformis* (*N. tomentosiformis* and *N. otophora*) and *N. sylvestris* (*N. sylvestris*, *N. acuminata* and *N. acuminata* var. multiflora). Tomato was included as an outgroup. Based on this tree and the calibration point of tomato–*Nicotiana* split at 27.7MYA, the S-genome in tobacco likely diverged from *N. sylvestris* 1.0MYA while the T-genome diverged from *N. tomentosiformis* 1.7MYA. It is thus concluded that the hybridization and tetraploidization event leading to tobacco likely occurred $\leq$1.0MYA. A previous, independent estimate, based on plastid DNA, places the event at 0.2MYA (Clarkson et al. 2005).

Structural differences within the genus *Nicotiana* and between the genera *Nicotiana* and *Solanum* (tomato) and inferences concerning genome evolution

Comparisons of two diploid and one tetraploid *Nicotiana* maps permitted identification of likely rearrangements within the genus *Nicotiana*, and inference of the timing of most events (Fig. 4b). Since the hybridization and tetraploidization events, the tobacco genome has possibly


Tmf2 T2 T5 Tmf10 T10r T10 Tmf8 T8 Tmf12 T12 Tmf6 T6 Tmf3 T3
0cM 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180
T1 T2 T3 T4 T5 T6 T7 T8 T9 T10 T11 T12
Tmf5 T4 Tmf4 T11r T11 Tmf5 T5 T9r Tmf9 T3r Tmf9 T1
0cM 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195
T7 Tmf7 T9r Tmf1 T11r Tmf1 T1r Tmf11 T11 Tmf11 T7
0cM 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195

◀ **Fig. 3** Comparative maps of Tmf and tomato genomes (see a close-up picture including the Acn genome in supplementary Fig. S2). Nomenclature of linkage groups follows Fig. 1. Some tomato chromosomes are placed in reversed order (indicated by a suffix "r", e.g., T1r) for a better depiction of synteny. Each tomato chromosome is assigned one *color* (see *color codes*) and the corresponding Tmf linkage group segment(s) are painted with the *same color*. Approximate tomato centromeric locations are indicated by *white dots* based on the comparison with the map by Frary et al. (2005). Orthologous markers are connected by a *line*, in which a *broken line* indicates that either or both of the markers were mapped at LOD < 2 and thus were not used for deduction of an inversion

experienced up to 12 chromosomal rearrangements. In the T-genome, four inversions occurred—one inversion occurred in Tob19 (orthologous to Tmf1), and Tob24 (Tmf10), and two in Tob22 (Tmf6); in the S-genome, up to five inversions occurred—one inversion occurred in Tob3 (Acn2), Tob6 (Acn3), and Tob8 (Acn6) and two in Tob16 (Acn4). Here, it should be noted that since the Acn map is based on *N. acuminata* and *N. acuminata* var. multiflora but not *N. sylvestris* (the diploid species contributing S-genome), some of the rearrangements assigned to the S-genome may have actually occurred in the ancestral genome of *N. sylvestris* prior to tetraploidization and subsequent to the divergence of *N. sylvestris* and *N. acuminata*. Therefore, the estimate of genome rearrangements in the S-genome should be considered as an upper limit. Besides these nine inversions, ancestral Tmf6 was split into two non-homologous chromosomes (Tob21 and Tob22), while ancestral Acn1 and Acn9 were fused into Tob7, and a reciprocal translocation has exchanged parts of ancestral Tmf4 and Tmf8 and formed Tob12 and Tob23.

