A Preliminary Study of Threonine Deaminase Duplication in Solanaceae

by

Jie Huang

Department of Biology
Duke University

Date:_____
Approved:

_____
Mark Rausher, Supervisor

_____
Mohamed Noor

_____
John Willis

Thesis submitted in partial fulfillment of
the requirements for the degree of
Master of Science in the Department of
Biology in the Graduate School
of Duke University

2013

ABSTRACT

A Preliminary Study of Threonine Deaminase Duplication in Solanaceae

by

Jie Huang

Department of Biology
Duke University

Date:_____
Approved:

_____
Mark Rausher, Supervisor

_____
Mohamed Noor

_____
John Willis

An abstract of a thesis submitted in partial
fulfillment of the requirements for the degree
of Master of Science in the Department of
Biology in the Graduate School of
Duke University

2013

Copyright by
Jie Huang
2013

Abstract

One of the most important questions in evolutionary biology is how new genes and new functions arise and evolve. Among the theories addressing this question, gene duplication is one of the most popular. Previous study has shown that two threonine deaminase (TD) gene copies exist in *Solanum lycopersicum*, and these two copies have very different functions and low sequence similarities. The primary objective of this study was to widen our understanding of this gene duplication and the subsequent evolutionary processes affecting the duplicate copies by first collecting additional TD sequences from related species, building a gene tree, and inferring the point of gene duplication. The evolutionary processes acting on this gene were then analyzed using the program PAML. Results indicate that 1) The TD duplication probably occurred in before the split of the Solanoideae from the Nicotianoidea; and 2) there is strong evidence for positive selection on one of the TD copies after gene duplication, while for the other TD copy, only weak evidence for positive selection was found; and 3) adaptive improvement of the copy with new function probably spanned a period of at least 25 million years.

Contents

List of Tables

List of Figures

Acknowledgements

Life and study here, in a new country, is not as easy as I imagined. I feel sorry that I give up the chance to be a doctor, while at the same time, I am happy to make it to the Master's degree. During my four years here, there are many people to thank.

First, and the most importantly, I want to give my warmest thank to my advisor Mark Rausher. You are always so nice and patient to me ever since one year before I came studying here, and even when I hadn't obtained any data for months, and even after I told you that I would stop my study. You are always so encouraging and never stint your praise --- I was praised and encouraged even after I obtained only a tiny piece of result. And you always revise my writing so carefully, patiently and quickly. I am so thankful to have you as my advisor, and so sorry that I am not continuing this glory under your supervision.

I'd like to thank my current and previous committee member, Mohamed Noor, John Willis, Greg Wray and Xinnian Dong. You gave me so many good suggestions, and are always so nice and supportive to me. I am so glad and honored to have chosen you as my committee.

I'd like to thank my lab mates, especially Tanya Duncan, Talline Martins and Rong-Chien Lin. You helped me so much in lab, and more importantly, I received so much happiness and encouragement talking with you. I'd also like to thank Cheng-Ruei

Lee for helping me with the PAML analysis and Eric Green for helping me with the likelihood ratio test.

I'd like to thank all the biology staff, especially our dear graduate coordinator Anne Lacey. You are so considerate about everything and every step we need to do regarding to continuing our study here. And you are so kind and patient to a procrastinated student, me. And when I was sad, you are the one I know for sure to give comfort.

I also want to thank my dear old friends who live faraway, and of course my parents. You are my major my source of happiness, and I feel so attached talking and being with you, and I miss you so much and so often during my studying here. I am not sociable, but I am very glad and thankful to make a few new friends here. Thanks to you all, who made my life here rich.

In the end, I want to give my sincerely thank to boyfriend, Junpeng. We met on our first flight to the United States, and ever since then, we've spent so much fun time together. You have the largest tolerance I've ever seen. You are not only always so patient and nice, but also, you are my personal cook, driver, dog walker, alarm, vent to anger, trip planner, joke producer, and happiness exporter. I am so thankful and appreciative to have met and known you.

1. Introduction

From the time the first single-cell life form appeared around 3.8 billion years ago in the sea (Mojzsis et al. 1996), to the present magnificent bio-diversity almost everywhere on earth, gene functions have become much more diversified. One important issue in evolutionary biology is how new genes and new functions arise and evolve (Long et al. 2003). Gene duplication is one of the most poplar theories to address this question. Ever since Bridges discovered in 1936 the first example of gene duplication in *Drosophila melanogaster* (Bridges 1936), many cases of gene duplication have been studied and many gene duplication models have been proposed (Innan & Kondrashov 2010). Three of the most popular models are the neo-functionalization model (NEOF), the sub-functionalization model (SUBF), and the escape from adaptive conflict model (EAC).

The neo-functionalization model (NEOF) states that after gene duplication, one copy maintains the original function and the other copy is freed from purifying selection, so that it can accumulate all kinds of mutations (Ohno 1970). Because most mutations are deleterious, one of the copies would in most cases end up as a pseudo-gene; however, there is a non-negligible chance that beneficial mutations will give a duplicated copy a new function. If beneficial, these mutations would further be fixed and maintained by selection. The sub-functionalization model postulates that after gene

duplication, selection is relaxed in both copies (Force et al. 1999). Consequently, deleterious mutations arise in both copies, but inactivate different sub-functions in the two copies, and are fixed by genetic drift. As a result, range of functions of each copy is reduced but they are complementary: together the two copies perform all of the original functions. Therefore, after these degenerate mutations are fixed by drift in both copies, both copies become necessary and thus would be maintained by purifying selection. The escape from adaptive conflict (EAC) model (Hughes 1994) postulates that if one gene has two functions, it will often be the case that neither function can be optimized because of a negative tradeoff: improvement of one function can only occur at the cost of decreasing the other function. However, gene duplication releases this constraint, allowing each copy to be optimized for one function by positive selection and after fixation be maintained by purifying selection. In this way, the adaptive conflict is solved.

These three models can be defined and distinguished by several different properties (Table 1). One property is the ancestral gene function. By definition, under NEOF the ancestral gene had only one function, while under EAC the ancestral gene had at least two functions. With SUBF, however, there are several possible cases: if the ancestor had one function, the daughter copies could be sub-functionalized in their regulatory domains and be expressed in different tissues; if, on the other hand, the ancestral copy had multiple functions, the coding region could be sub-functionalized, with different copies performing different functions.

Another property that can distinguish the models is whether there was adaptive conflict in the ancestral gene. For NEOF, adaptive conflict is irrelevant because the ancestral gene performs only one function. With SUBF, no adaptive conflict exists in the ancestral gene. In contrast, under EAC the ancestor exhibited adaptive conflict between multiple functions by definition.

A third property by which the models can be distinguished is the whether selection vs. genetic drift operates on the two copies after a gene duplication has become fixed. With NEOF, positive selection occurs only on the copy that is neo-functionalized, while purifying selection acts on the copy that maintains the ancestral function. By contrast, under EAC, positive selection occurs on both copies to optimize different functions in the two copies after duplication releases the adaptive constraint. Finally, under SUBF, selection does not act on either copy because drift fixes the mutations that lead to sub-functionalization.

The differences in these three properties provide a series of tests that allow one to determine which of these models best explains the post-duplication process in individual cases(Des Marais & Rausher 2008). In particular, asking whether the ancestral copy performed a single or multiple functions, whether the pre-duplication ancestral copy exhibited adaptive conflicts between its multiple functions, and whether positive selection occurred after duplication on one, both or neither copies should reveal a pattern that is specific to just one of the three processes (Table 1).

Table 1. Comparison of predictions of the NEOF, SUBF and EAC models.

	Ancestral gene functions	Adaptive conflict in ancestral gene	Observed positive selection
NEOF	One	NA	On one copy
SUBF	One or more	No	On neither copy
EAC	Two or more	Yes	On both copies

Threonine deaminase (TD) catalyzes the first step of the synthesis of Isoleucine (Ile) in most organisms (Umbarger 1956). In *Solanum lycopersicum* (cultivated tomato), a second copy of threonine deaminase (TD2) helps plants defend against lepidopteran herbivores (Chen et al. 2005).

Threonine deaminase 1 (TD1) in *Solanum lycopersicum* performs the function that the single copy in most other species performs, which is to catalyze threonine (Thr) to α-keto butyrate, the first step of Ile synthesis. This enzyme has 606 amino acids and 3 domains: an N′ catalytic sequence, a short neck region and a C′ regulatory sequence. High concentration of Ile interacts with the C′ regulatory sequence to inhibit the activity of TD1, so that a proper concentration of Ile is maintained.(Chen et al. 2007; Gonzales-Vigil et al. 2011)

In *Solanum lycopersicum*, threonine deaminase 2 (TD2) has 595 amino acids and shares 51% amino acid sequence identity to TD1. The primary TD2 protein has all three domains. However, in the midgut of lepidopteran larvae, the C′ regulatory sequence of TD2 is digested (producing pTD2) so that the activity of TD2 cannot be negatively inhibited by high concentrated Ile. As a result, Thr coverts most threonine to α-keto butyrate and the insects die because of the depletion of Thr. Unlike TD1, which is a housekeeping gene expressed in all tissues, TD2 is expressed primarily in a subset of tissues such as flower buds and unopened flowers. TD2 expression in leaves is induced by mechanical wounding and can also be induced by methyl jasmonate (MeJA). These

properties are consistent with the primary function of TD2 as defense because flowers are the most important reproductive organs and special protection is needed; MeJA is a defense hormone that plants produce under stress and can be induced by mechanical wounding by larvae. (Chen et al. 2005; Gonzales-Vigil et al. 2011)

In *Nicotiana attenuata* there is a single copy of TD. Kang J. *et al.* (2006) reported that this copy performs both functions: its housekeeping function and defense against herbivores. They also report that it functions as a defense in two ways: 1) it facilitates the production of Ile, which conjugates with jasmonic acid (JA-Ile) to induce the accumulation of direct defenses such as trypsin protease inhibitor and nicotine to defend against herbivores; and 2) it provides anti-nutritive defense in the same way as TD2 in *S. lycopersicum*, which is to deplete Thr in the gut of herbivores.

These two cases suggest that the duplication of TD that led to the production of TD1 and TD2 with separate function in *S. lycopersicum* may be an example of the EAC process. *Nicotiana* is a close relative of *Solanum* within the *Solanaceae*. The *N. attenuata* TD (NaTD) appears phylogenetically intermediate between two clades representing TD1 and TD2 function (Chen et al. 2007). Therefore, the single copy TD in *N. attenuata* might represent something close to the ancestral statue (Chen et al. 2005). An objective of my research is to identify TD genes in other *Solanaceae* species, and perhaps species from related families, in order to perform a phylogenetic reconstruction to (1) determine where on the gene tree duplication occurred and (2) to determine the evolutionary forces

experienced by each TD clade after duplication. Based on these results, I can preliminarily analyze which gene duplication fits this example the best.

2. Methods

2.1 Obtaining sequence data

In order to reconstruct the ancestral protein sequence and distinguish between drift and selection operating after duplication, where on a phylogeny the duplication occurred was studied. To do this, relatives at different phylogenetic distances from *S. lycopersicum*, according to the phylogeny of the Solanaceae (Olmstead et al. 2008; Bohs & Olmstead 1997; Rodriguez et al. 2009), was surveyed by performing BLAST searches and attempting to clone TD1 and TD2. First, BLAST (Sayers et al. 2012) online searching for TD sequences was performed using known TD sequences (Chen et al. 2007). Second, for species of which TD sequences were not found by BLAST, seeds were ordered and planted, and three-week young plants were induced by MeJA overnight, following which RNA was extracted from leaves (Chen et al. 2005). Based on the TD gene sequences from Chen et *al.* (2007), conservative primers were designed separately for TD1 and TD2. Bands resulting from RT-PCR were extracted, sequenced and aligned with TD1 and TD2 genes to confirm their homology.

2.2 Gene Tree construction and duplication point determination

Twenty-three Solanaceae TD sequences and five outgroup TD sequences with different evolutionary distances to Solanaceae were fist aligned using ClustalW (in MEGA 5.0), and then TD gene tree was constructed and compared using the Maximum

Likelihood (ML), Neighbor-joining (NJ), Minimum Evolution, and UPGMA methods in MEGA 5.0 (Tamura et al. 2011). Under ML methods, different nucleotide substitution models (Jukes & Cantor (JC), Kimura-2-parameter, Tamura-3-parameter, HKY, Tamura-Nei, and General Time Reversible model) were used separately and the resulting tree topologies were compared. All the trees were simulated under 100 bootstrap repetitions. Another TD tree was constructed using only ML method with JC model (500 bootstrap), but including several additional short TD sequences (less than 1kb) in *Petunia*, *Ipomoea* and *Mimulus*. The point of gene duplication was estimated by considering these two trees together based on the following criteria: A) most species within the duplicate clade have two TD copies (not all species because some duplicate copies may have been lost); B) the outgroup species have one TD copy or have multiple TD copies with different origin.

As shown below, the ML tree grouped each TD sequence from *Nicotiana* with one or the other of two Solanoidea clades that represent the two duplicates. To determine whether this was an appropriate placement of these sequences, I generated a second tree in which the *Nicotiana* sequences were an outgroup of the two *Solanoidea* clades. Using PUAP (Swofford 2003), I compared the log-likelihoods of the two trees to determine whether the constraint significantly reduced the likelihood. A significant likelihood ratio test would indicate that second tree should be rejected.

2.3 Positive selection tests

Determining whether substitutions following duplication are due to selection or drift was examined using the program CODMEL of PAML 4.6 package (Yang 2007). Lineage- and clade-specific analyses of the ratio of synonymous to nonsynonymous substitution rate (ω) was performed as in Des Marais and Rausher (Des Marais & Rausher 2008) and Fang et al. (2009). Only TD sequences with full coding regions were used, which resulted in 16 Solanaceae TD sequences and 3 outgroup TD sequences. These sequences were then aligned using ClustalW (codon) in MEGA 5.0. Regions with gaps (insertion or deletion) were excluded, and the tree was then constructed using ML method with JC model in MEGA 5.0 (Figure 3). Next, a branch-site model test (model=2, NSsites=2) was examined on all the branches within TD1 and TD2 clades. To test for significance, likelihood ratio test (LRT) was performed by calculating $\Lambda = 2\Delta l$ (twice the difference of the maximum log-likelihood values of the null and alternative hypotheses). This statistic has a chi-square distribution with degree of freedom (df) equal to the difference in number of parameters between the two hypotheses. The alternative hypothesis allows sites to be under positive selection, while the null hypothesis does not. Moreover, branch-site model tests (model=2, NSsites=2) on all the branches within TD1 clade and all branches with TD2 clade (or a subset of TD1 or TD2 clade) were performed to examine positively selected sites in multiple branches together as a group. Additionally, branch model (model=2, NSsites=0) tests were performed which allowed

all branches within one clade to be the same. ω of TD1 clade (ω_1), TD2 clade (ω_2) and the other branches (ω_0) were estimated, and LRT were performed as described above. In the ends, ancestral states were reconstructed by by Fitch Parsimony rules (Fitch 1971) on these five sites, base on which the branches these substitutions occurred were inferred.

3. Results

3.1 TD gene tree and duplication point

Based on the online BLAST searching, previous studies (Chen et al. 2007; Kang et al. 2006) and RT-PCR results, 16 TD sequences with full coding region and 6 TD sequences with incomplete coding region were found in plants in Solanaceae. Gene trees were constructed using different methods and nucleotide substitution models, and they all give very similar result (Appendix J), which are also consistent with the species tree (Appendix H). Figure 1 shows the TD tree obtained using the ML method with the JC model, in which almost all species of subfamily Solanoideae examined have both TD1 and TD2 sequences. TD2 sequences from the Solanoideae species form a single clade, which indicates a single origin. Nicotianoideae is the subfamily closest to the Solanoideae. Within this subfamily, there is only one TD copy in *N. attenuata* (Kang et al. 2006), but 2 TDs from online sequences of *N. tabacum* were found. However, *N. tabacum* is a tetraploid species resulting from artificial hybridization (Cameron 1952). Consequently, these two copies likely reflect tetraploidy rather than duplication. According to this result, there are two possibilities: 1) the TD gene duplication occurred after the divergence of Solanoideae and Nicotianoideae subfamilies; 2) TD gene duplication occurred before the divergence of Solanoideae and Nicotianoideae subfamilies but one TD copy was lost in the common ancestor of *N. tabacum* and *N. attenuata*.

In order to distinguish between these two hypotheses, three other outgroups, the Petunieae, Convolvulaceae and Phrymaceae, were examined, from which only short partial sequences (less than 1kb) were obtained from online database. Figure 2 shows the result of the tree that includes TDs from two *Petunia* species, *Ipomoea nil and Mimulus guttatus*. Although each has two TDs, none are orthologous to TD2 in the Solanoideae, which strongly suggests independent origination. Moreover, RT-PCR using different conserved primers for TD2 was performed on Petunia, but nothing was amplified, indicating that TD2 probably does not exist in this species.

In both Figures 1 and 2, the *Nicotiana* sequences fall within either TD clade 1 or TD clade 2. This pattern suggests that the TD duplication occurred before the split of the Solanoideae from the Nicotianoidea. Because the *N. tabacum* sequences fall within clade 1, while the *N. attenuata* sequence falls within clade 2, this tree implies that TD2 was lost in the progenitors of *N. tabacum*, while TD1 was lost in *N. attenuata* (hypothesis 1). A plausible alternative hypothesis is that the duplication occurred at the base of the Solanoideae (hypothesis 2), since both *Nicotiana* species examined contain only one copy of TD (Notice that *N. tabacum* has two TD copies due to polyploidization, instead of gene duplication). This hypothesis would require that the *Nicotiana* sequences form an outgroup to the *Solanoideae* (Appendix D, constrained tree).

I tested whether hypothesis 2 could be rejected by comparing the likelihood associated with the tree in Figure 1 with the constrained tree. According to the results,

reported in Appendix E, the constrained tree has a statistically lower likelihood than the original tree, which suggests that the constrained tree is probably incorrect. This result implies that hypothesis 2, duplication occurring at the base of Solanoideae, is probably wrong. However, because of the small number of Nicotianoidea species examined, I am suspicious of this result. Therefore, in the following analyses, the same tests were performed on both of the two hypothesized trees.