A number of chromosomal rearrangements also occurred to the diploid genomes of *N. tomentosiformis* and *N. acuminata* since their divergence from the last common ancestor. *N. tomentosiformis* experienced three inversions (Tmf1, Tmf4 and Tmf 10) prior to the tetraploidization and one inversion (Tmf3) subsequent to the tetraploidization. *N. acuminata* experienced one inversion (Acn2) before its

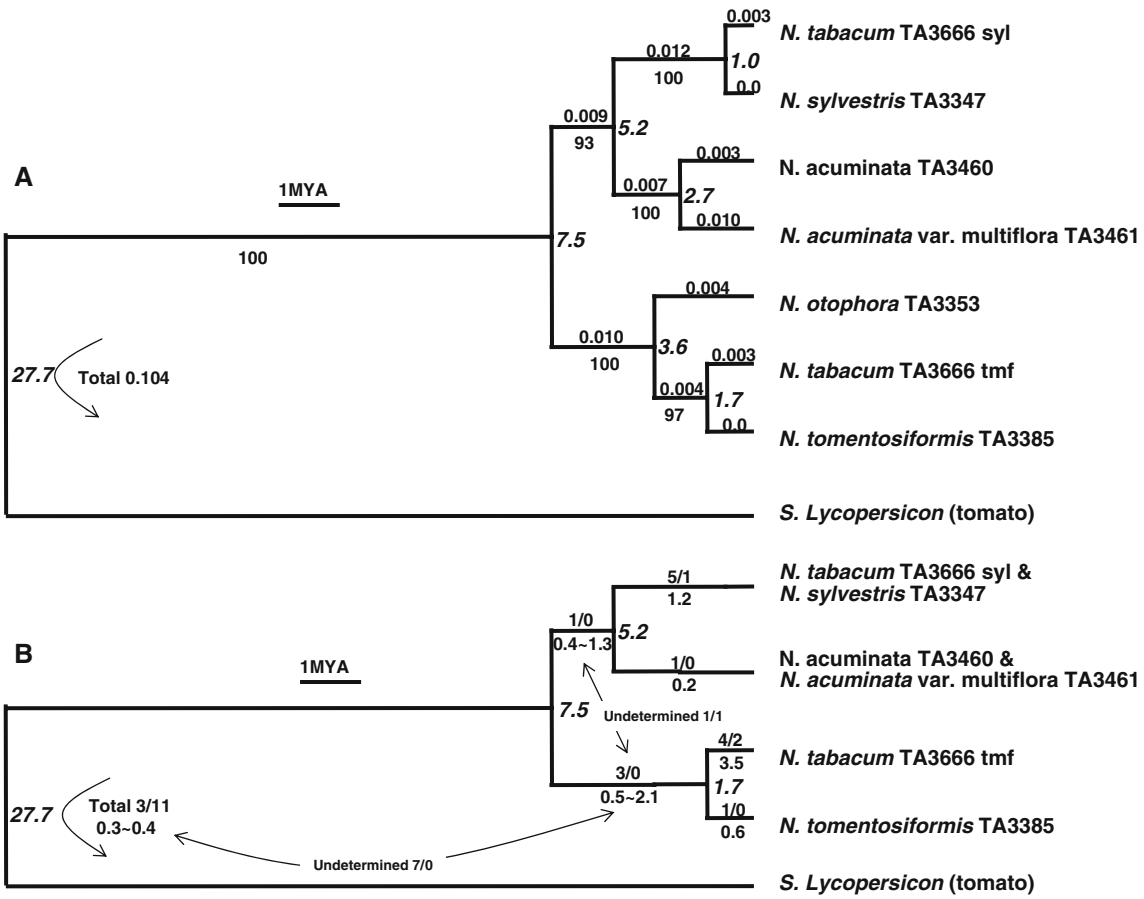


Fig. 4 Estimation of *Nicotiana* divergence time and placement of chromosomal rearrangements over phylogeny. Tobacco alleles were named *N. tabacum* TA3666 tmf for T-genome allele and *N. tabacum* TA3666 syl for S-genome allele. The calibration point was placed at 27.7MYA for the split of *Nicotiana* and tomato. The estimated divergence time (*bold*, *italic*) is listed next to each internal node. **a** The tree was reconstructed using maximum likelihood (ML) method. The substitution rate (substitution/bp) is placed *above branch* and the bootstrap value *below branch*. **b** The tree was modified from **a**. The number of rearrangements (inversions/translocations) is placed *above branch* and the evolutionary rate (rearrangements/million years) *below branch*

divergence from *N. sylvestris* and one inversion (Acn3) thereafter. In addition to these inversions, there are one inversion between Tmf5 and Acn5, and one translocation between linkage groups 10 and 12 but it is unknown whether they occurred along the Tmf or Acn lineage. It should be noted that all these estimations are based on the presented maps. Since some of these maps are based on rather small populations, it is possible that especially some inversions could be mapping artifacts. Nevertheless, we think that the overall picture provides an accurate estimation of the synteny levels between the two diploid species in relation to the tetraploid tobacco genome.

In total, 11 *Nicotiana* linkage groups (except for Tmf2/Acn2) are differentiated from the tomato genome by one or more translocation events, two for Tmf1/Acn1, Tmf5/Acn5, Tmf9/Acn9 and Tmf11/Acn11 while one for the other seven linkage groups. It requires a minimum of 11 reciprocal translocations for the transition from the tomato karyotype of these 11 chromosomes to that of either Tmf or Acn genome (Fig. 5). However, the order and timing of these events remain to be determined. In addition, numerous inversions differentiate the tomato genome from that of the last common ancestor of these studied *Nicotiana* species. The ancestral form of Tmf/Acn1 (the segment orthologous to T1), Tmf/Acn8 (T8) and Tmf10/Acn12 (T5) each has experienced one inversion. There are another seven inversions between the Tmf and tomato maps—one in each of Tmf3 (T3), Tmf5 (T5), Tmf6 (T6), Tmf8 (T10), Tmf9 (T9), Tmf11 (T11) and Tmf12 (T12). Since the involved markers were not mapped (or mapped at LOD < 2) in the Acn map, we cannot rule out the possibility that some of them may have occurred in the lineage leading to the Tmf genome and thus do not represent common differences between *Nicotiana* and tomato.

It was unexpected to identify fewer inversions than translocations (10 vs. 11) between *Nicotiana* and tomato. In other pairwise comparative mapping studies in the Solanaceae, the number of inversions was always much higher than that of translocations, e.g., 5 inversions and no