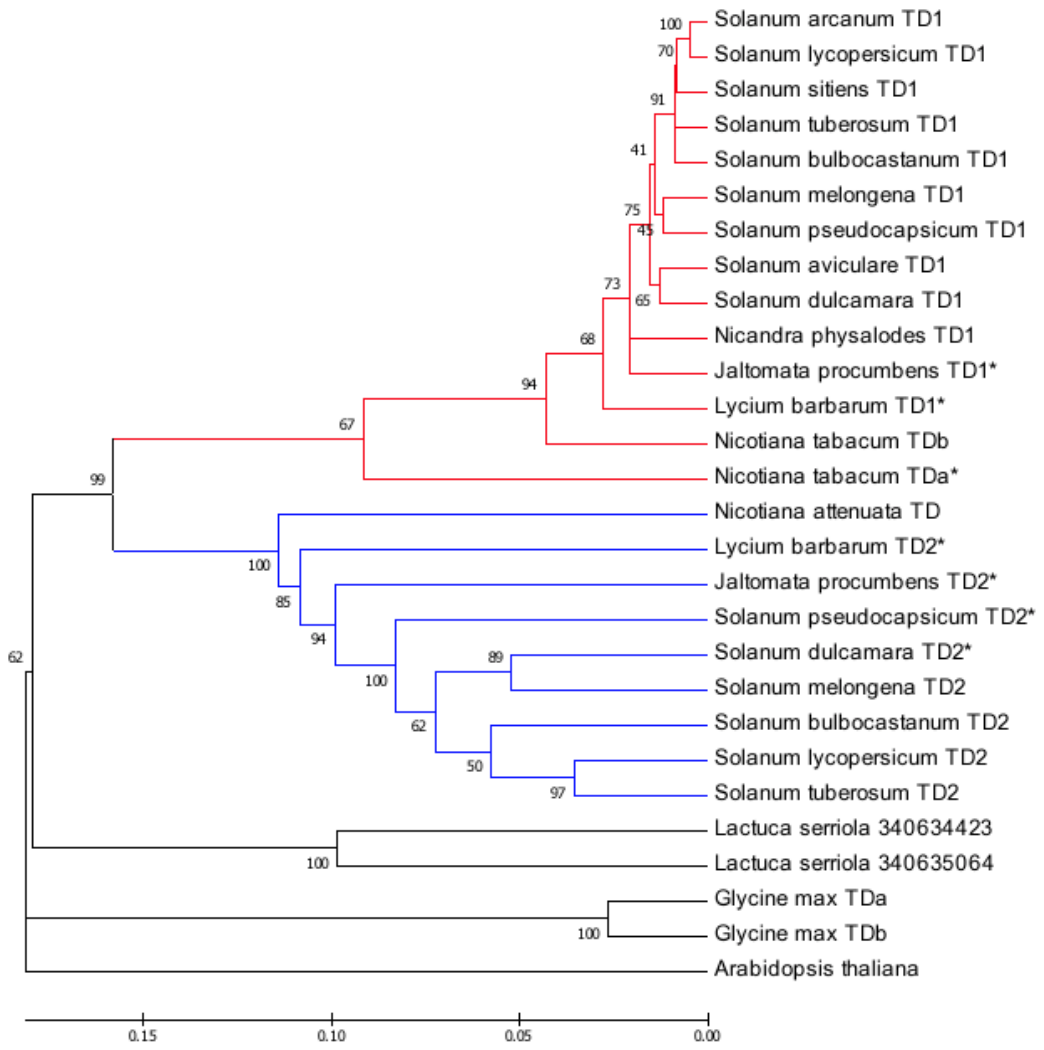


Figure 1. TD gene tree constructed using MEGA 5.0. Maximum likelihood method (JC69 model) was used with 100 bootstrap replications. The numbers above each branch represent bootstrap percentages. The red and blue lines represent branches in TD1 clade and TD2 clade respectively. Asterisks represent sequences of which incomplete coding sequences were obtained but were used to build the alignment and tree.

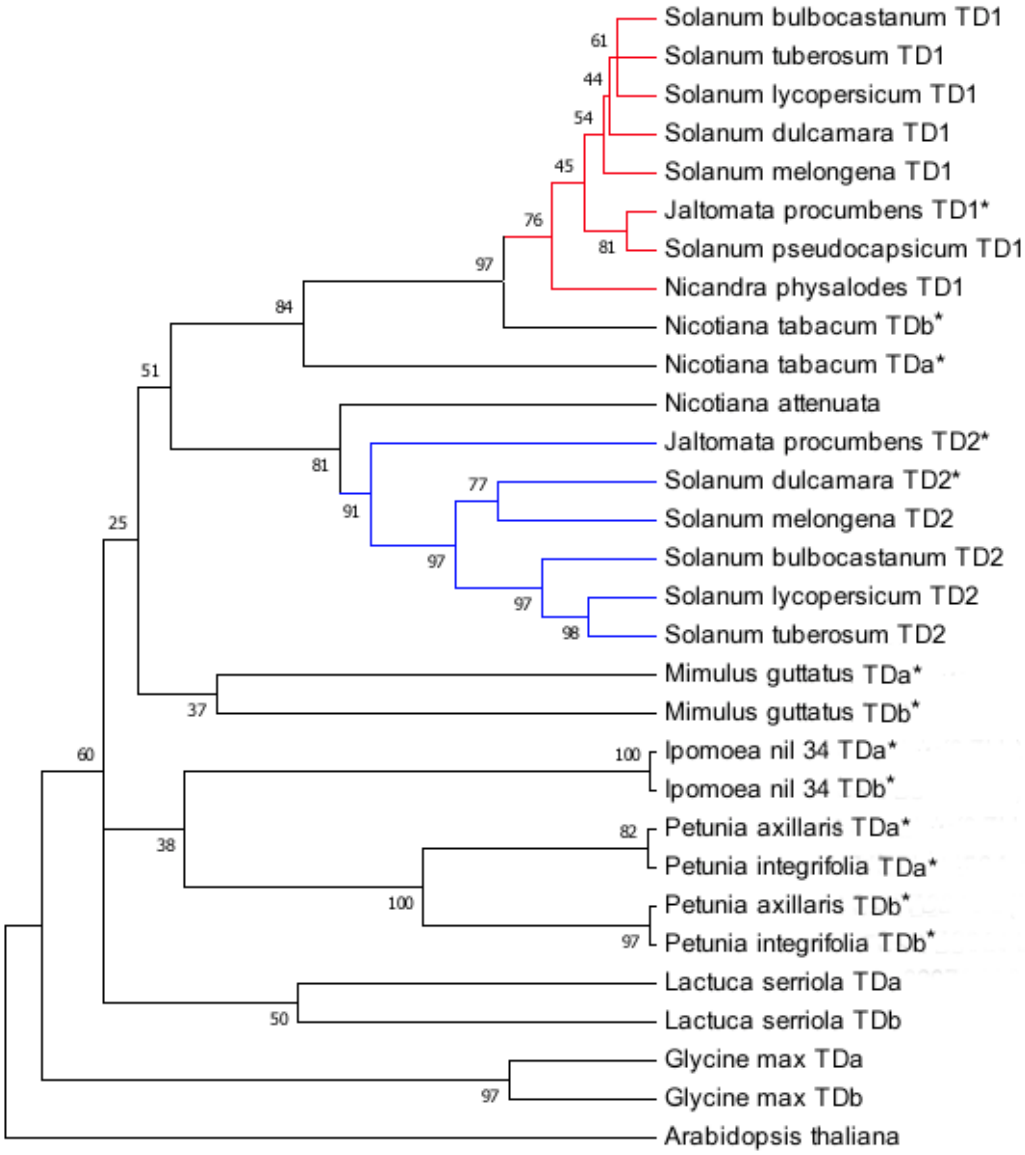


Figure 2. TD gene tree including *Ipomoea nil, Petunia axillaris, Petunia integrifolia* and *Mimulus guttatus*. Method and color notation are the same as in Figure 1.

3.2 Analysis of positive selection

For this analysis, I examined two gene trees. One, depicted in Figure 3, was constructed using the same method as Figure 1, but contains only the sequences with complete coding regions (subsequently I refer to this tree as Tree 1). This tree reflects the hypothesis that duplication of TD occurred prior to the split of the Nicotianoideae and the Solanoideae, but that TD2 was lost in the two *Nicotiana* species examined (hypothesis 1). The second tree (Tree 2) constrains the *Nicotiana* sequences to be outgroups to the two Solanoidea clades (Figure 4). This tree reflects the hypothesis that the TD duplication occurred after the split of the two subfamilies on the basal branch of the Solanoideae (hypothesis 2). The following analyses are applied to each tree because I believe I cannot distinguish between hypotheses (1) and (2) with current data.

3.2.1 PAML analysis based on the original gene tree

Branch model tests which allow different branches to have different values of ω were performed first. All branches in the tree in Figure 3 were divided into three groups: group 1 consisted of all branches in the TD1 clade (red branches) and are assigned an ω value of ω_1; group 2 consisted of all branches in the TD2 clade (blue branches) with ω value of ω_2; and group 0 consisted of all the remaining branches (black branches) with $\omega = \omega_0$. ω values were set to be the same for branches within each group. The average ω values for group 1, group 2 and group 0 are 0.18, 0.48 and 0.15 respectively, which

suggests an elevated ω in clade TD2. In order to test whether this difference is statistically significant, the ω values were compared pairwise by calculating the difference in log-likelihood ratios (Λ) for a model in which the two groups are allowed to have different ω values to a model in which the two groups are constrained to have equal ω values (Table 2). These analyses reveal that ω_0, ω_1 and ω_2 are all pairwise different, which indicates that the both TD1 and TD2 clade have an elevated dN/dS ratio, which could be caused either by an elevated rate of positive selection, by relaxation of purifying selection, or both.

To specifically test whether positive selection acted on any branches, branch-site model tests were performed. Under this model, the clade corresponding to TD2 (subtended by branch I, Figure 3) did not exhibit significant positive selection, as judged by the log-likelihood ratio statistic (Table 3, Group 2). However, all of its subclades (subtended by branch G, F or C, Figure 3) exhibited significant positive selection (Table 3). Moreover, individual branches (branch A to I) all exhibit significant positive selection with sites under positive selection found, except for the branch associated with *N. attenuata*, for which no sites under positive selection was found. The analysis identified 83 sites subject to positive selection in TD2 clade (Appendix F). By contrast, only one positively selected site was detected in the clade corresponding to TD1 (subtended by branch J, Figure 3; Table 3, Group 1), but no other positive sites were found on any of the subclades of TD1, or on any of the individual branches within TD1 clade.

3.2.2 PAML analysis based on the constrained tree

Results of the analysis using the constrained tree were generally similar to those described above. Average ω values for group 1, group 2 and group 0 are 0.19, 0.41 and 0.15 respectively, very similar to the values found in the previous analysis. However, unlike previously, ω_1 does not differ significantly from ω_0, while ω_2 is significantly higher than both ω_0 and ω_1 (Table 2), suggesting elevated positive selection or relaxed purifying selection in the TD2 clade, but not in the TD1 clade.

Branch-site model tests also revealed a similar pattern (Table 4). When group 2 (the TD2 clade) is analyzed, there is no statistical evidence of positive selection. However, when the basal branch G is omitted, the remaining clade exhibits highly significant positive selection, as do all of its subclades and individual branches. For the full TD2 clade, a total of 53 sites are identified as subject to positive selection. 50 of these were also identified in the previous analysis, although 33 sites identified in the previous analysis were not identified by this analysis, and 3 new sites were identified by this analysis. In this analysis, group 1 (the TD1 clade, subtended by branch K) exhibits significant positive selection, but only two sites are identified as being subject to selection. And similarly, no other TD1 subclades or individuals are identified as being subject to positive selection.

In summary, the gene trees depicted in Figures 3 and 4 give consistent results: extensive positive selection occurred in the TD2 clade but not in the TD1 clade. Many

positively selected amino acid residues on TD2 branches were found using both methods, and a large proportion of them were common to both analyses. By contrast, little evidence of positive selection was found for the TD1 clade.

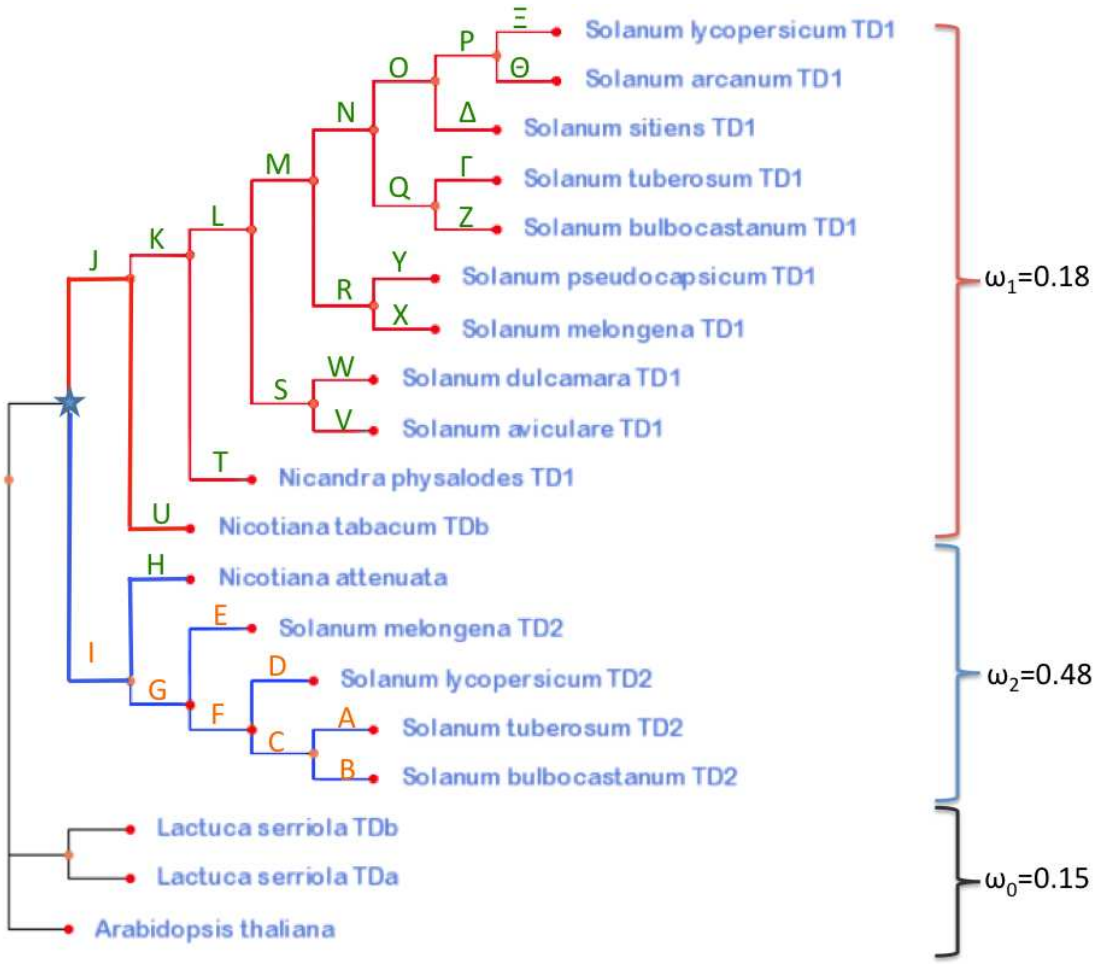


Figure 3. TD gene tree used for the first PAML analysis (Tree 1). Blue and red lines show branches within TD2 and TD1 clades respectively. Orange letters represent branches having positively selected sites under branch-site model. Green letters present branch names for the rest branches in TD1 and TD2 clades. Blue star shows the point where duplication occurred. ω Values on the right show the average ω of each branch group.

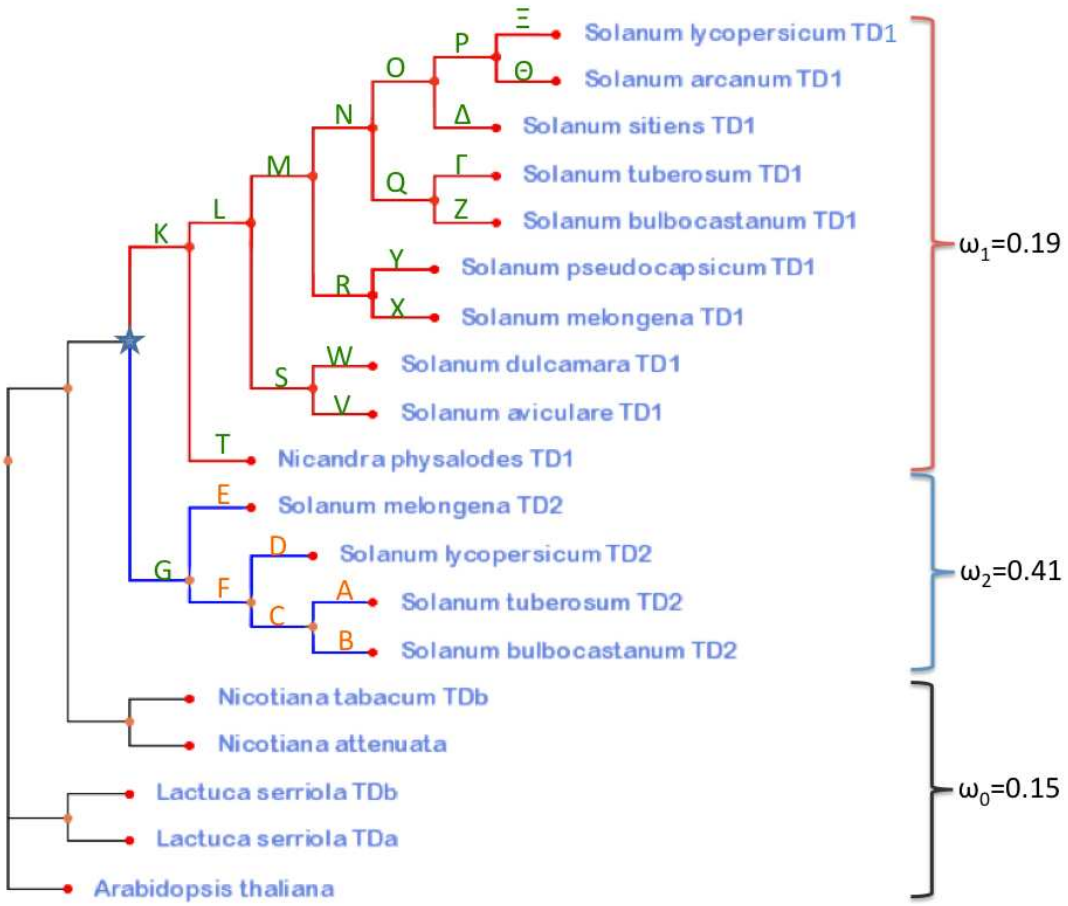


Figure 4. The constrained TD tree used for PAML analysis (Tree 2). Letter and color notations are the same as in Figure 3.

Table 2. Significance tests of 3 branch groups based on two phylogenies in Figure 3 and Figure 4. ω_1, ω_2 and ω_3 represents ω values of branches in group1 (red lines), group 2(blue lines) and group 0 (black lines) respectively. All the tests have degree of freedom equal to one (df=1) and share the same H1, which is the three ω's are pairwise different($\omega_1 \neq \omega_2$, $\omega_0 \neq \omega_1$, $\omega_0 \neq \omega_2$).

Tree	Groups	H_0	$2\Delta\ell$	P Value	Significance
Tree 1	ω_1 versus ω_2	$\omega_1=\omega_2$, $\omega_0\neq\omega_1$, $\omega_0\neq\omega_2$	45.06	0	Yes
	ω_1 versus ω_0	$\omega_1\neq\omega_2$, $\omega_0=\omega_1$, $\omega_0\neq\omega_2$	14.26	0.0001	Yes
	ω_2 versus ω_0	$\omega_1\neq\omega_2$, $\omega_0\neq\omega_1$, $\omega_0=\omega_2$	162.62	0	Yes
Tree 2	ω_1 versus ω_2	$\omega_1=\omega_2$, $\omega_0\neq\omega_1$, $\omega_0\neq\omega_2$	35.61	2.45×10^{-9}	Yes
	ω_1 versus ω_0	$\omega_1\neq\omega_2$, $\omega_0=\omega_1$, $\omega_0\neq\omega_2$	2.95	0.086	No
	ω_2 versus ω_0	$\omega_1\neq\omega_2$, $\omega_0\neq\omega_1$, $\omega_0=\omega_2$	117.94	0	Yes

Table 3. Branches under positive selection based on Tree 1 in Figure 3. All the tests have degree of freedom equal to one (df=1). For tests on single branch, only branches with statistical significance are listed, and all the other branches not listed are insignificant. For tests on cluster of branches, results of all the tests are listed.