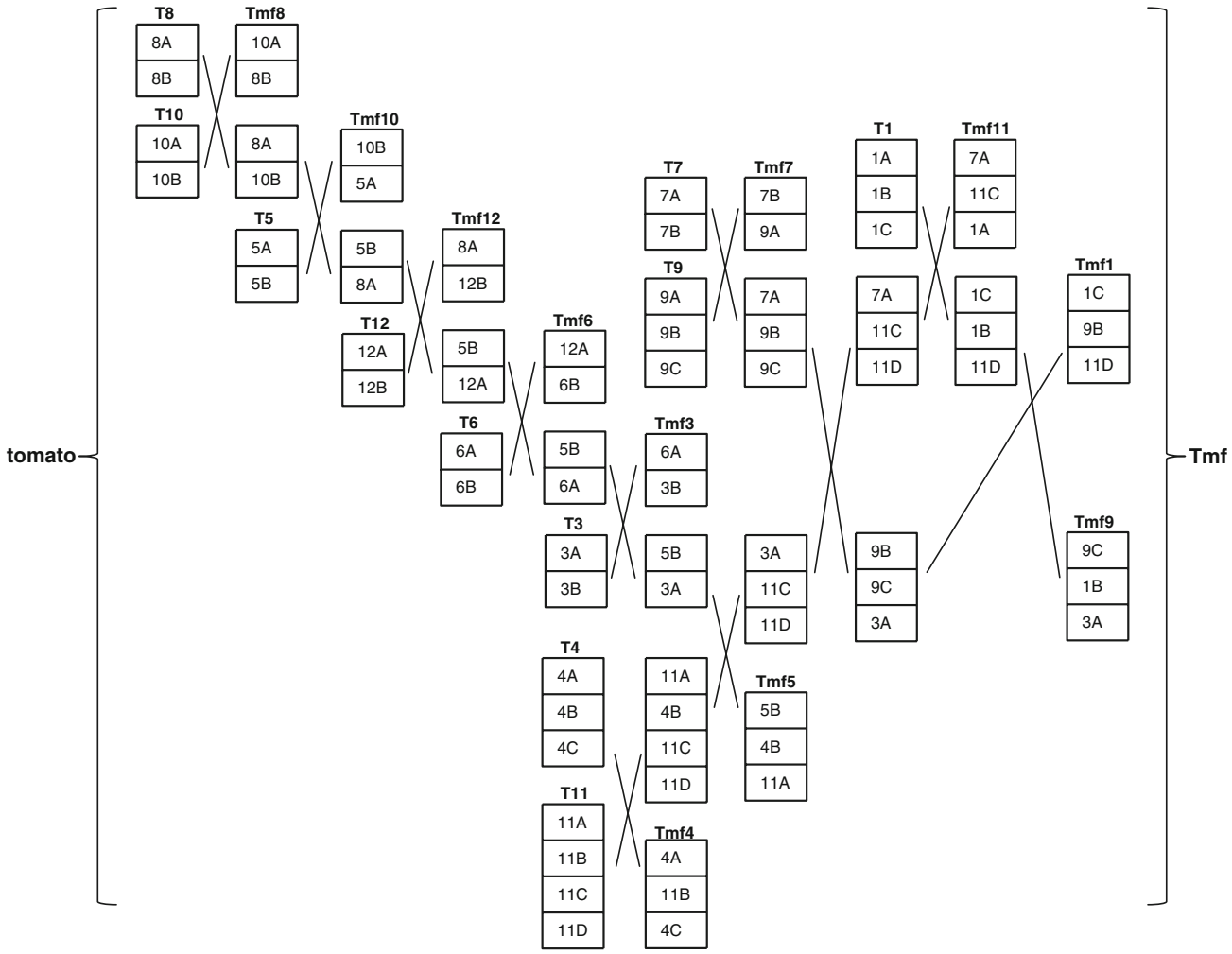


Fig. 5 One possible evolutionary pathway showing a minimum of 11 reciprocal translocations between the genomes of Tmf and tomato. *A–D* represents chromosome segments that were exchanged during translocations (see supplementary Fig. S2 for details). A *cross* indicates a translocation with an unknown direction

translocations between potato and tomato, 24 inversions and 5 translocations between eggplant and tomato, and 19 inversions and 6 translocations between pepper and tomato (Tanksley et al. 1992; Wu et al. 2009a, b). Since the synteny marker density in the latter two studies is similar to that between the Tmf and tomato maps, the exception in this study is possibly due to the following reasons. First, these 10 inversions and 11 translocations are differences between tomato and the last common ancestor of these studied *Nicotiana* species—an extinct *Nicotiana* species living probably around 7.5MYA. If we compare two extant species, *N. tomentosiformis* differs from tomato by a minimum of 14 inversions and 11 translocations. Secondly, the ratio of inversions versus translocations seems to decrease when the two compared species are more divergent, e.g., from potato versus tomato to *N. tomentosiformis* versus tomato. It is possible that multiple, sequential rearrangements shuffle the marker order to an extent that the individual inversions can no longer be deciphered using the method employed in this study. For instance, the inversion between Tmf5 and Acn5 involved markers from different tomato chromosomes—C2At5g16710 from T11 and C2At5g14520 from T7. Such an inversion could not be recognized in either Tmf–tomato or Acn–tomato comparison.

With the estimated divergence time, it was possible to calculate the absolute rates for chromosomal evolution in these species (Fig. 4b). After these studied *Nicotiana* species diverged from their common ancestor, *N. tomentosiformis* evolved at 0.5–2.1 rearrangements/MY before the tobacco speciation and thereafter at 0.6 rearrangements/MY (the range is due to the rearrangements with undetermined timing). *N. acuminata* experienced 0.4–1.3 rearrangements/MY before its divergence from *N. sylvestris* and thereafter 0.2 rearrangements/MY. In contrast, the chromosomal evolution has been much faster in the tetraploid tobacco genome. The T-genome of tobacco experienced 3.5 rearrangements/MY versus 0.6 in *N. tomentosiformis* in the same period. As for the S-genome of tobacco, the resultant 1.2 rearrangements/MY is the upper limit because some of the identified rearrangements may have occurred to the *N. sylvestris* genome before tobacco speciation. However, it is likely that the S-genome also experienced a much higher rate than *N. acuminata* (0.2 rearrangements/MY) in the same period. Consistently, the DNA substitution rates estimated by ML method (Fig. 4a) also suggested that T- and S-genome evolved much faster than their diploid relatives. Accelerated genetic change in the allotetraploid tobacco has been characterized from several aspects, including rDNA homogenization, retrotransposon activity and distribution, intergenomic translocations and so on (Dadejova et al. 2007; Kovarik et al. 2004, 2008; Lim et al. 2004; Melayah et al. 2004; Petit et al. 2007). Our results support the notion with additional evidence—each subgenome of the allotetraploid tobacco experienced chromosomal rearrangements more frequently than their diploid relatives. Such rapid genome changes in tobacco as well as other polyploidy species may be attributed to homeologous recombination, altered patterns of DNA methylation, enhanced transposable element activity and nuclear–cytoplasmic interaction (Doyle et al. 2008; Gaeta et al. 2007; Leitch et al. 2006; Wendel 2000).

Putative single gene transpositions

In addition to the translocations and inversions, a significant number of single marker transpositions disrupted the synteny between the *Nicotiana* and tomato genomes. For example, 36 tomato markers each mapped to a non-homologous chromosome on the Tmf map, and so did 13 tomato markers on the Acn map. A similar situation also existed between the two diploid *Nicotiana* genomes with four markers involved. This phenomenon was reported in both eggplant–tomato (5 cases) and pepper–tomato (12 cases) comparisons (Wu et al. 2009a, b), where it occurred with a lower frequency possibly due to shorter divergence time. A detailed analysis in the pepper–tomato work suggested that at least some of these transpositions had been facilitated by transposons or other mechanisms that are capable of moving small portions of genome from one site to another. A recent study revealed that such gene transposition was a common form of genome rearrangement in plant evolution (Freeling et al. 2008).

Conclusions

In this work, we constructed two diploid *Nicotiana* maps based on species closely related to *N. tomentosiformis* and *N. sylvestris* that are thought to be the closest living relatives to the diploid progenitors of the allotetraploid *N. tabacum*. Both maps contain more than 300 COSII and SSR markers and 12 linkage groups equal to chromosome number of the haploid *Nicotiana* genome. We compared the two diploid maps using common COSII markers and identified seven inversions and one reciprocal translocation since the divergence from their last common ancestor. We also compared the diploid maps with the map of tetraploid tobacco and inferred six chromosomal rearrangements in the tobacco T-genome but only one in the *N. tomentosiformis* genome, suggesting that the tetraploidization had led to a significantly higher rate of genome evolution. Furthermore, those mapped COSII markers on the diploid *Nicotiana* maps are distributed across the entire tomato map, and thus permitted inference of syntenic relationships between tomato and the last common ancestor of these studied *Nicotiana* species. Based on this comparison, we

inferred the *Nicotiana* and tomato genomes to differ by a minimum of 3–10 inversions and 11 reciprocal translocations as well as numerous putative single gene transpositions. The results of these comparative maps have made it possible to identify the orthologous tomato genome counterpart to every part of the tobacco genome especially a set of 25 CSSs within which the marker/gene order is well preserved. Thus, the tomato genome sequence and related resources can now be utilized by researchers working with species in the genus *Nicotiana*.

Acknowledgments This work was supported by Philip Morris International, Switzerland. We thank Xiaomin Jia, Ingrid S. Phillips, Doris Kriseleit and Katja Wendehake for technical assistance, Nick J. Van Eck for greenhouse work, and Dr. Lukas A. Mueller (Boyce Thompson Institute for Plant Research, Ithaca, NY, USA) and his bioinformatics group for hosting the maps and the marker data in Solanaceae Genomics Network (http://www.sgn.cornell.edu).