Branch	2Δl	P Value[d]	# of Sites[b]	Sites Under Positive Selection[c]
Sub by J[a] (group1)	9.85	0.0008**	1	63P
Sub by K	0.00	0.5	0	Not found
Sub by L	0.00	0.5	0	Not found
Sub by M	0.00	0.5	0	Not found
Sub by S	0.00	0.5	0	Not found
Sub by N	0.00	0.5	0	Not found
Sub by R	0.00	0.5	0	Not found
Sub by O	0.00	0.5	0	Not found
Sub by Q	0.00	0.5	0	Not found
Sub by P	0.00	0.5	0	Not found
Branch J	4.16	0.0207	0	Not found
Sub by I (group2)	0.00	0.5	43	75P,76K,84V,100V,124T,139S,154N,184N,188T,193E,200V,209N,211V,215Q,222T,225V,231D,235Y,253T,300A,304L,315E,319T,332Y,335A,349N,356I,379A,407H,410T,414E,417S,418D,428I,473D,480K,498A,508S,511I,529S,536S,584D,587N
Sub by G	10.29	0.0007**	48	70A,139S,154N,159E,180G,193E,209N,211V,215Q,222T,228S,231D,234T,235Y,241A,243G,253T,267V,276S,280A,281F,285I,315E,332Y,335A,356I,379A,407H,414E,417S,418D,428I,435K,439K,449V,470E,480K,486M,498A,505V,508S,521F,529S,530T,536S,584D,587N,591N

Continued

Sub by F	44.36	0**	28	139S,180G,186T,193E,211V,215Q,222T,231D,280A,285I,315E,335A,379A,407H,414E,428I,439K,449V,480K,498A,521F,527A,584D,585N,586L,587N,588E,592I
Sub by C	48.42	0**	19	55S,92P,139S,186T,193E,211V,222T,227L,231D,232K,297Y,498A,584D,586L,587N,588E,589A,591N,592I
Branch A	9.05	0.0013**	2	193E,232K
Branch B	54.46	0**	12	55S,211V,222T,227L,231D,297Y,584D,587N,588E,589A,591N,592I
Branch C	9.11	0.0013**	6	241A,449V,481L,491F,554G,555N
Branch D	6.24	0.0062**	1	315E
Branch E	13.21	0.0001**	2	190V,435K
Branch F	14.21	0.0001**	6	228S,407H,408K,428I,498A,411I,
Branch G	10.77	0.0005**	1	536S
Branch H	5.23	0.0111*	0	Not found
Branch I	4.25	0.0196*	2	166T,200V

Notes:

a. "Sub by J" represents branch-site model test on all the branches subtended by J (in Figure 3). The following lines are defined in the same way.

b. Number of sites under positive selection under Bayes empirical Bayes (BEB) test (Yang et al. 2005). Only sites with ω larger than 1 and BEB probability larger than 0.95 were considered as positively selected sites.

c. Numbers and letters represent the positions and the names of amino acid residues in *Solanum tuberosum* TD2 respectively.

d. Asterisks following p values show significance level after sequential Bonferroni correction. *** $p<0.005$; ** $p<0.01$; * $p<0.05$.

Table 4. Branches under positive selection based on Tree 2 in Figure 4. Methods and notations are the same as in Table 3.

Branch	2Δl	P Value	# of sites	Sites Under Positive Selection
Sub by K (group1)	7.95	0.0048**	2	63P,430Q
Sub by L	0.00	0.5	0	Not found
Sub by M	1.34	0.1235	0	Not found
Sub by S	0.00	0.5	0	Not found
Sub by N	0.00	0.5	0	Not found
Sub by R	0.00	0.5	0	Not found
Sub by O	0.00	0.5	0	Not found
Sub by Q	0.00	0.5		Not found
Sub by P	0.00	0.5	0	Not found
Sub by G (group2)	0.29	0.59	28	70A,139S,154N,193E,211V,215Q,222T,231D,235Y,280A,315E,332Y,335A,379A,414E,418D,428I,449V,470E,480K,486M,498A,508S,521F,529S,536S,584D,587N
Sub by F	46.17	0***	26	92P,139S,180G,186T,193E,211V,215Q,222T,231D,280A,285I,315E,379A,414E,428I,449V,480K,498A,521F,527A,584D,585N,586L,587N,588E,592I
Sub by C	49.67	0***	14	92P,139S,186T,193E,211V,222T,231D,498A,584D,586L,587N,588E,589A,592I
Branch A	10.14	0.0007***	2	193E,232K
Branch B	53.85	0***	14	186T,211V,222T,227L,231D,297Y,310H,584D,586L,587N,588E,589A,591N,592I
Branch C	8.59	0.0017***	7	124T,251A,449V,481L,491F,529S,554G
Branch D	8.70	0.0016***	2	139S,315E
Branch E	10.67	0.0005***	1	435K
Branch F	20.40	0***	7	228S,335A,349N,407H,408K,428I,498A

3.3.3 Evolution of enzyme stability

According to Gonzales-Vigil et al. (2011), the fact that TD2 in cultivated tomato is more stable than TD1 in the insect gut is likely due to the evolution of three extra ion-pair bonds: 100D-245K, 116E-133R, and 93E-335K (93E, 100D, 116E, 245K and 335K, corresponds to 94E, 101D, 117E, 246K and 336K in *S. tuberosum*, respectively). Of these six amino acid sites, five differ between TD1 and TD2 in the amino acid present, which suggests that substitutions at these sites may have been adaptive and may thus have experienced positive selection. However, none of these sites are among the amino acid sited that are detected to be subject to positive selection (Appendix F). This lack of detection of positive selection may not be surprising, though, if these sites remained under purifying selection after the single adaptive substitution at each site. A more powerful test would be to determine whether dN/dS is > 1 on the branch(es) on which these substitutions occurred. This first requires identifying those branches.

To identify these branches, I reconstructed ancestral states at internal nodes using Fitch Parsimony rules (Fitch 1971) for each of these sites. These reconstructions indicate that the substitutions occurred on different branches. At position 1 (94E), the substitution of amino acid E occurred before the separation of the Solanoidea and the Nicotianoidea (Appendix I, Figures A, B, K and L. The paired substitution at position 2 (336K) did not occur until substantially later, well within the diversification of the genus *Solanum* (Appendix I, Figures C, D, K and L). Thus, the pair bond was not created until

the substitution at position 2, which indicates that the substitution 1 may have been fixed by drift and represents a pre-adaptation.

There is some ambiguity for the location of the substitution at positions 3 (101D). It either occurred on the branch between the node at the *Lycium*-other Solanoidea split and the node at the base of *Solanum* or at the base of the Solanoidea TD2 clade (Appendix I, Figures E, F, K, and L). The paired substitution at position 4 (246K) localizes to just above the *Lycium* branch point and thus may be have occurred on the same branch as the substitution at position 3. It is therefore not possible to determine the order of these two substitutions and hence which is nor likely to have been adaptive.

Finally, the third pair bond was created by substitution at position 5 (117E) because its pair at position 6 (134R) has the same amino acid in all sequences examined. This substitution was thus presumably adaptive and co-localizes on the same branch as the substitutions at positions 3 and 4 (Appendix I, Figures I, J, K and L). Thus, there is a reasonable expectation that substitutions 2 and 5 and either 3 or 4 were adaptive because they created ion-pair bonds.

These four substitutions all co-localize on branch G on Trees 1 and 2 (Figures 3 and 4). I therefore estimated the dN/dS ratio for these sites along branch G in each tree (Table 5). Again, none of these sites have ω that are statistically larger than 1 (i.e., with $Pr(\omega>1)>0.95$), which is consistent with the result that none of these sites were found under positive selection under branch site model (Tables 3 and 4). However, the power

to detect significance on a single branch is limited, and there are hints for positive selection on some of these sites. The strongest possibility is for site 336, which experienced a transition from I to K and completed the evolution of an ion-pair bond. On tree 1 $\omega = 4.95$, with a probability of $\omega > 1.0$ equal to 0.928, while on tree 2 $\omega = 1.10$, with a probability of $\omega > 1.0$ equal to 0.715. Two other sites (246K and 117E) also have $\omega > 1.0$ on Tree 1, though the probability that $\omega > 1.0$ is substantially lower. However, these substitutions at these two sites also completed ion-pair bonds. These results are consistent with the hypothesis that substitutions at 94E, 101D and 134R were the first associated with an ion pair and were fixed by drift, while 117E, 246K and 336K occurred late to complete the ion pairs and were fixed by positive selection. Support for this hypothesis is, however, admittedly weak.

Table 5. Average ω calculated from Bayes Empirical Bayes (BEB) posterior, and the probability that is larger than one, under branch-site model on branch G. Numbers and letters represent the positions and amino acid names in TD2 respectively.

Position in *S. lycopersicum* TD2	Position in *S. lycopersicum* TD2	ω from Tree 1	Pr(ω >1)	ω from Tree 2	Pr(ω >1)
93 E	94E	1.06	0.014	1.01	0.026
100 D	101D	0.20	0.019	0.14	0.035
116 E	117E	1.34	0.215	0.27	0.060
245 K	246K	1.22	0.053	1.00	0.033
335 K	336K	4.95	0.928	1.10	0.715

4. Discussion

This investigation had two major objectives. The first was to identify where in the phylogeny duplication of the TD gene arose and fixed. The second was to determine, by analysis of positive selection, which model of gene duplication the post-duplication fate of this gene best fit. Below I discuss my conclusions regarding these two objectives.

4.1 Gene Tree construction

According to the most likely tree (Tree 1; Figures 1 and 2), most species examined in Solanoideae subfamily have two TD copies that form two distinct clades: the TD1 clade and the TD2 clade. At the same time, only one TD copy (or two copies due to hybridization) is found in the two *Nicotiana* species that have been examined. These patterns are consistent with the TD duplication having occurred in the basal lineage of the Solanoideae. However, it is also possible that the duplication occurred in the common ancestor of the Solanoidea and the Nicotianoidea, but that the two species examined from the latter subfamily subsequently lost one copy (Tree 2). Although a statistical comparison of these alternative hypotheses indicates the first hypothesis should be rejected, I believe that because my analysis included only three *Nicotianoid* sequences, I can not be greatly confident in this conclusions. Further sampling from the

Nicotianoidea will be required to distinguish definitively between these two possibilities.

Several species outside the Solanoideae appear to have two copies of TD (*Petunia axillaris, P. integrifolia, M. guttatus* and *I. nil*). It is not clear whether these truly represent different loci or different alleles at the same locus. If they do represent different loci, it is clear that they represent independent duplications from that which occurred in the Solanoideae/Nicotianoideae since neither copy groups with either the TD1 or the TD2 clade. This pattern suggests that single-copy TD loci may frequently undergo duplication.

4.2 Positive selection

Due to my inability to resolve completely where the duplication of TD occurred within the Solanoideae/Nicotianoideae clade, I examined positive selection on two gene trees, one of which assumed that the duplication occurred at the base of the Solanoideae (Tree 2), while the other assumed the duplication occurred before the split of the two subfamilies (Tree 1).

Both gene trees produced similar results, and lead to the following conclusions:

(1) *Substantial positive selection has occurred on TD2.* With Tree 1 (Figure 3), all sub-branches of TD2 exhibited significant positive selection except the branch leading to *N. attenuata* (branch H). This result provides additional evidence that the position of this species in this tree is inappropriate. In particular, it indicates that TD1 was lost in this

species and that either (a) this species can survive without a core metabolic enzyme, or (b) that TD2 in this species was converted back to perform its ancestral function. The former possibility seems very unlikely, while the latter is inconsistent with my failure to detect positive selection on this branch.

In Tree 2 (Figure 4), the basal branch of the TD2 clade (branch G, Figure 4) exhibits no evidence of positive selection, while all clades and branches above it do. On its face, this pattern seems to imply that most adaptation in TD2 occurred after diversification of the species in this clade. One might expect, therefore, that the TD2 sequences of the different species would exhibit different ways of adapting to the new conditions in the insect gut in which this copy of the enzyme functions. However, I have also shown that the five substitutions that created the presumably stabilizing ion-pair bonds evolved on branch G. While my analysis did not identify any of these sites as being subject to positive selection, this may not be surprising if they each involved only one nucleotide substitution and were subsequently subject to strong purifying selection. Moreover, because my analysis did not examine regulatory substitutions, it is silent about whether any of them may have occurred on branch G. It is thus possible that a number of initially important substitutions that created a proto-TD2 function may well have evolved soon after the gene duplication, before diversification of the TD2 clade. The substitutions that occurred subsequently, and that I detected, would then represent fine-tuning of the TD2 functions.

2. *Little positive selection occurred in the clade corresponding to TD1.* The analysis with Tree 1 revealed elevated ω value and one sites under positive selection, while that with Tree 2 identified no change in the ω value and only two sites under positive selection. This result was expected because this copy of TD retains the ancestral housekeeping functions. It is thus expected to be subject to strong purifying selection. Moreover, because it performs an ancient function, something close to its optimal sequence presumably evolved long before the origin of angiosperms. Substitutions that improve function and thus are fixed by positive selection are thus expected to be very rare, in accordance with my results.

4.3 Gene duplication model discussion

There are three criteria for distinguishing among the three gene duplication models neofunctionalization, subfunctionalization, and escape from adaptive conflict (Table 1): the number of ancestral gene functions, adaptive conflict in ancestral gene and observed positive selection after duplication.

Because I did not reconstruct the ancestral gene sequence and examine the properties of the corresponding enzyme, I was not able to examine the first two criteria. However, my analyses allow me to use the third criteria, which differs for each model.

The subfunctionalization model, as described by Force et al. (1999), assumes that genetic drift alone is responsible for the early evolution of gene duplicates, although it does not preclude the possibility that once the copies have become subfunctionalized,

further adaptive fine-tuning of function may occur (Innan & Kondrashov 2010). My results are not consistent with this expectation. While I found no evidence of selection acting on the TD1 copy, I found that substantial positive selection acted on TD2. While much of this may constitute later fine-tuning, the presumably adaptive substitutions associated with formation of the stabilizing ion-pair bonds were almost surely adaptive and occurred early in the evolution of TD2. This adaptive evolution is inconsistent with the subfunctionalization model.

The Escape from Adaptive Conflict model also does not appear to be consistent with my results. Under this model, one expects positive selection to result in adaptive substitutions in both copies as selection resolves the adaptive conflicts present in the single-copy ancestor (Des Marais and Rausher 2008). While such selection clearly occurred on TD2, I have little evidence of it having occurred on TD1. One caveat to this conclusion is that the type of selection analysis used may fail to identify positive selection at sites where only one or a small number of adaptive substitutions occurred in a clade. Nevertheless, the TD1 clade fails to exhibit the type of "fine-tuning" substitutions seen in the TD2 clade and that would be expected as optimal TD1 function is restored.

The Neofunctionalization model seems most consistent with my results. Under this model, after duplication, one copy of the gene retains its ancestral function and is thus expected to be subject primarily to purifying selection, while the other copy evolves

a novel function and experiences many adaptive substitutions. However, the evolution of TD in the Solanaceae does not conform completely to the canonical Neofunctionalization model. Specifically, the ancestral single-copy TD may well have performed a rudimentary defensive function, simply by virtue of the fact that it deaminates threonine. It is conceivable that when ingested into an insect's gut, this function may have served to partially reduce the amount of threonine available for insect growth, even though it was not stable to degredation under the conditions present in the gut. If this were true, then the "defensive" function of threonine was already present in the ancestral single-copy version of the gene. Evolution of TD2 after duplication would then represent improvement of defensive function rather than the evolution of a completely new function.

It is likely that this is the situation for many duplicated genes that have been classified as having undergone neofunctionalization. Recent work (Khersonsky & Tawfik 2010) shows that many enzymes perform non-cannonical functions, usually at very low levels. When these genes become duplicated, one copy often is selected to improve these non-cannonical functions (Khersonsky & Tawfik 2010), presumably because environmental conditions change to favor use of this function. This process actually combines elements of two of the three major duplication models. The duplicate evolves a new function in the sense that a function that was not previously utilized is enhanced to become an important contributor to organismal fitness. There is thus an

element of neofunctionalization. There is also an element of subfunctionalization, in that one duplicate retains the canonical function, while the other perfects the original non-cannonical function. Because this combination of neo- and sub-functionalization seems likely to be a common fate of duplicate genes, as it appears to be for threonine deaminase in the Solanaceae, it should be recognized as a separate type of post-duplication fates.

4.4 The Tempo of Duplicate Evolution

Most recent analyses of the fates of duplicated genes focus primarily on what model best describes the evolutionary processes following duplication. Another interesting issue, however, is how long it takes for novel functions to be perfected after duplication. My results indicate that fine-tuning of the duplicate copy with novel function continues for a considerable period of time after duplication.

Regardless of whether the duplication occurred prior to the split of the Solanoidea and Nicotianoidea or at the base of the Solanoidea, the last ion-pair bond formed within the genus *Solanum*, after the divergence of the *Solanum* lineage from both *Lycium* and *Jaltomata.* The genus *Lycium* is estimated to be approximately 29.4 years in age (Fukuda *et al.* 2001), while *Solanum* is estimated to have originated approximately 17.95 million years ago (Poczai *et al.* 2011). Consequently, the last ion-pair bond did not evolve until at least 11 million years after the duplication. Moreover, The all branches examined within *Solanum* exhibited evidence of positive selection on TD2, suggesting

that fine-tuning occurred until at most very recently. The split between *S. lycopersicum* and *S. tuberosum* is estimated to have occurred approximately 4 million years ago (Poczai *et al.* 2011), which suggests that fine-tuning of TD2 appears to have occurred over a period of at least 25 million years. It will be of interest to determine whether other neofunctionalized genes require this long to perfect their functions.

Appendix A: PCR primers.