References

Bindler G, van der Hoeven R, Gunduz I, Plieske J, Ganal M, Rossi L, Gadani F, Donini P (2007) A microsatellite marker based linkage map of tobacco. Theor Appl Genet 114:341–349

Bogani P, Lio P, Intrieri MC, Buiatti M (1997) A physiological and molecular analysis of the genus *Nicotiana*. Mol Phylogenet Evol 7:62–70

Chaleff RS, Ray TB (1984) Herbicide-resistant mutants from tobacco cell-cultures. Science 223:1148–1151

Clarkson JJ, Lim KY, Kovarik A, Chase MW, Knapp S, Leitch AR (2005) Long-term genome diploidization in allopolyploid Nicotiana section Repandae (Solanaceae). New Phytol 168:241–252

Clemente T (2006) *Nicotiana* (*Nicotiana tobaccum*, *Nicotiana benthamiana*). Methods Mol Biol 343:143–154

Dadejova M, Lim KY, Souckova-Skalicka K, Matyasek R, Grandbastien MA, Leitch A, Kovarik A (2007) Transcription activity of rRNA genes correlates with a tendency towards intergenomic homogenization in Nicotiana allotetraploids. New Phytol 174:658–668

Daniell H, Streatfield SJ, Wycoff K (2001) Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. Trends Plant Sci 6:219–226

Doganlar S, Frary A, Daunay MC, Lester RN, Tanksley SD (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

Doyle JJ, Flagel LE, Paterson AH, Rapp RA, Soltis DE, Soltis PS, Wendel JF (2008) Evolutionary genetics of genome merger and doubling in plants. Annu Rev Genet 42:443–461

Frary A, Xu YM, Liu JP, Mitchell S, Tedeschi E, Tanksley S (2005) Development of a set of PCR-based anchor markers encompassing the tomato genome and evaluation of their usefulness for genetics and breeding experiments. Theor Appl Genet 111:291–312

Freeling M, Lyons E, Pedersen B, Alam M, Ming R, Lisch D (2008) Many or most genes in Arabidopsis transposed after the origin of the order Brassicales. Genome Res 18:1924–1937

Fulton TM, van der Hoeven R, Eannetta NT, Tanksley SD (2002) Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. Plant Cell 14:1457–1467

Gaeta RT, Pires JC, Iniguez-Luy F, Leon E, Osborn TC (2007) Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. Plant Cell 19:3403–3417

Goodin MM, Zaitlin D, Naidu RA, Lommel SA (2008) *Nicotiana benthamiana*: its history and future as a model for plant-pathogen interactions. Mol Plant Microbe Interact 21:1015–1026

Julio E, Verrier JL, de Borne FD (2006) Development of SCAR markers linked to three disease resistances based on AFLP within *Nicotiana tabacum* L. Theor Appl Genet 112:335–346

Kenton A, Parokonny AS, Gleba YY, Bennett MD (1993) Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. Mol Gen Genet 240:159–169

Kessler A, Halitschke R (2007) Specificity and complexity: the impact of herbivore-induced plant responses on arthropod community structure. Curr Opin Plant Biol 10:409–414

Kessler D, Gase K, Baldwin IT (2008) Field experiments with transformed plants reveal the sense of floral scents. Science 321:1200–1202

Knapp S, Chase MW, Clarkson JJ (2004) Nomenclatural changes and a new sectional classification in *Nicotiana* (Solanaceae). Taxon 53:73–82

Konieczny A, Ausubel FM (1993) A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. Plant J 4:403–410

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

Kovarik A, Matyasek R, Lim KY, Skalicka K, Koukalova B, Knapp S, Chase M, Leitch AR (2004) Concerted evolution of 18-5.8-26S rDNA repeats in Nicotiana allotetraploids. Biol J Linn Soc 82:615–625

Kovarik A, Dadejova M, Lim YK, Chase MW, Clarkson JJ, Knapp S, Leitch AR (2008) Evolution of rDNA in Nicotiana allopolyploids: a potential link between rDNA homogenization and epigenetics. Ann Bot 101:815–823