Forward primers	Sequences
Po-TD2-L(begin)	ATGGAATTTCTTTGTTTAGCCCCA
Po-TD2-D1-L	AGGAGGAATTAGCTAAAGGGGTTA
po-TD2-5'1	TTCACGTAACTCCCCTTCTTC
po-TD2-5'2	TTCACGTAACTCCCCTTCTTCT
po-TD2-5'3	GAAGTTCACGTAACTCCCCTTC
po-TD2-5'4	CCCTTCAACTTAGTGCAAGTGAT
po-TD2-5'5	GCAAGTGATCATATCATAAAGAAAAAT
Po-TD1-2L	AATATGATGGCAAAACTCCCTAAA
T-TD1-1L	CCTTCATCTGCAGTAACCCTTG
po-TD1-5'1	CCGCCGGAAAATATCTACATC
po-TD1-5'2	GCCACCGACGGATTTGTA
po-TD1-5'3	ATAACCCTTTGCCAACGACCT
Reverse primers	**Sequences**
Po-TD2-L(end)	AAAAGGCCATTACATTGGATACAT
Po-TD2-D1-R	ACCATATGGCTCAACTCCAATAAT
po-TD2-3'1	AAGCTGATCCGTACAACACG
po-TD2-3'2	CTGATCCGTACAACACGGTTA
po-TD2-3'3	GCCATTACATTGGATACATGC
po-TD2-3'4	TCCGTACAACACGGTTATTTTT
po-TD2-3'5	ACTTAAAAGGCCATTACATTGGA
Po-TD1-2R	CACCAACCACTTTAACAGCTACAC
po-TD1-3'1	CCAAAACAGGCAACACCTTT
po-TD1-3'2	GTCCCAGAACAAGCCAAAAC
po-TD1-3'3	AGTCGGCGAATACCACGTTA

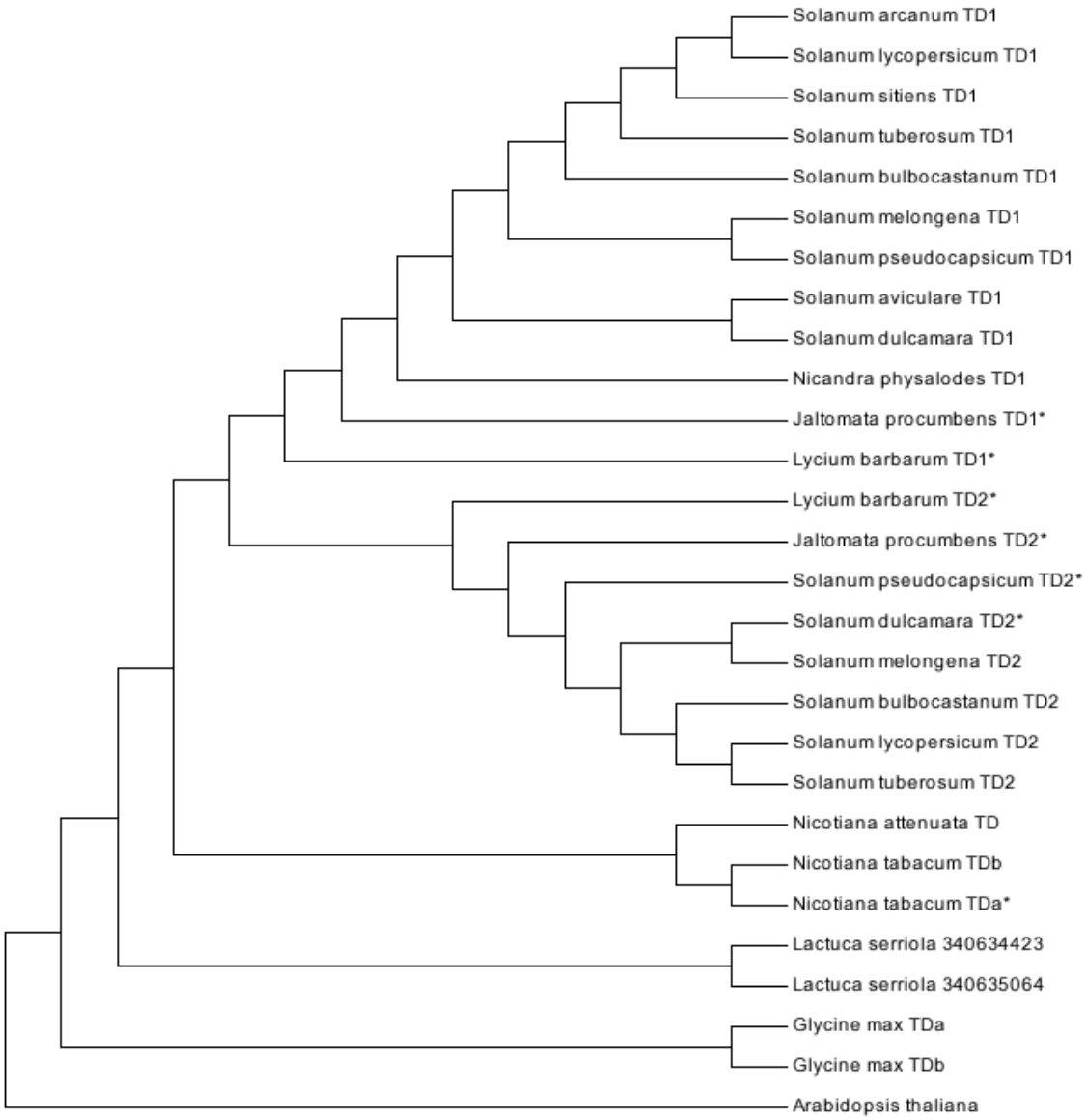
Appendix B: Primer pair abbreviations.

Abbreviation	Primer pairs	Amplified region
1L0	T-TD1-1L and Po-TD1-2R	left CDS of TD1
1L1	po-TD1-5'1 and Po-TD1-2R	left CDS of TD1
1L2	po-TD1-5'2 and Po-TD1-2R	left CDS of TD1
1L3	po-TD1-5'3 and Po-TD1-2R	left CDS of TD1
1R2	Po-TD1-2L and po-TD1-3'2	right CDS of TD1
1R3	Po-TD1-2L and po-TD1-3'3	right CDS of TD1
2L1	po-TD2-5'1 and Po-TD2-D1-R	left CDS of TD2
2L2	po-TD2-5'2 and Po-TD2-D1-R	left CDS of TD2
2L3	po-TD2-5'3 and Po-TD2-D1-R	left CDS of TD2
2L4	po-TD2-5'4 and Po-TD2-D1-R	left CDS of TD2
2L5	po-TD2-5'5 and Po-TD2-D1-R	left CDS of TD2
2R1	Po-Td2-D1-L and po-TD2-3'1	right CDS of TD2
2R2	Po-Td2-D1-L and po-TD2-3'2	right CDS of TD2
2R3	Po-Td2-D1-L and po-TD2-3'3	right CDS of TD2
2R4	Po-Td2-D1-L and po-TD2-3'4	right CDS of TD2
2R5	Po-Td2-D1-L and po-TD2-3'5	right CDS of TD2
TD2_B+E	Po-TD2-L(begin) and Po-TD2-L(end)	full CDS of TD2

Appendix C: Primer pairs used for each TD sequences.

TD sequences	Primer pairs
S. sitiens TD1	1L1,1L2,1R3
S. melongena TD1	1L0,1L1,1L3,1R3
Nicandra physalodes TD1	1L0,1L3,1R3
S. bulbocastanum TD1	1L0,1R3
S. pseudocapsicum TD1	1L0,1R3
S. arcanum TD1	1L0,1R1,1R2,1R3
S. aviculare TD1	1L0,1L1,1L3,1R2,1R3
S. dulcamara TD1	1L1,1R1,1R3
Jaltomata procumbens TD1	1L1,1R2,1R3
Lycium barbarum TD1	1L0,1L1,1L3,1R2
S. bulbocastanum TD2	2L3,2L4,2L5,2R5
S. melongena TD2	TD2_B+E
S. dulcamara TD2	TD2_B+E,2L3,2L5,2R3
Jaltomata procumbens TD2	2R1,2R2
Lycium barbarum TD2	2L3,2L4,2L5,2R1,2R4
S. pseudocapsicum TD2	TD2_B+E,2L5,2R1,2R3

Appendix D: The constrained tree



Appendix E: LRT between the two hypothesis

Likelihood-ration analysis of whether adding constraint to the maximum likelihood tree (Fig. 1) significantly reduces the tree likelihood. The constraint places the Nicotiana TD sequences as an outgroup to the two Solanoidea TD clades (Appendix D). LogL: log-likelihood of tree. Diff LogL: difference in log-likelihood between trees. SH: Shimodaira-Hasegawa test using RELL bootstrap (one-tailed test). wtd-SH: weighted Shimodaira–Hasegawa test. AU: Shimodaira Approximately Unbiased test.

Tree	LogL	Diff LogL	SH	wtd-SH	AU
Unconstrained (Figure 1)	-22200.81	(best)			
Constrained (Appendix D)	-22795.97	595.15	0.0000*	0.0000*	~0*

Appendix F: Positively selected sites detected

Tree 1 (84 amino acid sites in total):

55S,63P*,70A,75P,76K,84V,92P,100V,124T,139S,154N,159E,166T,180G,184N,186T,188T,190V,193E,200V,209N,211V,215Q,222T,225V,227L,228S,231D,232K,234T,235Y,241A,243G,253T,267V,276S,280A,281F,285I,297Y,300A,304L,315E,319T,332Y,335A,349N,356I,379A,407H,408K,410T,411I,414E,417S,418D,428I,439K,449V,470E,473D,480K,481L,486M,491F,498A,505V,508S,511I,521F,527A,529S,530T,536S,554G,555N,584D,585N,586L,587N,588E,589A,591N,592I

Tree 2 (55 amino acid sites in total):

63P*,70A,92P,124T,139S,154N,180G,186T,193E,211V,215Q,222T,227L,228S,231D,232K,235Y,251A,280A,285I,297Y,310H,315E,332Y,335A,349N,379A,407H,408K,414E,418D,428I,430Q*,435K,449V,470E,480K,481L,486M,491F,498A,508S,521F,527A,529S,536S,554G,584D,585N,586L,587N,588E,589A,591N,592I

Sites shared between Tree 1 and Tree 2 (51 sites in total):

63P*,70A,92P,124T,139S,154N,180G,186T,193E,211V,215Q,222T,227L,228S,231D,232K,235Y,280A,285I,297Y,315E,332Y,335A,349N,379A,407H,408K,414E,418D,428I,449V,470E,480K,481L,486M,491F,498A,508S,521F,527A,529S,536S,554G,584D,585N,586L,587N,588E,589A,591N,592I

Sites only detected in Tree 1 (4 sites in total):

251A, 310H, 430Q*, 435K

Sites only detected in Tree 2 (33 sites in total):

55S,75P,76K,84V,100V,159E,166T,184N,188T,190V,200V,209N,225V,234T,241A,243G,253T,267V,276S,281F,300A,304L,319T,356I,410T,411I,417S,439K,473D,505V,511I,530T,555N

Note: Asterisks represent positively selected sites detected in TD1 clade, all the rest shows positively selected sites detected in TD2 clade.

Appendix G: Codon for the six amino acids forming the tree ion pairs.

	First ion pair				Second ion pair				Third ion bond			
	94E[c]		336K[c]		101D[c]		246K[c]		117E[c]		134R[c]	
S. arcanum TD1	GTG	V[a]	CTC	L[a]	AAC	N[a]	GTA	V[a]	CAG	Q[a]	AGA	R[a]
S. lycopersicum TD1	GTG	V	CTG	L	AAC	N	GTA	V	CAG	Q	AGA	R
S. sitiens TD1	GTG	V	CTC	L	AAC	N	GTA	V	CAG	Q	AGA	R
S. tuberosum TD1	GTG	V	CTC	L	AAC	N	GTA	V	CAG	Q	AGA	R
S. bulbocastanum TD1	GTG	V	CTC	L	AAC	N	GTA	V	CAG	Q	AGA	R
S. melongena TD1	TTG	L	ATC	I	AAC	N	GCC	A	CAG	Q	AGA	R
S. pseudocapsicum TD1	TTG	L	ATC	I	AAC	N	GCT	A	CAG	Q	AGA	R
S. aviculare TD1	TTG	L	CTC	L	AAC	N	GTA	V	CAA	Q	AGA	R
S. dulcamara TD1	TTG	L	CTC	L	AAC	N	GTA	V	CAG	Q	AGA	R
N. physalodes TD1	GTG	V	CTC	L	AAC	N	GTA	V	CAG	Q	AGA	R
J. procumbens TD1	AAG	K	ATC	I	AAC	N	GTA	V	CAG	Q	AGA	R
L. barbarum TD1	GTG	V	NA		AGC	S	ATA	I	CAG	Q	AGA	R
N. tabacum TDb	AAG	K	CTC	L	AAT	N	ATA	I	CAG	Q	AGA	R
N. tabacum Tda	GAG	E	CTT	L	GAT	D	GTA	V	CAG	Q	CGA	R
N. attenuata TD	GAA	E	CTT	L	GAA	E	ATA	I	CAA	Q	AGA	R
L. barbarum TD2	GAA	E	NA		GAA	E	ATA	I	CAG	Q	AGA	R
J. procumbens TD2	NA[b]		ATT	I	NA		AAA	K	NA		NA	
S. pseudocapsicum TD2	GAG	E	ATA	I	GAC	D	AAA	K	GAA	E	AGA	R
S. dulcamara TD2	NA		AAA	K	NA		AAA	K	NA		NA	
S. melongena TD2	GAA	E	AAA	K	GAT	D	AAA	K	GAA	E	AGA	R
S. bulbocastanum TD2	GAA	E	AAA	K	GAC	D	AAA	K	GAG	E	AGA	R
S. tuberosum TD2	GAA	E	AAA	K	GAC	D	AAA	K	GAG	E	AGA	R

Continued

S. lycopersicum TD2	GAA	E	AAA	K	GAT	D	AAA	K	GAG	E	AGA	R
L. serriola TDa	GCT	A	TTA	L	AAT	N	ATG	M	TCG	S	CGA	R
L. serriola TDb	CCG	P	TTA	L	AAC	N	TCT	S	CAA	Q	AGA	R
A. thaliana TD	AGC	S	ATA	I	AAT	N	GCT	A	CAA	Q	AGA	R

Notes:

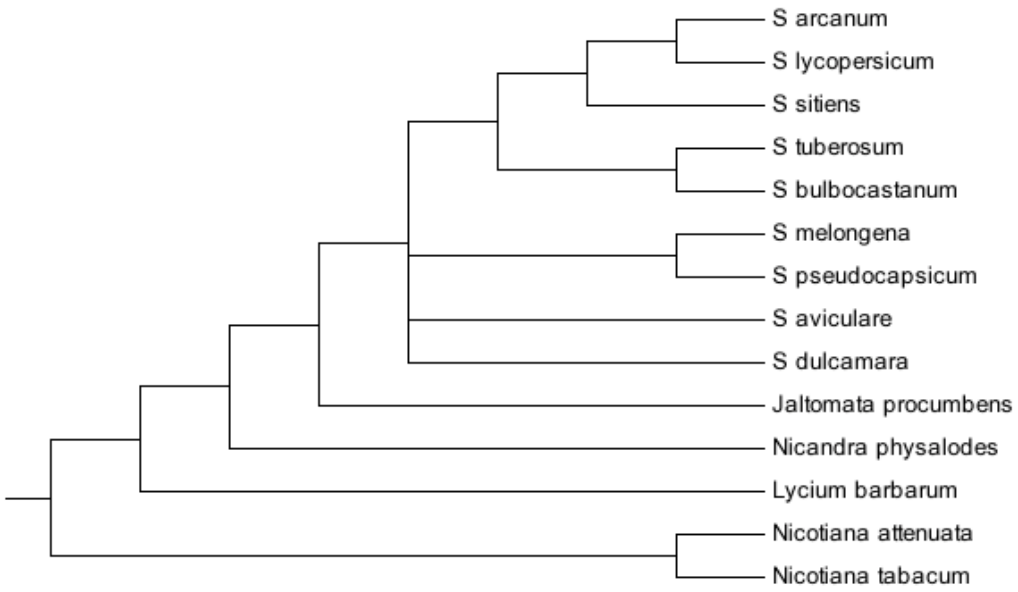
a. Letters in these columns shows the amino acids coded by codons in former columns.

b. NA means not available.

c. Numbers and letters represent the positions and the names of amino acid residues in *Solanum tuberosum* TD2 respectively.

They correspond to 93E-335K, 100D-245K, and 116E-133R in *S. lycopersicum* respectively.

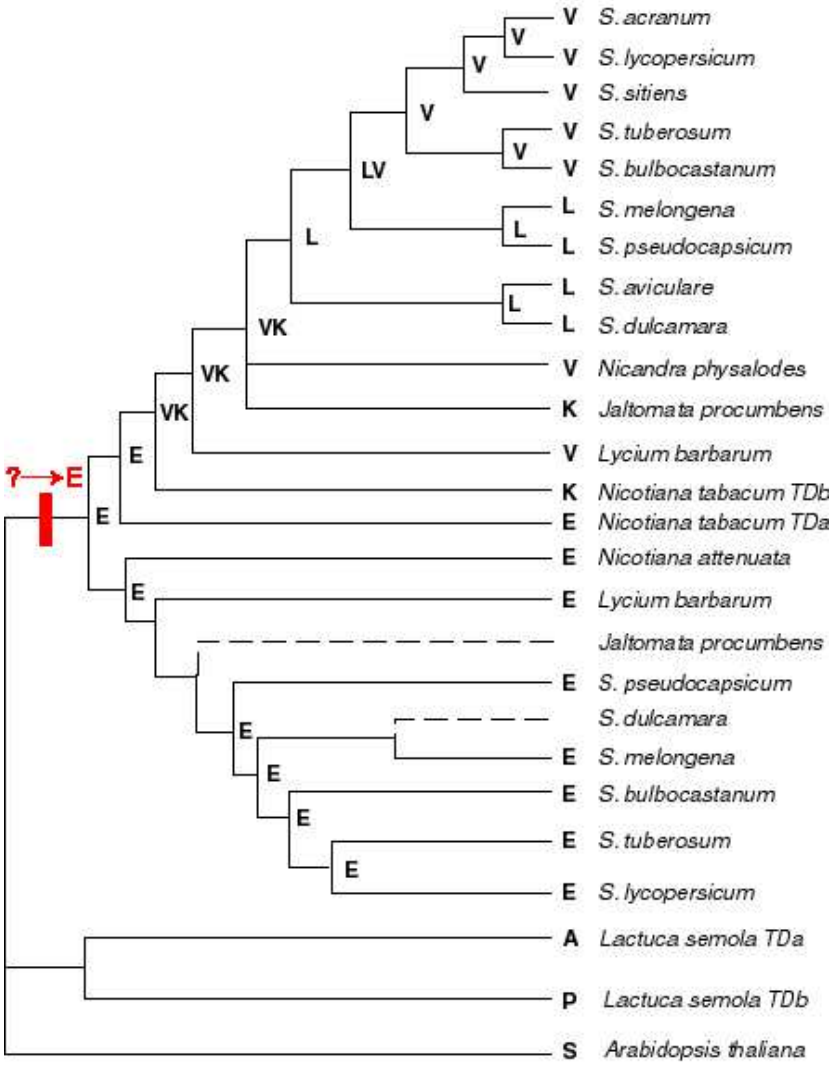
Appendix H: Solanaceae species tree



The species of the Solanaceae species used in this study based on (Olmstead et al. 2008; Bohs & Olmstead 1997; Rodriguez et al. 2009).

Appendix I : Ancestral state reconstructions for the five amino-acid positions.

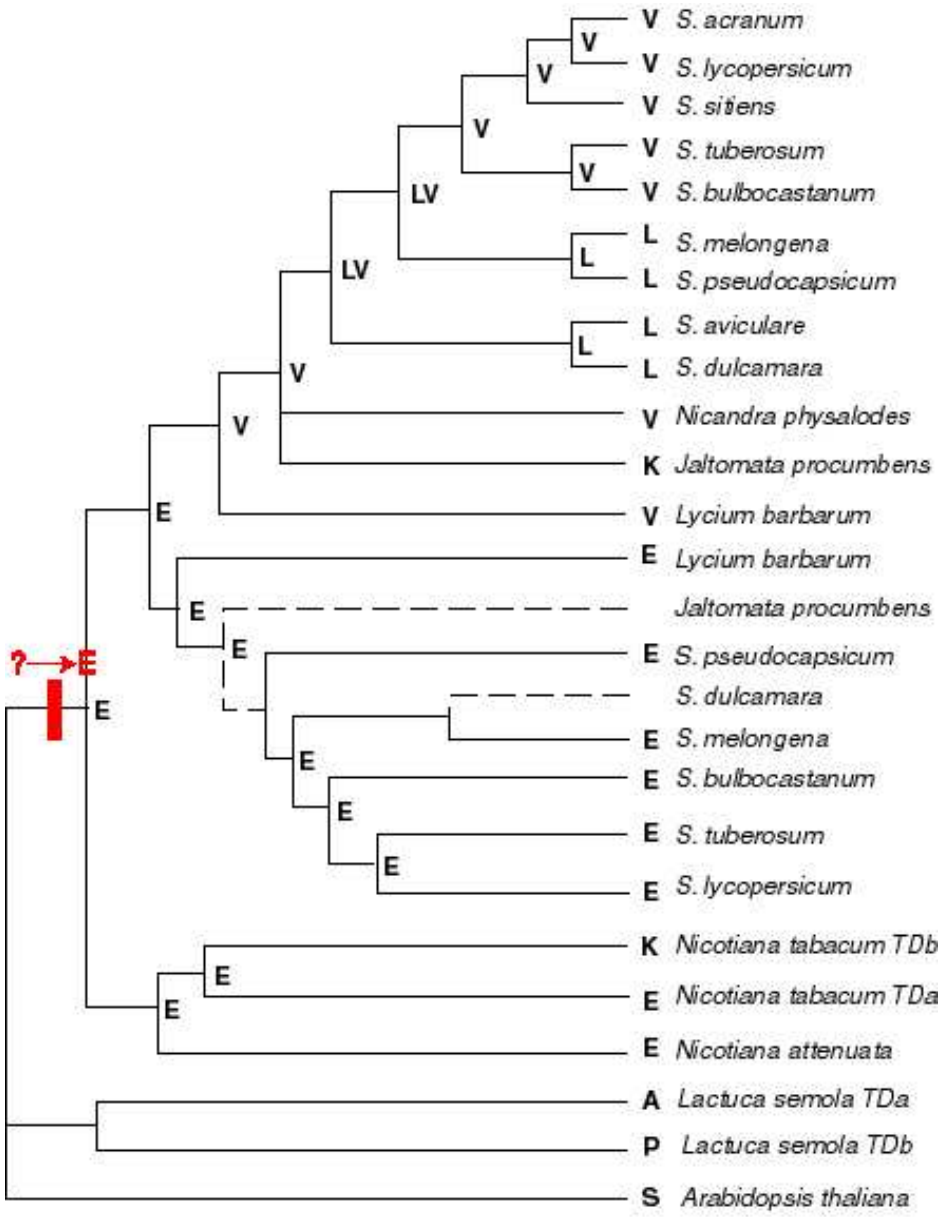
Figure A1. Ancestral state reconstructions for the five amino-acid positions at which substitutions created ion-pair bonds. Positions 1 – 5 refer to 94E, 336K, 101D, 246K, and 117E, respectively, in *Solanum tuberosum* TD2. A,C,E,G,I: reconstructions of positions 1 – 5 on Tree 1. B,D,F,H, J: reconstructions of positions 1 – 5 on Tree 2. Letters at terminal nodes indicate amino acids present in indicated species. Letters at internal nodes indicate amino acids inferred by Fitch Parsimonay rules. Red bars indicate branches on which substitutions at indicated positions (in red) are inferred to have occurred. Stippled bars indicate alternative inferences. Double letters or question mark for ancestral state indicate uncertainty. K,L: Inferred positions on Trees 1 and 2 of all five substitutions.



Tree 1

Position 1

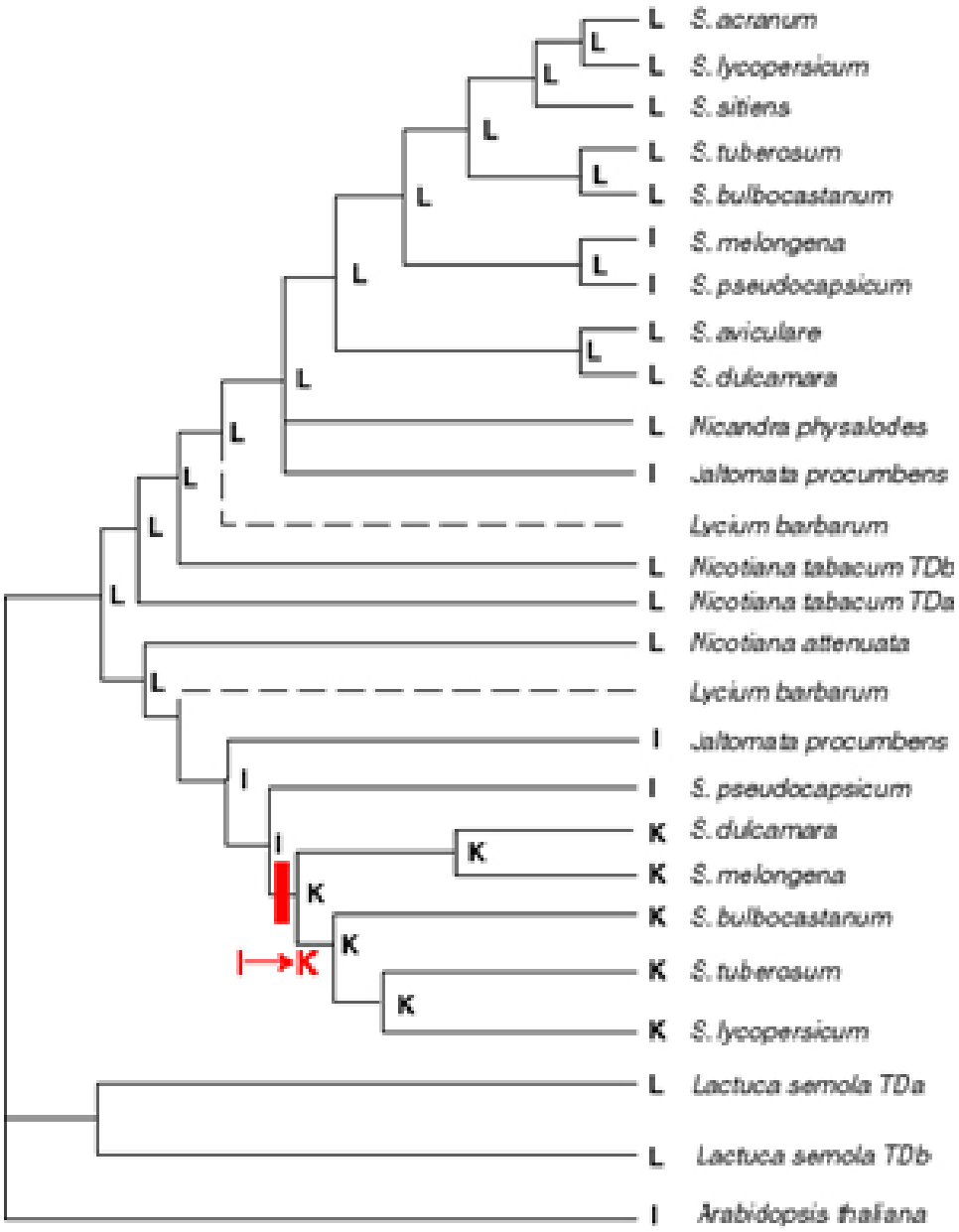
Appendix I. A



Tree 2

Position 1

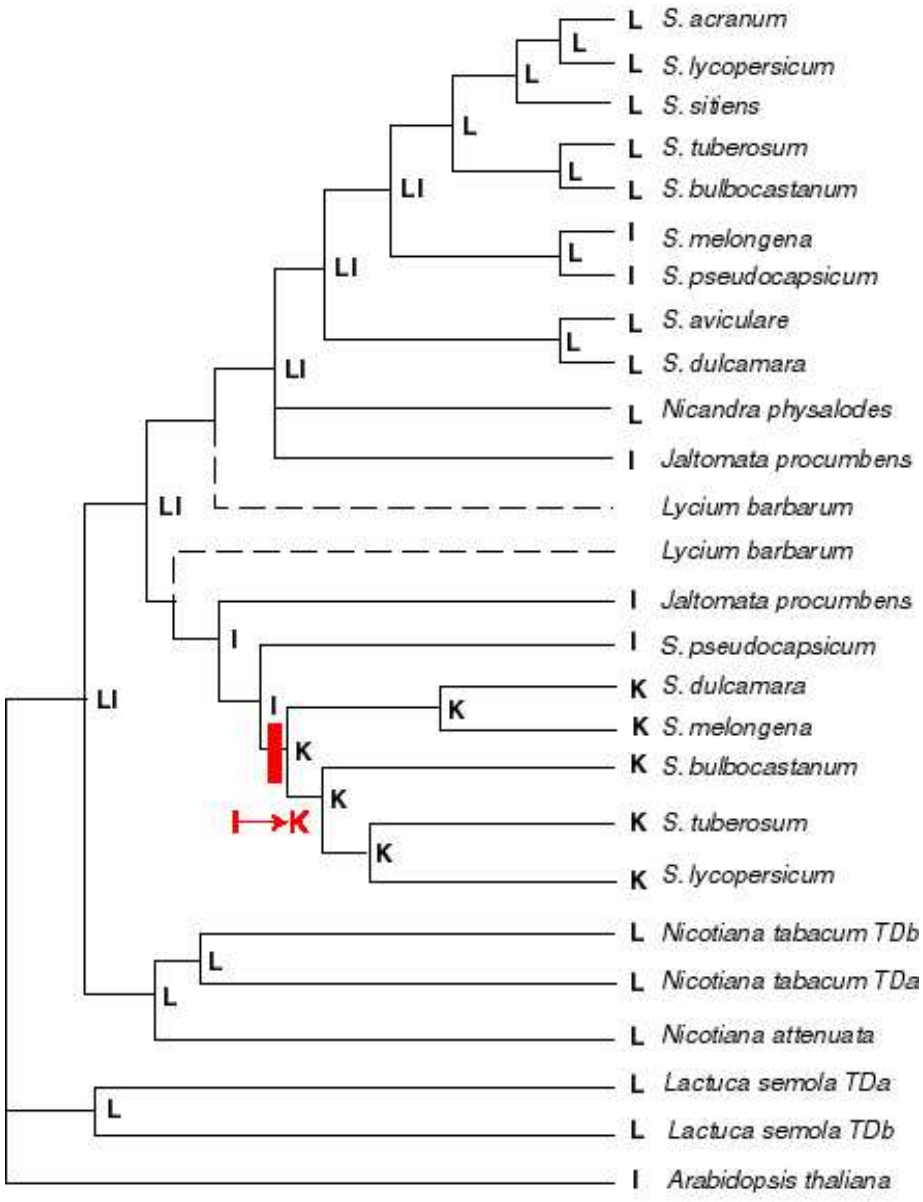
Appendix I. B



Tree 1

Position 2

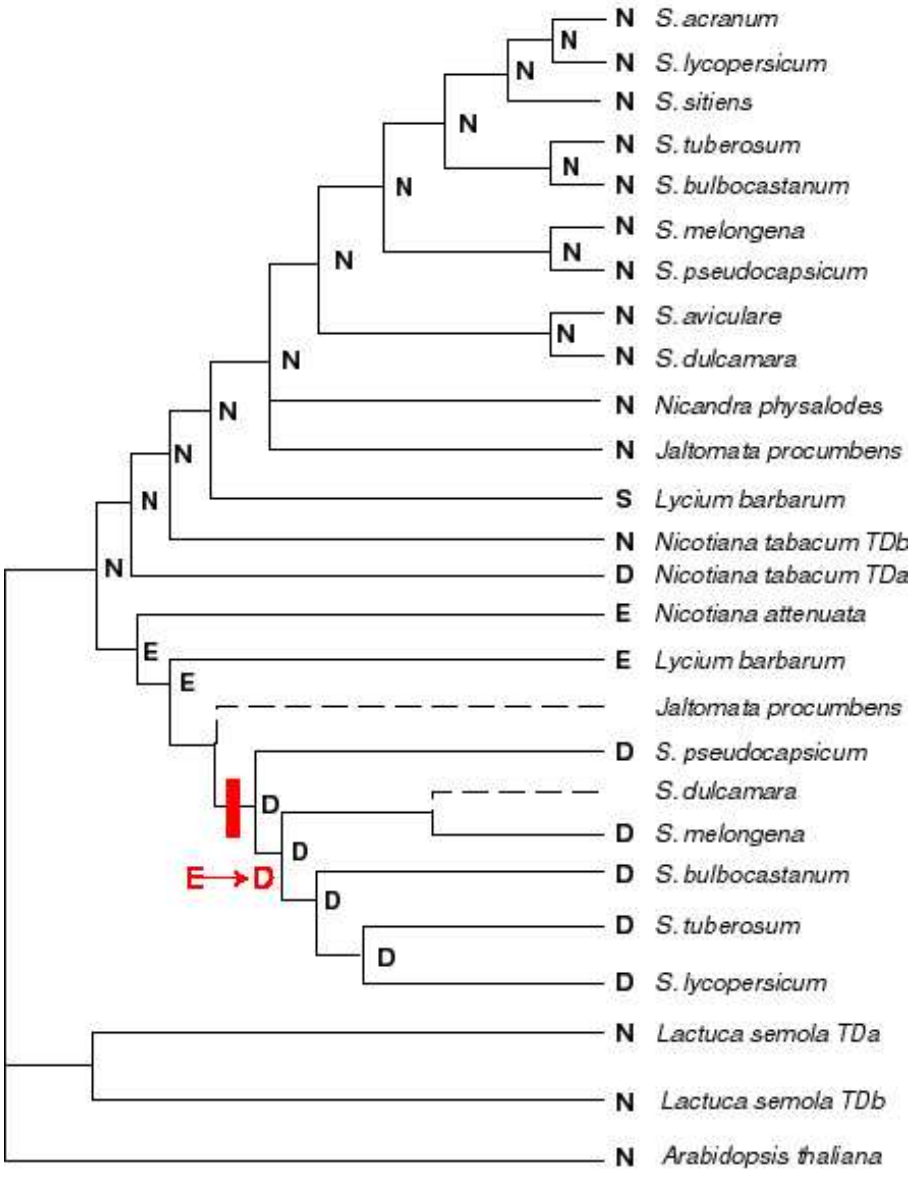
Appendix I. C



Tree 2

Position 2

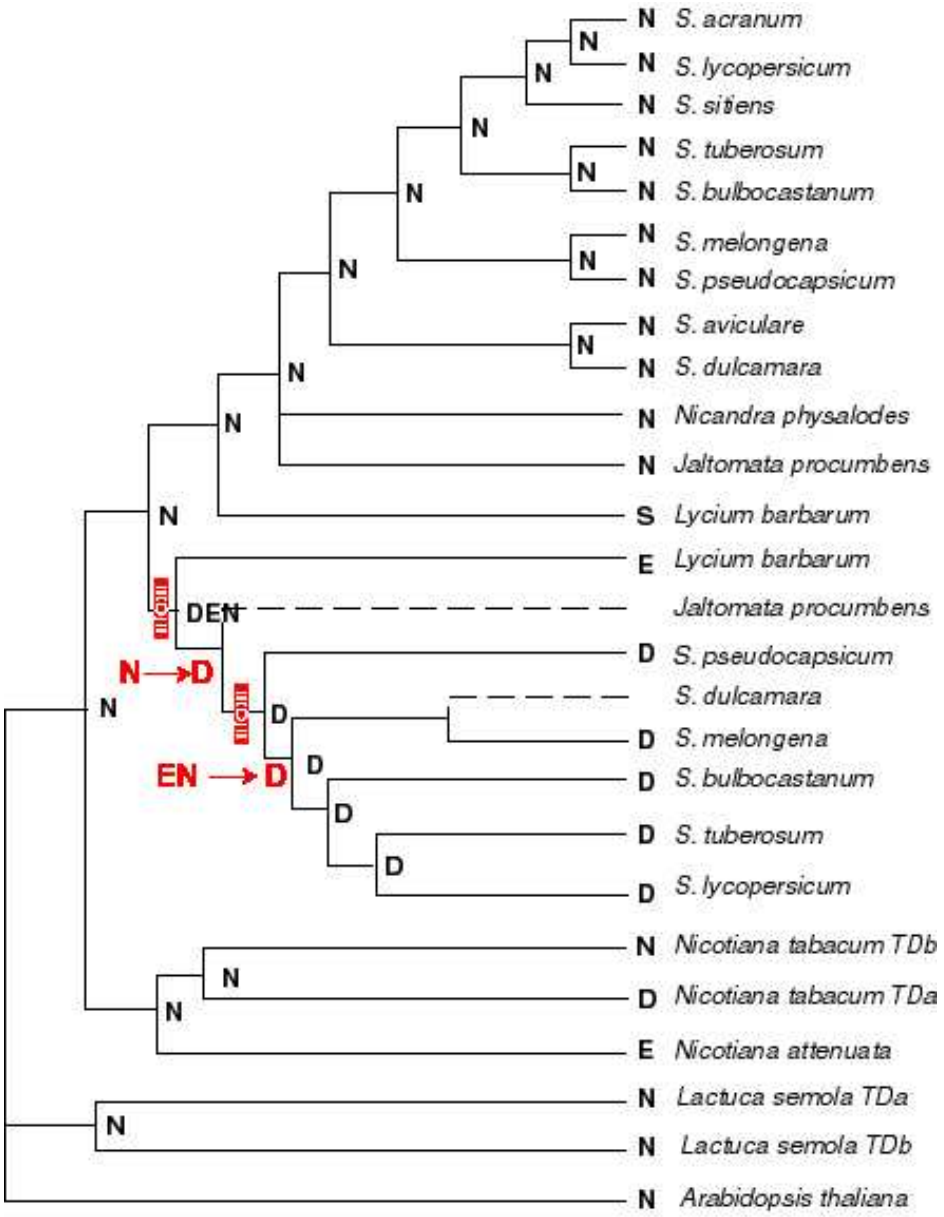
Appendix I. D



Tree 1

Position 3

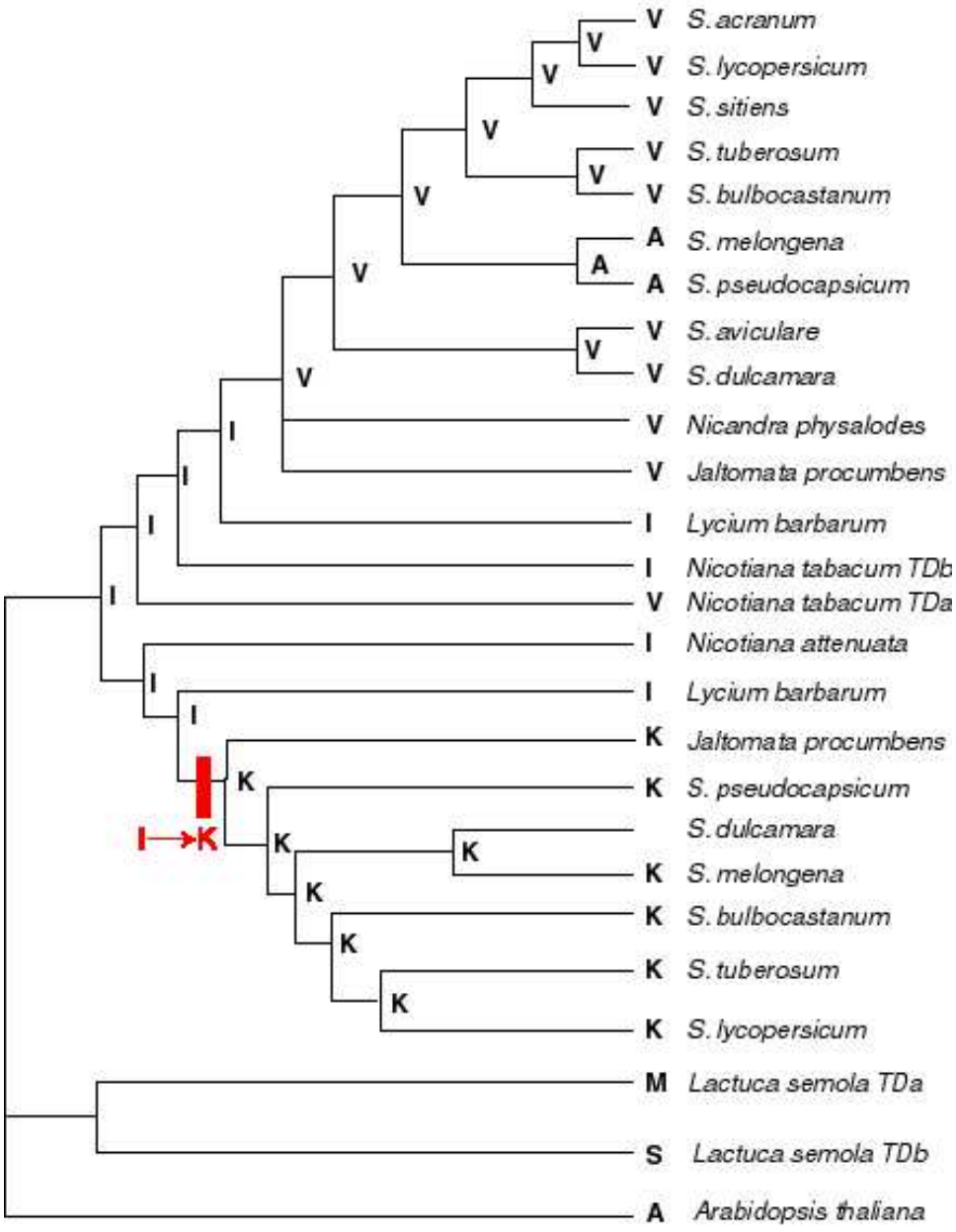
Appendix I. E



Tree 2

Position 3

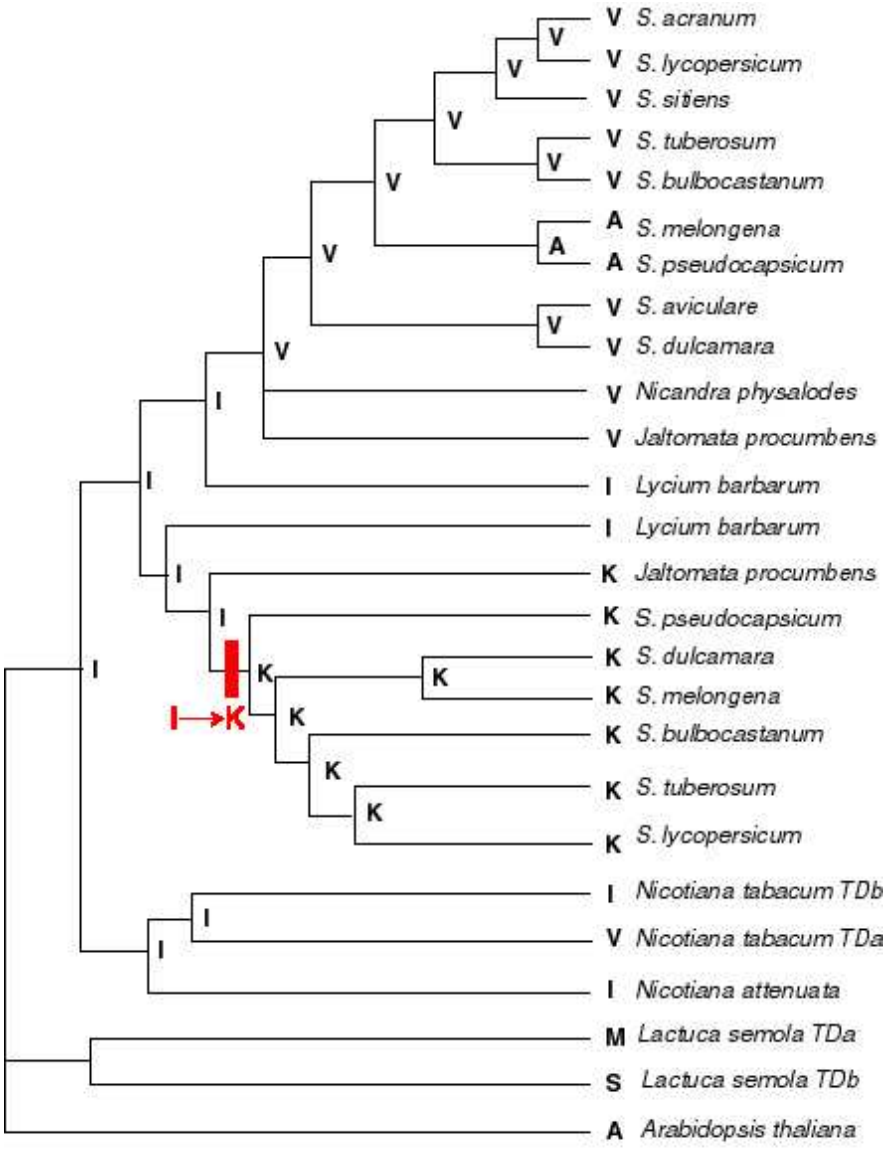