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

Leitch A, Lim KY, Skalicka K, Kovarik A (2006) Nuclear cytoplasmic interaction hypothesis and the role of translocations in Nicotiana allopolyploids. In: Cigna AA, Durante M (eds) Radiation risk estimates in normal and emergency situations. Springer, The Netherlands, pp 319–326

Lewis RS, Nicholson JS (2007) Aspects of the evolution of *Nicotiana tabacum* L. and the status of the United States *Nicotiana* Germplasm Collection. Genet Resour Crop Evol 54:727–740

Lim KY, Matyasek R, Kovarik A, Leitch AR (2004) Genome evolution in allotetraploid *Nicotiana*. Biol J Linn Soc 82:599–606

Lin TY, Kao YY, Lin S, Lin RF, Chen CM, Huang CH, Wang CK, Lin YZ, Chen CC (2001) A genetic linkage map of *Nicotiana plumbaginifolia*/*Nicotiana longiflora* based on RFLP and RAPD markers. Theor Appl Genet 103:905–911

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

Melayah D, Lim KY, Bonnivard E, Chalhoub B, De Borne FD, Mhiri C, Leitch A, Grandbastien M (2004) Distribution of the Tnt1 retrotransposon family in the amphidiploid tobacco (*Nicotiana tabacum*) and its wild *Nicotiana* relatives. Biol J Linn Soc 82:639–649

Neff MM, Neff JD, Chory J, Pepper AE (1998) dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thaliana* genetics. Plant J 14:387–392

Neff MM, Turk E, Kalishman M (2002) Web-based primer design for single nucleotide polymorphism analysis. Trends Genet 18:613–615

Nishi T, Tajima T, Noguchi S, Ajisaka H, Negishi H (2003) Identification of DNA markers of tobacco linked to bacterial wilt resistance. Theor Appl Genet 106:765–770

Okamuro JK, Goldberg RB (1985) Tobacco single-copy DNA is highly homologous to sequences present in the genomes of its diploid progenitors. Mol Gen Genet 198:290–298

Olmstead RG, Sweere JA, Spangler RE (1999) Phylogeny and provisional classification of the Solanaceae based on chloroplast DNA. In: Nee M, Symon DE, Lester RN, Jessop JP (eds) Solanaceae IV. Royal Botanic Gardens, Kew, pp 111–137

Petit M, Lim KY, Julio E, Poncet C, de Borne FD, Kovarik A, Leitch AR, Grandbastien MA, Mhiri C (2007) Differential impact of retrotransposon populations on the genome of allotetraploid tobacco (*Nicotiana tabacum*). Mol Gen Genom 278:1–15

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817–818

Ren N, Timko MP (2001) AFLP analysis of genetic polymorphism and evolutionary relationships among cultivated and wild *Nicotiana* species. Genome 44:559–571

Rossi L, Bindler G, Pijnenburg H, Isaac PG, Giraud-Henry I, Mahe M, Orvain C, Gadani F (2001) Potential of molecular marker analysis for variety identification in processed tobacco. Plant Var Seeds 14:89–101

Sanderson MJ (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol Biol Evol 14:1218–1231

Swofford DL (2003) PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, MA

Tanksley SD, Ganal MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160

Wang Y, Diehl A, Wu FN, Vrebalov J, Giovannoni J, Siepel A, Tanksley SD (2008) Sequencing and comparative analysis of a conserved syntenic segment in the solanaceae. Genetics 180:391–408

Wendel JF (2000) Genome evolution in polyploids. Plant Mol Biol 42:225–249

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 and systematic studies: a test case in the euasterid plant clade. Genetics 174:1407–1420

Wu FN, Eannetta NT, Xu YM, Durrett R, Mazourek M, Jahn MM, Tanksley SD (2009a) A COSII genetic map of the pepper genome provides a detailed picture of synteny with tomato and new insights into recent chromosome evolution in the genus *Capsicum*. Theor Appl Genet 118:1279–1293

Wu FN, Eannetta NT, Xu YM, Tanksley SD (2009b) A detailed synteny map of the eggplant genome based on conserved ortholog set II (COSII) markers. Theor Appl Genet 118:927–935