Appendix I. F



Tree 1

Position 4

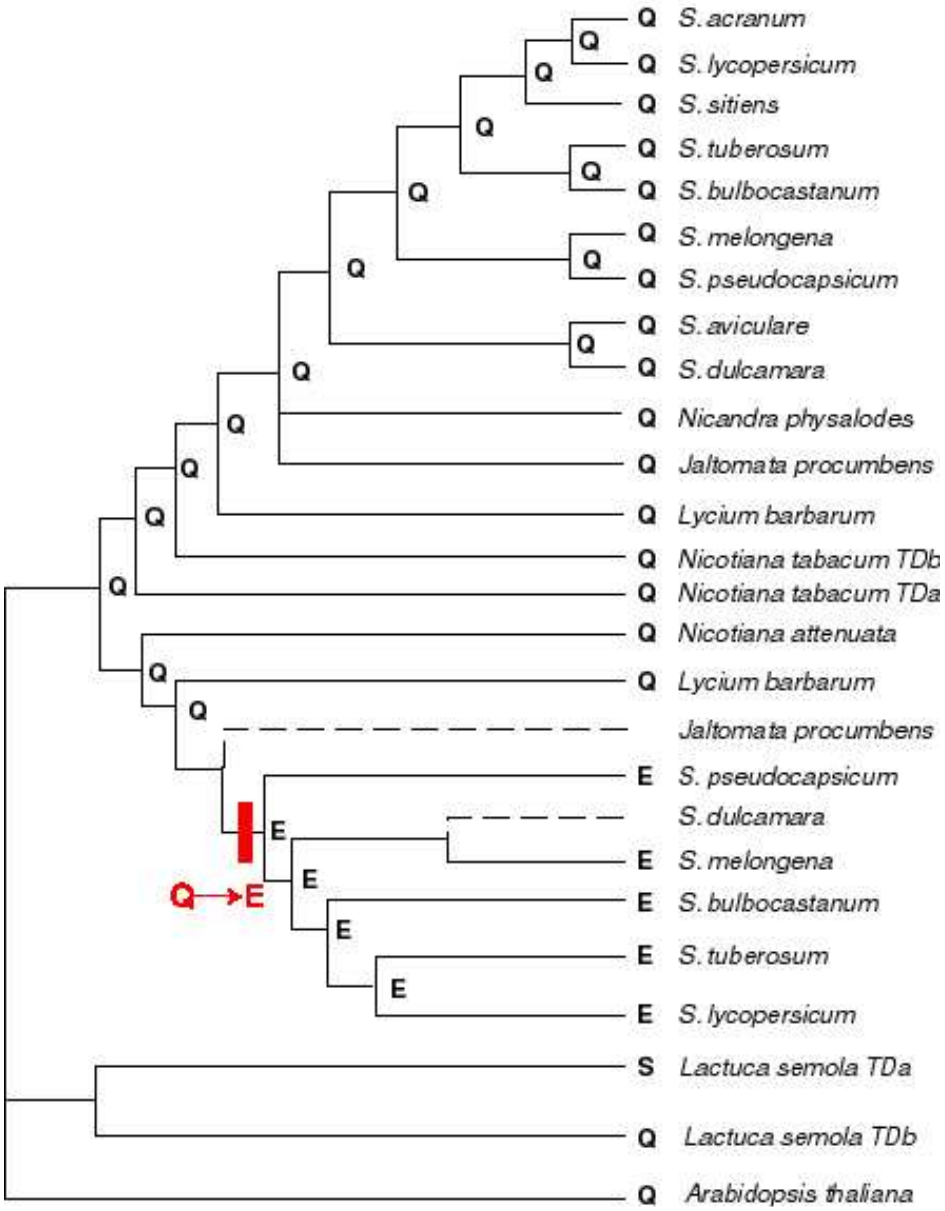
Appendix I. G



Tree 2

Position 4

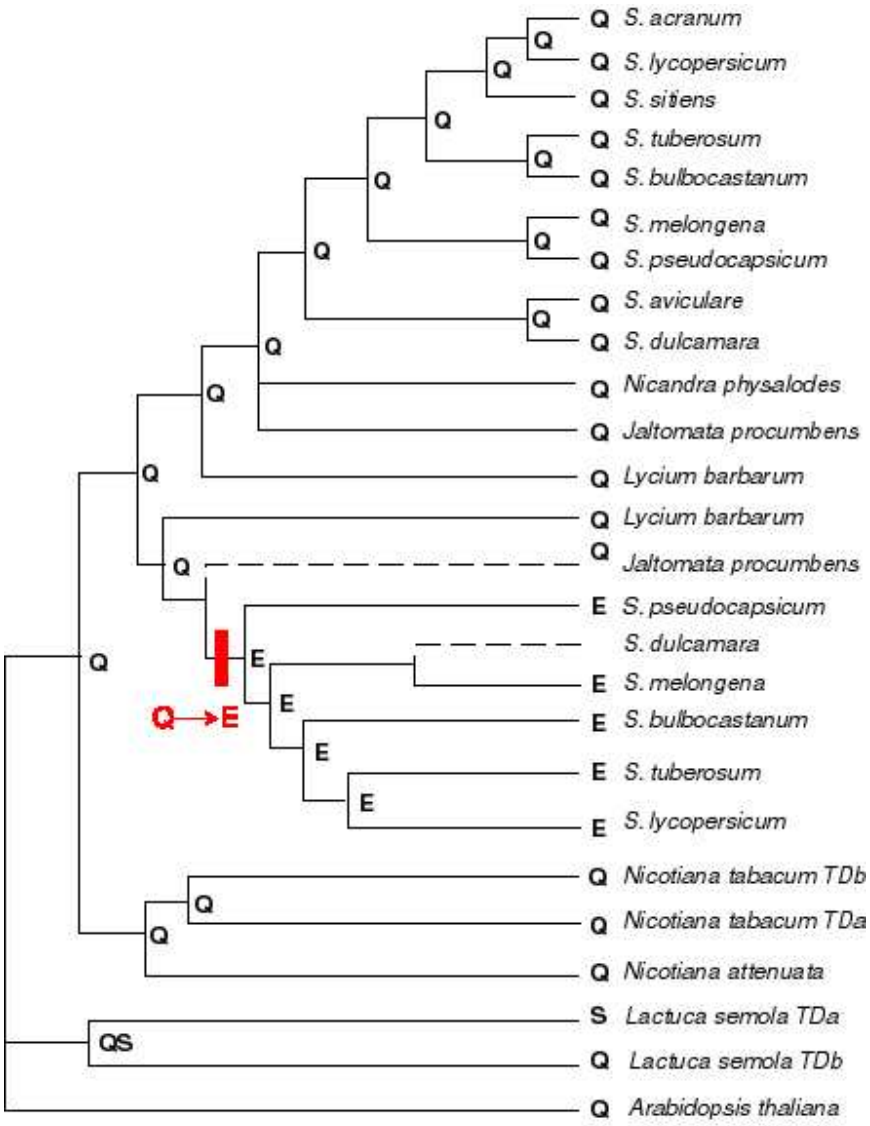
Appendix I. H



Tree 1

Position 5

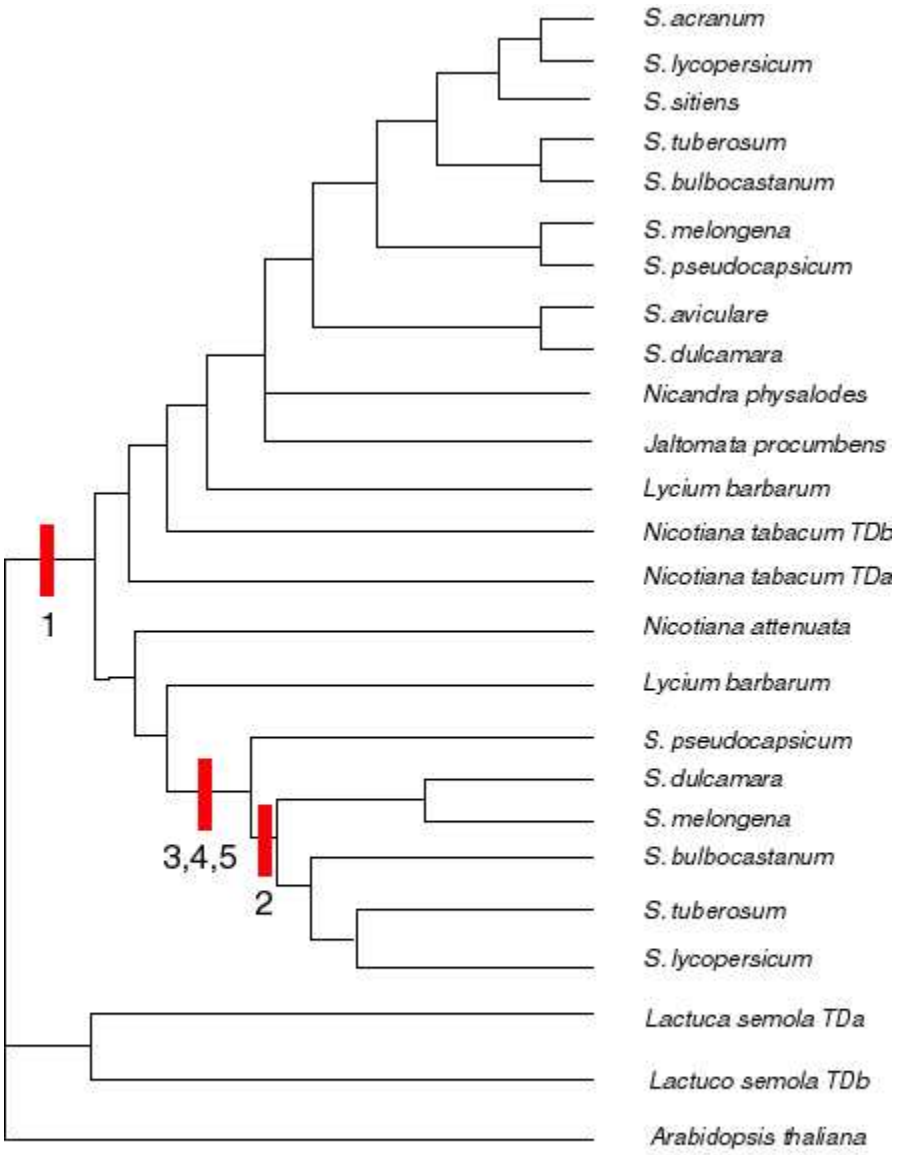
Appendix I. I



Tree 2

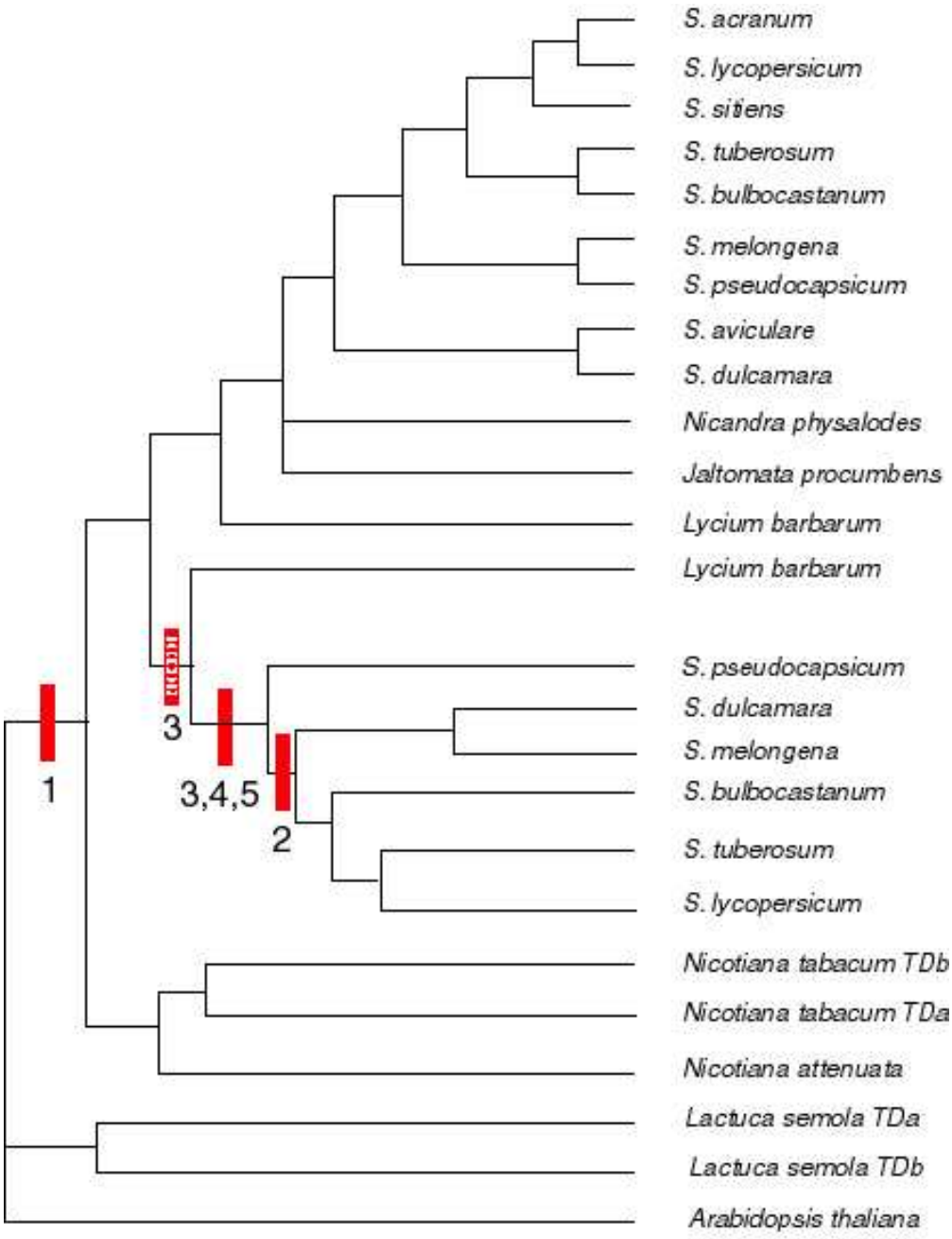
Position 5

Appendix I. J



Tree 1

Appendix I. K



Tree 2

Appendix I. L

Appendix J: TD trees under different methods and models.

1. Method=ML, Model=JC, Gap/missing data treatment=complete deletion, Log Likelihood=-5497.51,
Tree=((((((((((((Solanum_arcanum_TD1:0.01188016,Solanum_lycopersicum_TD1:0.00514437)1.0000:0.00777926,Solanum_sitiens_TD1:0.00361882)0.7000:0.00026077,Solanum_tuberosum_TD1:0.00815845)0.4200:0.00000000,Solanum_bulbocastanum_TD1:0.00777485)0.9100:0.00663755,(Solanum_melongena_TD1:0.02021587,Solanum_pseudocapsicum_TD1:0.00996930)0.4500:0.00412535)0.4100:0.00000010,(Solanum_aviculare_TD1:0.01028061,Solanum_dulcamara_TD1:0.01224407)0.6500:0.00747549)0.7500:0.00511580,Nicandra_physalodes_TD1:0.01033415)0.5200:0.00000010,Jaltomata_procumbens_TD1*:0.01195153)0.7300:0.00827384,Lycium_barbarum_TD1*:0.01990663)0.6800:0.01581219,Nicotiana_tabacum_TDb:0.03048244)0.9400:0.01821581,Nicotiana_tabacum_TDa*:0.09203134)0.6700:0.01493781,(Nicotiana_attenuata_TD:0.06574482,(Lycium_barbarum_TD2*:0.14946617,(Jaltomata_procumbens_TD2*:0.03669641,(Solanum_pseudocapsicum_TD2*:0.12473224,((Solanum_dulcamara_TD2*:0.07764018,Solanum_melongena_TD2:0.07779381)0.8900:0.01932462,(Solanum_bulbocastanum_TD2:0.07132798,(Solanum_lycopersicum_TD2:0.05490253,Solanum_tuberosum_TD2:0.02916445)0.9700:0.02200214)0.5000:0.01917520)0.6200:0.01254518)1.0000:0.04081549)0.9400:0.03509267)0.8500:0.01447048)1.0000:0.11619330)0.9900:0.04782399,(Lactuca_serriola_340634423:0.12609862,Lactuca_serriola_340635064:0.11802694)1.0000:0.07197458)0.6200:0.01068024,(Glycine_max_TDa:0.01409217,Glycine_max_TDb:0.03233965)1.0000:0.11371211,Arabidopsis_thaliana:0.14540756);

2. Method=ML, Model=JC, Gap/missing data treatment=use all sites, Log Likelihood=-22777.33,
Tree=((((((((((((Solanum_arcanum_TD1:0.00590068,Solanum_lycopersicum_TD1:0.00532335)1.0000:0.00687592,Solanum_sitiens_TD1:0.00793347)0.9600:0.00157799,(Solanum_bulbocastanum_TD1:0.00776218,Solanum_tuberosum_TD1:0.00762186)0.7000:0.00247215)1.0000:0.01885876,Solanum_melongena_TD1:0.04064171)0.5000:0.00000010,Solanum_dulcamara_TD1:0.01773452)0.4900:0.00000011,Solanum_aviculare_TD1:0.01692638)0.5500:0.00418187,(Jaltomata_procumbens_TD1*:0.02527945,Solanum_pseudocapsicum_TD1:0.01839806)0.9000:0.00926718)0.9700:0.00536758,Nicandra_physalodes_TD1:0.02799474)0.8900:0.00000010,Lycium_barbarum_TD1*:0.10563695)0.8400:0.02221636,Nicotiana_tabacum_TDb:0.04544833)1.0000:0.06516175,Nicotiana_tabacum_TDa*:0.14562744)0.5800:0.02609214,(Nicotiana_attenuata_TD:0.09874137,(Lycium_barbarum_TD2*:0.17019469,(Jaltomata_procumbens_TD2*:0.01350148,(((Solanum_lycopersicum_TD2:0.08379444,Solanum_tuberosum_TD2:0.03816199)0.7500:0.00253852,Solanum_bulbocastanum_TD2:0.04644996)1.0000:0.08097209,(Solanum_pseudocapsicum_TD2*:0.14304144,(Solanum_dulcamara_TD2*:

0.09706961,Solanum_melongena_TD2:0.10143460)0.5100:0.00023634)1.0000:0.02321045)1.0000:0.06764565)0.9800:0.00000010)1.0000:0.06479624)1.0000:0.11711864)1.0000:0.05220830,(Lactuca_serriola_340634423:0.15775620,Lactuca_serriola_340635064:0.15334619)1.0000:0.06383880)1.0000:0.03419873,(Glycine_max_TDa:0.02205107,Glycine_max_TDb:0.02714736)1.0000:0.14693472,Arabidopsis_thaliana:0.20556998);

3. Method=ML, Model=JC, Gap/missing data treatment=partial deletion(site coverage cutoff-95%), Log Likelihood-7193.28, Tree=((((((((((((Solanum_arcanum_TD1:0.00912314,Solanum_lycopersicum_TD1:0.00564866)0.9900:0.00844819,Solanum_sitiens_TD1:0.00632274)0.6400:0.00000010,Solanum_tuberosum_TD1:0.00634318)0.4700:0.00000000,Solanum_bulbocastanum_TD1:0.00716182)0.9900:0.00961234,(Solanum_melongena_TD1:0.01873299,Solanum_pseudocapsicum_TD1:0.01250355)0.6500:0.00605118)0.6500:0.00000010,(Solanum_aviculare_TD1:0.00956972,Solanum_dulcamara_TD1:0.01096758)0.4700:0.00880087)0.8800:0.00629277,Jaltomata_procumbens_TD1*:0.01665835)0.3800:0.00000020,Nicandra_physalodes_TD1:0.01347236)0.9000:0.01024969,Lycium_barbarum_TD1*:0.07128627)0.9600:0.01472598,Nicotiana_tabacum_TDb:0.02783747)0.9800:0.02273380,Nicotiana_tabacum_TDa*:0.09285317)0.4900:0.01645996,(Nicotiana_attenuata_TD:0.07127298,(Lycium_barbarum_TD2*:0.13925207,(Jaltomata_procumbens_TD2*:0.04286998,(Solanum_pseudocapsicum_TD2*:0.12487519,((Solanum_dulcamara_TD2*:0.13665357,Solanum_melongena_TD2:0.07401802)0.8200:0.01928303,(Solanum_bulbocastanum_TD2:0.06406576,(Solanum_lycopersicum_TD2:0.05751240,Solanum_tuberosum_TD2:0.02467377)0.9900:0.01955468)0.9100:0.02477991)0.5500:0.01550845)1.0000:0.03578169)0.9700:0.04292617)0.8800:0.01482182)1.0000:0.10405742)1.0000:0.05323334,(Lactuca_serriola_340634423:0.10578725,Lactuca_serriola_340635064:0.11532416)1.0000:0.06391968)0.7400:0.01113080,(Glycine_max_TDa:0.01908556,Glycine_max_TDb:0.02680462)1.0000:0.10283586,Arabidopsis_thaliana:0.13989467);

4. Method=ML, Model=Kimura-2-parameter, Gap/missing data treatment=complete deletion, Log Likelihood=-5426.97, Tree=((((((((((((Solanum_arcanum_TD1:0.01170175,Solanum_lycopersicum_TD1:0.00559537)0.9600:0.00789225,Solanum_sitiens_TD1:0.00379719)0.5700:0.00025505,Solanum_tuberosum_TD1:0.00815408)0.3100:0.00000000,Solanum_bulbocastanum_TD1:0.00763209)0.8400:0.00666511,(Solanum_melongena_TD1:0.02011823,Solanum_pseudocapsicum_TD1:0.00994791)0.3900:0.00401338)0.3600:0.00000010,(Solanum_aviculare_TD1:0.01029151,Solanum_dulcamara_TD1:0.01226796)0.5400:0.00723591)0.7200:0.00489258,Nicandra_physalodes_TD1:0.01045129)0.4600:0.00003585,Jaltomata_procumbens_TD1*:0.01609286)0.7800:0.00830500,Lycium_barbarum_TD1*:0.01994638)0.7300:0.01590466,Nicotiana_tabacum_TDb:0.03039422)0.9200:0.01854685,Nicotiana_tabacum_TDa*:0.09211841)0.5100:0.01227124,(Nicotiana_attenuata_TD:0.06583507,(Lycium_barbarum_TD2*:0.15055858,(Jaltomata_procumbens_TD2*:0.03803346,(Solanum_pseudocapsicum_TD2*:0.12706519,((Solanum_d

ulcamara_TD2*:0.07759788,Solanum_melongena_TD2:0.07876980)0.9200:0.02009931,(Solanum_bulbocastanum_TD2:0.07203734,(Solanum_lycopersicum_TD2:0.05422986,Solanum_tuberosum_TD2:0.02964367)0.9700:0.02234310)0.4800:0.01893765)0.6400:0.01307057)1.0000:0.04280945)0.9100:0.03442748)0.7900:0.01608673)1.0000:0.12053444)1.0000:0.05255807,(Lactuca_serriola_340634423:0.12649997,Lactuca_serriola_340635064:0.11709995)1.0000:0.07308072)0.5500:0.00942477,(Glycine_max_TDa:0.01499890,Glycine_max_TDb:0.03190359)1.0000:0.11343858,Arabidopsis_thaliana:0.14702089);

5. Method=ML, Model=Kimura-2-parameter, Gap/missing data treatment=Use all sites, Log Likelihood=-22524.03,
Tree=((((((((((((Solanum_arcanum_TD1:0.00584923,Solanum_lycopersicum_TD1:0.00541503)1.0000:0.00688634,Solanum_sitiens_TD1:0.00791212)0.9800:0.00172253,(Solanum_bulbocastanum_TD1:0.00776765,Solanum_tuberosum_TD1:0.00764244)0.7700:0.00233472)1.0000:0.01874949,Solanum_melongena_TD1:0.04081259)0.5700:0.00000010,Solanum_dulcamara_TD1:0.01771099)0.5500:0.00000011,Solanum_aviculare_TD1:0.01661666)0.6300:0.00449307,(Jaltomata_procumbens_TD1*:0.02507398,Solanum_pseudocapsicum_TD1:0.01877991)0.9500:0.00872085)0.9900:0.00512829,Nicandra_physalodes_TD1:0.02816727)0.9200:0.00000010,Lycium_barbarum_TD1*:0.10539289)1.0000:0.02414309,Nicotiana_tabacum_TDb:0.04369830)1.0000:0.08748748,(Nicotiana_tabacum_TDa*:0.15346383,(Nicotiana_attenuata_TD:0.09792782,(Lycium_barbarum_TD2*:0.16887161,(Jaltomata_procumbens_TD2*:0.01749325,(((Solanum_lycopersicum_TD2:0.08343857,Solanum_tuberosum_TD2:0.03525761)0.9800:0.00294910,Solanum_bulbocastanum_TD2:0.04596605)1.0000:0.08296287,(Solanum_dulcamara_TD2*:0.05799213,(Solanum_melongena_TD2:0.09154400,Solanum_pseudocapsicum_TD2*:0.13247642)0.5600:0.02480098)1.0000:0.01552946)1.0000:0.07168652)1.0000:0.00000010)1.0000:0.06903472)1.0000:0.13484918)0.9500:0.00000010)1.0000:0.05845969,(Lactuca_serriola_340634423:0.16031411,Lactuca_serriola_340635064:0.15293330)1.0000:0.06435815)1.0000:0.03517101,(Glycine_max_TDa:0.02230027,Glycine_max_TDb:0.02673058)1.0000:0.14875675,Arabidopsis_thaliana:0.19729963);

6. Method=ML, Model=Kimura-2-parameter, Gap/missing data treatment=partial deletion(site coverage cutoff-95%), Log Likelihood=-7112.4,
Tree=((((((((((((Solanum_arcanum_TD1:0.00895601,Solanum_lycopersicum_TD1:0.00593581)0.9800:0.00853564,Solanum_sitiens_TD1:0.00631679)0.6900:0.00000010,Solanum_tuberosum_TD1:0.00634111)0.5500:0.00000000,Solanum_bulbocastanum_TD1:0.00705562)1.0000:0.00965148,(Solanum_melongena_TD1:0.01856005,Solanum_pseudocapsicum_TD1:0.01242519)0.5900:0.00596671)0.5400:0.00000010,(Solanum_aviculare_TD1:0.00956788,Solanum_dulcamara_TD1:0.01097399)0.5200:0.00868068)0.8800:0.00624076,Jaltomata_procumbens_TD1*:0.01661947)0.5600:0.00000010,Nicandra_physalodes_TD1:0.01341650)0.9200:0.01043389,Lycium_barbarum_TD1*:0.07174730)0.9600:0.01485136,Nicotiana_tabacum_TDb:0.02779987)1.0000:0.02298468,Nicotiana_tabacum_TDa*:0.09324352)0.4500:0.0140352

3,(Nicotiana_attenuata_TD:0.07072141,(Lycium_barbarum_TD2*:0.14029068,(Jaltomata_procumbens_TD2*:0.04257097,(Solanum_pseudocapsicum_TD2*:0.12678369,((Solanum_dulcamara_TD2*:0.13751124,Solanum_melongena_TD2:0.07431367)0.9500:0.02074737,(Solanum_bulbocastanum_TD2:0.06457373,(Solanum_lycopersicum_TD2:0.05732389,Solanum_tuberosum_TD2:0.02451792)1.0000:0.01994830)0.8500:0.02452244)0.6100:0.01709043)1.0000:0.03814734)0.9000:0.04361559)0.9500:0.01706564)1.0000:0.10919088)1.0000:0.05795990,(Lactuca_serriola_340634423:0.10588592,Lactuca_serriola_340635064:0.11470029)1.0000:0.06588417)0.6800:0.00999183,(Glycine_max_TDa:0.01967807,Glycine_max_TDb:0.02655426)1.0000:0.10242505,Arabidopsis_thaliana:0.14062470);

7. Method=ML, Model=Tamura-3-parameter, Gap/missing data treatment=complete deletion, Log Likelihood=-5412.82, Tree=((((((((((((Solanum_arcanum_TD1:0.01170565,Solanum_lycopersicum_TD1:0.00560442)0.9500:0.00790505,Solanum_sitiens_TD1:0.00383060)0.5600:0.00025517,Solanum_tuberosum_TD1:0.00815314)0.3000:0.00000000,Solanum_bulbocastanum_TD1:0.00762602)0.8000:0.00666348,(Solanum_melongena_TD1:0.02001890,Solanum_pseudocapsicum_TD1:0.00993834)0.4200:0.00401398)0.4500:0.00000010,(Solanum_aviculare_TD1:0.01024858,Solanum_dulcamara_TD1:0.01225571)0.4800:0.00723789)0.7300:0.00490170,Nicandra_physalodes_TD1:0.01045468)0.5100:0.00003178,Jaltomata_procumbens_TD1*:0.01609256)0.8200:0.00830564,Lycium_barbarum_TD1*:0.01946688)0.7500:0.01592742,Nicotiana_tabacum_TDb:0.03004830)0.8900:0.01854231,Nicotiana_tabacum_TDa*:0.09243376)0.5300:0.01209106,(Nicotiana_attenuata_TD:0.06630921,(Lycium_barbarum_TD2*:0.14956552,(Jaltomata_procumbens_TD2*:0.03935748,(Solanum_pseudocapsicum_TD2*:0.12727501,((Solanum_dulcamara_TD2*:0.07781250,Solanum_melongena_TD2:0.08001899)0.9400:0.02015630,(Solanum_bulbocastanum_TD2:0.07312864,(Solanum_lycopersicum_TD2:0.05463008,Solanum_tuberosum_TD2:0.02963604)0.9700:0.02235985)0.4700:0.01894757)0.7400:0.01308106)1.0000:0.04302168)0.8600:0.03446720)0.8800:0.01619327)1.0000:0.12103741)1.0000:0.05290391,(Lactuca_serriola_340634423:0.12797988,Lactuca_serriola_340635064:0.11719534)1.0000:0.07329355)0.6100:0.00940965,(Glycine_max_TDa:0.01455157,Glycine_max_TDb:0.03191650)1.0000:0.11358248,Arabidopsis_thaliana:0.14723918);

8. Method=ML, Model=Tamura-3-parameter, Gap/missing data treatment=Use all sites, Log Likelihood=-22847.84, Tree=((((((((((((Solanum_arcanum_TD1:0.00584259,Solanum_lycopersicum_TD1:0.00541161)1.0000:0.00689466,Solanum_sitiens_TD1:0.00790536)1.0000:0.00172879,(Solanum_bulbocastanum_TD1:0.00777520,Solanum_tuberosum_TD1:0.00762597)0.7900:0.00232726)1.0000:0.01873938,Solanum_melongena_TD1:0.04080688)0.5700:0.00000010,Solanum_dulcamara_TD1:0.01748579)0.5700:0.00000011,Solanum_aviculare_TD1:0.01732885)0.6100:0.00450995,(Jaltomata_procumbens_TD1*:0.02495610,Solanum_pseudocapsicum_TD1:0.01938576)0.9200:0.00870342)0.9900:0.00513370,Nicandra_physalodes_TD1:0.02786185)0.9700:

0.00000010,Lycium_barbarum_TD1*:0.10410082)1.0000:0.02440565,Nicotiana_tabacum_TDb:0.04438390)1.0000:0.06576664,Nicotiana_tabacum_TDa*:0.14743569)0.8000:0.02519674,(Nicotiana_attenuata_TD:0.10152940,(Lycium_barbarum_TD2*:0.16725950,(Jaltomata_procumbens_TD2*:0.01478713,(((Solanum_lycopersicum_TD2:0.08444570,Solanum_tuberosum_TD2:0.03906131)0.9700:0.00297823,Solanum_bulbocastanum_TD2:0.04634664)1.0000:0.08319861,(Solanum_dulcamara_TD2*:0.05780930,(Solanum_melongena_TD2:0.09299505,Solanum_pseudocapsicum_TD2*:0.13507752)0.5000:0.02508546)1.0000:0.01549477)1.0000:0.07222938)0.9400:0.00000010)1.0000:0.06958039)1.0000:0.12274032)1.0000:0.05651913,(Lactuca_serriola_340634423:0.16344617,Lactuca_serriola_340635064:0.15106205)1.0000:0.06463221)1.0000:0.03535441,(Glycine_max_TDa:0.02231347,Glycine_max_TDb:0.02677019)1.0000:0.14912018,Arabidopsis_thaliana:0.19799879);

9. Method=ML, Model=Tamura-3-parameter, Gap/missing data treatment=partial deletion(site coverage cutoff-95%), Log Likelihood=-7155.63, Tree=((((((((((((Solanum_arcanum_TD1:0.00895605,Solanum_lycopersicum_TD1:0.00594097)1.0000:0.00854634,Solanum_sitiens_TD1:0.00631476)0.6100:0.00000010,Solanum_tuberosum_TD1:0.00633837)0.4900:0.00000000,Solanum_bulbocastanum_TD1:0.00705181)0.9800:0.00964967,(Solanum_melongena_TD1:0.01850348,Solanum_pseudocapsicum_TD1:0.01241797)0.5800:0.00597070)0.5300:0.00000010,(Solanum_aviculare_TD1:0.00954582,Solanum_dulcamara_TD1:0.01095575)0.4200:0.00868780)0.8300:0.00625067,Jaltomata_procumbens_TD1*:0.01663024)0.4800:0.00000010,Nicandra_physalodes_TD1:0.01377171)0.9100:0.01039250,Lycium_barbarum_TD1*:0.07146409)0.9400:0.01491989,Nicotiana_tabacum_TDb:0.02800643)0.9800:0.03522966,(Nicotiana_tabacum_TDa*:0.09510200,(Nicotiana_attenuata_TD:0.07003950,(Lycium_barbarum_TD2*:0.13969705,(Jaltomata_procumbens_TD2*:0.04360790,(Solanum_pseudocapsicum_TD2*:0.12696634,((Solanum_dulcamara_TD2*:0.13846586,Solanum_melongena_TD2:0.07515700)0.8900:0.02086361,(Solanum_bulbocastanum_TD2:0.06509545,(Solanum_lycopersicum_TD2:0.05778160,Solanum_tuberosum_TD2:0.02447885)1.0000:0.01997089)0.8500:0.02450079)0.5600:0.01716485)1.0000:0.03835665)0.9900:0.04372963)0.9400:0.01718203)1.0000:0.11824728)0.4000:0.00000010)1.0000:0.05967095,(Lactuca_serriola_340634423:0.10687464,Lactuca_serriola_340635064:0.11480742)1.0000:0.06609464)0.7600:0.00997539,(Glycine_max_TDa:0.01923477,Glycine_max_TDb:0.02656105)1.0000:0.10254050,Arabidopsis_thaliana:0.14080957);

10. Method=ML, Model=Hasegawa-Kishino-Yano, Gap/missing data treatment=complete deletion, Log Likelihood=-5408.36, Tree=((((((((((((Solanum_arcanum_TD1:0.01172812,Solanum_lycopersicum_TD1:0.00560158)0.9000:0.00791538,Solanum_sitiens_TD1:0.00384701)0.5500:0.00025565,Solanum_tuberosum_TD1:0.00816126)0.4100:0.00000000,Solanum_bulbocastanum_TD1:0.00762381)0.8400:0.00666043,(Solanum_melongena_TD1:0.02005207,Solanum_pseudocapsicum_TD1:0.00992863)0.4600:0.00401817)0.4100:0.00000010,(Solanum_aviculare_TD1:0.01026169,Sola

num_dulcamara_TD1:0.01224404)0.5700:0.00724795)0.7600:0.00491953,Nicandra_physalo
des_TD1:0.01044944)0.4200:0.00002534,Jaltomata_procumbens_TD1*:0.01609178)0.7200:0
.00830444,Lycium_barbarum_TD1*:0.01935078)0.6700:0.01594916,Nicotiana_tabacum_T
Db:0.03020946)0.9100:0.01852276,Nicotiana_tabacum_TDa*:0.09285166)0.5600:0.01196135
,(Nicotiana_attenuata_TD:0.06607245,(Lycium_barbarum_TD2*:0.15019269,(Jaltomata_p
rocumbens_TD2*:0.03848104,(Solanum_pseudocapsicum_TD2*:0.12742619,((Solanum_d
ulcamara_TD2*:0.07842222,Solanum_melongena_TD2:0.07902087)0.9100:0.02019372,(Sol
anum_bulbocastanum_TD2:0.07312654,(Solanum_lycopersicum_TD2:0.05468868,Solanu
m_tuberosum_TD2:0.02960774)0.9800:0.02236564)0.5700:0.01896306)0.7500:0.01307099)1.
0000:0.04319872)0.8400:0.03452877)0.8500:0.01626110)1.0000:0.12146274)1.0000:0.0531445
1,(Lactuca_serriola_340634423:0.12761647,Lactuca_serriola_340635064:0.11732859)1.0000:
0.07349051)0.5900:0.00943452,(Glycine_max_TDa:0.01473093,Glycine_max_TDb:0.031944
98)1.0000:0.11374059,Arabidopsis_thaliana:0.14741815);

11. Method=ML, Model=Tamura-Nei, Gap/missing data treatment=complete deletion, Log Likelihood=-5408.31,
Tree=((((((((((((Solanum_arcanum_TD1:0.01172888,Solanum_lycopersicum_TD1:0.0055
9805)0.9600:0.00791606,Solanum_sitiens_TD1:0.00384405)0.5900:0.00025648,Solanum_tub
erosum_TD1:0.00816079)0.4100:0.00000000,Solanum_bulbocastanum_TD1:0.00762097)0.8
700:0.00665636,(Solanum_melongena_TD1:0.02004883,Solanum_pseudocapsicum_TD1:0
.00993155)0.4600:0.00402210)0.3700:0.00000010,(Solanum_aviculare_TD1:0.01026152,Sola
num_dulcamara_TD1:0.01224349)0.5900:0.00725294)0.8600:0.00491980,Nicandra_physalo
des_TD1:0.01045094)0.5400:0.00002556,Jaltomata_procumbens_TD1*:0.01609188)0.7700:0
.00829608,Lycium_barbarum_TD1*:0.01934514)0.7000:0.01595381,Nicotiana_tabacum_T
Db:0.03021632)0.9000:0.01852814,Nicotiana_tabacum_TDa*:0.09286551)0.4900:0.01198308
,(Nicotiana_attenuata_TD:0.06606600,(Lycium_barbarum_TD2*:0.15019459,(Jaltomata_p
rocumbens_TD2*:0.03847204,(Solanum_pseudocapsicum_TD2*:0.12738781,((Solanum_d
ulcamara_TD2*:0.07845149,Solanum_melongena_TD2:0.07902764)0.9300:0.02019886,(Sol
anum_bulbocastanum_TD2:0.07313390,(Solanum_lycopersicum_TD2:0.05467963,Solanu
m_tuberosum_TD2:0.02961152)1.0000:0.02236117)0.5100:0.01895449)0.7300:0.01306148)1.
0000:0.04315845)0.8100:0.03451348)0.8200:0.01624048)1.0000:0.12142647)1.0000:0.0531130
4,(Lactuca_serriola_340634423:0.12764752,Lactuca_serriola_340635064:0.11739533)1.0000:
0.07349406)0.6000:0.00942881,(Glycine_max_TDa:0.01471177,Glycine_max_TDb:0.031928
48)1.0000:0.11371792,Arabidopsis_thaliana:0.14743482);

12. Method=ML, Model=General Time Reversible, Gap/missing data treatment=complete deletion, Log Likelihood=-5405.32,
Tree=((((((((((((Solanum_arcanum_TD1:0.01196236,Solanum_lycopersicum_TD1:0.0056
5246)0.9700:0.00792209,Solanum_sitiens_TD1:0.00383944)0.5400:0.00025243,Solanum_tub
erosum_TD1:0.00817817)0.4200:0.00000000,Solanum_bulbocastanum_TD1:0.00770162)0.8

800:0.00677678,(Solanum_melongena_TD1:0.02035607,Solanum_pseudocapsicum_TD1:0
.01001614)0.3800:0.00395284)0.4200:0.00000010,(Solanum_aviculare_TD1:0.01030600,Sola
num_dulcamara_TD1:0.01226792)0.4300:0.00706902)0.7900:0.00497597,Nicandra_physalo
des_TD1:0.01027864)0.5400:0.00002614,Jaltomata_procumbens_TD1*:0.01601278)0.8100:0
.00839673,Lycium_barbarum_TD1*:0.01935014)0.7000:0.01606927,Nicotiana_tabacum_T
Db:0.03012637)0.9100:0.01850735,Nicotiana_tabacum_TDa*:0.09332335)0.5400:0.01251414
,(Nicotiana_attenuata_TD:0.06677085,(Lycium_barbarum_TD2*:0.15147465,(Jaltomata_p
rocumbens_TD2*:0.03714132,(Solanum_pseudocapsicum_TD2*:0.12584405,((Solanum_d
ulcamara_TD2*:0.07860558,Solanum_melongena_TD2:0.07743539)0.9100:0.02046189,(Sol
anum_bulbocastanum_TD2:0.07302875,(Solanum_lycopersicum_TD2:0.05471525,Solanu
m_tuberosum_TD2:0.03005441)0.9900:0.02146704)0.6400:0.01930121)0.6700:0.01424239)1.
0000:0.04192236)0.9400:0.03491663)0.8200:0.01512142)1.0000:0.12205246)1.0000:0.0523924
8,(Lactuca_serriola_340634423:0.12500486,Lactuca_serriola_340635064:0.11882444)1.0000:
0.07309769)0.6000:0.00976752,(Glycine_max_TDa:0.01513582,Glycine_max_TDb:0.032688
69)1.0000:0.11661540,Arabidopsis_thaliana:0.14948680);

13. Method=Neighbor-joining, Model=JC, Gap/missing data treatment=complete deletion, Log Likelihood=N.A.,
Tree=((((((((((((Solanum_arcanum_TD1:0.01091062,Solanum_lycopersicum_TD1:0.00549
468)0.9600:0.00777926,Solanum_sitiens_TD1:0.00563106)0.5600:0.00026077,Solanum_tube
rosum_TD1:0.00918204)0.4900:-
0.00011333,Solanum_bulbocastanum_TD1:0.00777485)0.8800:0.00663755,(Solanum_melo
ngena_TD1:0.02109104,Solanum_pseudocapsicum_TD1:0.00996930)0.6100:0.00412535)0.4
900:-
0.00289083,(Solanum_aviculare_TD1:0.01095404,Solanum_dulcamara_TD1:0.01169696)0.
7300:0.00747549)0.8300:0.00518681,(Jaltomata_procumbens_TD1*:0.01211532,Nicandra_p
hysalodes_TD1:0.00636611)0.4900:0.00434561)0.9600:0.00813700,Lycium_barbarum_TD1*
:0.02792065)0.5900:0.01581219,Nicotiana_tabacum_TDb:0.03610781)0.9700:0.01821581,Ni
cotiana_tabacum_TDa*:0.08509662)0.8100:0.01493781,(Nicotiana_attenuata_TD:0.070503
12,(Lycium_barbarum_TD2*:0.15118873,(Jaltomata_procumbens_TD2*:0.05236574,(Solan
um_pseudocapsicum_TD2*:0.12473224,((Solanum_dulcamara_TD2*:0.07385396,Solanum
_melongena_TD2:0.07608583)0.8300:0.01932462,(Solanum_bulbocastanum_TD2:0.070162
19,(Solanum_lycopersicum_TD2:0.05216962,Solanum_tuberosum_TD2:0.02916445)0.960
0:0.02200214)0.7500:0.01917520)0.7900:0.01254518)1.0000:0.04081549)0.8300:0.03509267)0.
8700:0.01447048)1.0000:0.11619330)0.9900:0.04918773,(Glycine_max_TDa:0.02016604,Glyc
ine_max_TDb:0.03233965)1.0000:0.12247952)0.4600:0.00511617,(Lactuca_serriola_3406344
23:0.10743193,Lactuca_serriola_340635064:0.11802694)1.0000:0.06626195,Arabidopsis_tha
liana:0.14917149);

14. Method=Minimum Evolution, Model=JC, Gap/missing data treatment=complete deletion, Log Likelihood=N.A.,
Tree=((((((((((((Solanum_arcanum_TD1:0.01091062,Solanum_lycopersicum_TD1:0.00549468)0.9900:0.00777926,Solanum_sitiens_TD1:0.00563106)0.6100:0.00026077,Solanum_tuberosum_TD1:0.00918204)0.5700:-0.00011333,Solanum_bulbocastanum_TD1:0.00777485)0.9100:0.00663755,(Solanum_melongena_TD1:0.02109104,Solanum_pseudocapsicum_TD1:0.00996930)0.6900:0.00412535)0.5400:-0.00289083,(Solanum_aviculare_TD1:0.01095404,Solanum_dulcamara_TD1:0.01169696)0.6900:0.00747549)0.8600:0.00511580,Nicandra_physalodes_TD1:0.01035014)0.5000:-0.00001388,Jaltomata_procumbens_TD1*:0.01609507)0.9600:0.00827384,Lycium_barbarum_TD1*:0.02792065)0.6700:0.01581219,Nicotiana_tabacum_TDb:0.03610781)0.9500:0.01821581,Nicotiana_tabacum_TDa*:0.08509662)0.7600:0.01493781,(Nicotiana_attenuata_TD:0.07050312,(Lycium_barbarum_TD2*:0.15118873,(Jaltomata_procumbens_TD2*:0.05236574,(Solanum_pseudocapsicum_TD2*:0.12473224,((Solanum_dulcamara_TD2*:0.07385396,Solanum_melongena_TD2:0.07608583)0.9200:0.01932462,(Solanum_bulbocastanum_TD2:0.07016219,(Solanum_lycopersicum_TD2:0.05216962,Solanum_tuberosum_TD2:0.02916445)0.9500:0.02200214)0.8200:0.01917520)0.8200:0.01254518)0.9900:0.04081549)0.8500:0.03509267)0.8500:0.01447048)1.0000:0.11619330)1.0000:0.04782399,(Lactuca_serriola_340634423:0.10743193,Lactuca_serriola_340635064:0.11802694)1.0000:0.07197458)0.5000:0.01068024,(Glycine_max_TDa:0.02016604,Glycine_max_TDb:0.03233965)1.0000:0.11371211,Arabidopsis_thaliana:0.14540756);

15. Method=UPGMA, Model=JC, Gap/missing data treatment=complete deletion, Log Likelihood=N.A.,
Tree=((((((((((Jaltomata_procumbens_TD1*,Nicandra_physalodes_TD1)0.7400,(Solanum_aviculare_TD1,Solanum_dulcamara_TD1)0.7100)0.2000,Solanum_pseudocapsicum_TD1)0.1700,((Solanum_arcanum_TD1,Solanum_lycopersicum_TD1)0.7700,(Solanum_tuberosum_TD1,(Solanum_bulbocastanum_TD1,Solanum_sitiens_TD1)0.4600)0.7700)0.5900)0.3700,Solanum_melongena_TD1)0.7600,Lycium_barbarum_TD1*)0.6600,Nicotiana_tabacum_TDb)1.0000,Nicotiana_tabacum_TDa*)1.0000,(Glycine_max_TDa,Glycine_max_TDb)1.0000)0.5200,((Lactuca_serriola_340634423,Lactuca_serriola_340635064)1.0000,(((Jaltomata_procumbens_TD2*,Nicotiana_attenuata_TD)0.9500,Lycium_barbarum_TD2*)0.6600,(Solanum_pseudocapsicum_TD2*,((Solanum_dulcamara_TD2*,Solanum_melongena_TD2)0.8100,(Solanum_bulbocastanum_TD2,(Solanum_lycopersicum_TD2,Solanum_tuberosum_TD2)1.0000)0.9000)0.9200)0.7800)0.9900)0.8600,Arabidopsis_thaliana);

Note: All the tests have 100 bootstrap repetitions.

Appendix K: TD sequences obtained from this study

>Solanum_sitiens_TD1

ATGGAGGTTCTTCGGTTTACCGCCGTGAAATCATTGAATTCCTGCGTCCGT
CCAGAATTCACGGCGACGAGTTCCGTCATTGTTCCTATTAACACTGTGAGAGTTT
GCGGAACGAGAGAAAGCAAAAAGAAGGCGTTTATTCGTGCTAAGGCGACGGAG
ATCTTGTCATCGCCAATTACGGTAACGGAACCGTTGAAGGCGGAGCCGGCAGAG
GCAACGGTTCCGTTGCTACGAGTTTCTCCGAGCTCGTTGCAGTGTGAGCCGGGAT
ATTTGATACCGAATAGTCCGGTTTTAGGTGCGGGTGGTGTGACCGGGTATGAGTA
TTTGACGAACATCTTGTCATCGAAAGTTTATGATGTAGCTTATGAGACGCCTTTG
CAGAAAGCTCCTAAGCTGTCAGAAAGATTGGGGGTTAATGTATGGCTAAAAAG
AGAGGATCTTCAGCCGGTGTTTTCGTTCAAAATCAGAGGAGCTTACAACATGAT
GGCAAAACTCCCTAAAGAGCAGTTGGAAAAAGGGGTTATATGCTCATCAGCTG
GAAACCATGCACAAGGTGTTGCATTATCTGCTCAGAGACTTGGTTGCGATGCTGT
GATTGTTATGCCTGTTACTACACCAGACATCAAATGGAAATCAGTTAAGAGATT
GGGTGCTACTGTTGTTCTTGTTGGGGACTCATATGATGAAGCTCAAGCATATGCC
AAAAAGCGGGCTGAAGCTGAAGGCCGCACATTCATCCCTCCTTTTGATCACCCA
GATGTCATCGTAGGGCAAGGTACAGTAGGGATGGAGATAAATCGCCAACTCAA
AGATAACATTCATGCAATCTTTGTGCCTGTTGGAGGAGGGGGGCTTATAGCTGGT
ATTGCTGCTTATTTGAAAAGGGTTGCCCCAGATATAAAGATTATTGGAGTTGAAC
CACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGCCAGAGAGTAATGC
TGGACCAAGTTGGGGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAG
AGACTTATCGTCTCTGCGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGA
TGCTATATGTGCATCTATAAAGGACATGTTTGAAGAGAAAAGGAGCATACTAGA
GCCTGCAGGTGCACTTGCTCTTGCTGGAGCCGAGGCATACTGCAAGTATTATGGC
CTCAAGGGTGAAAACGTAGTAGCAATAACTAGTGGAGCCAACATGAACTTTGA
CAGGCTTAGATTGGTAACAGAACTCGCAGATGTTGGTAGACAGCGGGAAGCTGT
TCTTGCTACTTTTATGCCAGAAGACCCAGGAAGCTTCAAAAAGTTCGCTGAAAT
GGTAGGACCAATGAATATCACTGAATTCAAGTACAGATACAACTCTGATAAAGA
AAGAGCTCTTGTACTTTACAGTGTTGGACTTCACACAATATTAGAACTTGAAGGA
ATGGTGGAGAGGATGGAATCAGCAGATCTGCAAACCATTAATCTTACAGACAAT
GACTTGGTCAAAGATCATCTTAGGCATTTGATGGGTGGTAGAACAAATGTTCAT
AATGAGCTTTTGTGTCGATTCACTTTTCCCGAGAAGCCTGGTGCTTTGATGAAGTT
TTTAGATGCTTTCAGCCCACGTTGGAATATAAGTTTGTTTCATTATCGTGCACAGG
GAGATACTGGTGCAAATGTGCTAGTCGGAATCCAAGTTCCACAGGATGAGGTAG
TTGAGTTTGAGGGACGGGCAGACAGTCTTGGTTATGAATATGCAGTGGAGAGTC
TCAATGAGGCATACCAGCTCATAATGCAT

>Solanum_melongena_TD1

ATGGAGGTTCTTCGGTTTACAGCCGTAAAGTCATTGAATTCCTGCGTCCGT
CCCGAATTATCGGGGACGAGCTCCGTCATTGTTCCCCTTAACACTGTGAAAGCTT
CCGGAACGAGGAAAACCAAAAAGAAGGCGTTCATTCGTGCTAAGGCAACGGAG
ATCTTGTCATCCCCAGCTAGGGTTACAGAGTCATTGCAGGCGGCGCCGGTAGAA
GCACCGTCGGGTAAGGTTCCGTCGGTGCGAGTTTCTCCCAGCTCGTTGCAGTGTG
AGCCGGGATATTTGATACCGAATAGTCCGGTTCTAGGTAGTGGTGGTTTGATCGG
TTACGAGTATCTGACGAACATCTTGTCGTCGAAAGTGTATGATGTAGCTTATGAG
ACGCCGTTGCAGAAAGCTCCTAAGCTGTCAGAAAGATTGGGAGTTAATGTGTGG
CTCAAAAGAGAGGATCTTCAGCCGGTCTTCTCATTCAAAATCAGAGGAGCTTAC
AATATGATGGCAAAACTCTCTAAAGAGCAGTTGGAAAAAGGGGTAATATGCTC
ATCAGCTGGAAACCATGCACAAGGTGTTGCATTATCCGCTCAGAGACTTGGCTG
CGATGCTGTGATTGCCATGCCTGTTACTACGCCAGACATCAAATGGAAATCAGTT
AAGAGATTGGGTGCTACTGTTGTTCTTGTGGGCGACTCGTATGATGAAGCTCAAG
CATATGCCAAAGAGCGGGCTGTAGCTGAAGGCCGCACATTCATCCCTCCTTTTG
ATCACCCAGATGTCATCGCCGGGCAAGGTACAGTAGGGATGGAGATAAATCGC
CAACTCAAAGATAACATTCATGCAATCTTTGTGCCTGTTGGAGGAGGGGGGCTT
ATAGCTGGTATTGCTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTG
GAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGA
GAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGT
TGGTGAAGAAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTT
GGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAGAAAAGAAGC
ATACTAGAGCCTGCAGGTGCACTTGCTCTTGCGGGAGCCGAGGCATATTGCAAG
TACTATGGACTCGAGGGTGAAAATGTGGTAGCAATAACTAGTGGAGCCAACATG
AACTTTGACAGACTTAGATTGGTCACTGAACTCGCAGATGTTGGTAGACAGCGG
GAAGCTGTTCTTGCTACTTTTCTGCCAGAAGAGGCAGGAAGCTTCAAAAAGTTTG
CTGAAATGGTTGGACCAATTAATATCACTGAATTCAAGTACAGATACAATTCTG
ATAAAGAAAGAGCTCTAGTACTTTACAGTGTTGGTCTTCACACAATATTAGAACT
TGAAGGAATGGTGGAGAGGATTGAAGCAGCAGATCTGCAAACCATTAATCTTAC
AGACAATGACTTGGTGAAAGATCATCTTAGGCATTTGATGGGTGGGAGAACAAA
TGTTCAGAATGAGCTTTTGTGTCGATTCACTTTTCCCGAGAAGCCTGGTGCCTTGA
TGAAGTTTTTAGATGCTTTCAGTCCACGTTGGAATATAAGTTTGTTTCATTATCGT
GCACAGGGAGATACTGGTGCAAATGTGCTAGTCGGAATCCAAGTTCCACACGAT
GAGGAAGTTGAGTTCCAGGGACGGGCAGATAGTCTTGGTTATGAATATACAGTG
GAGAGTCTCAACGAGGCATTGCAGCTCATAATGCAT

>Nicandra_physalodes_TD1

ATGGAGGTTCTTCGATTTACCGCCGTGAAGTCATTGAATTCTTACGTCCGT
CCCGAATTCACAGCGACGACCTACGCCATTGTTCCCGTTAAGACTGTGAAAACC
TCTGGAACGAGGAAAAACAAGAAGAAGGCGTTCATTCGTGCTAAGGCGACGGA
GATCTTGACATCTCCAGCTACGGTAGCGGAGTCGTTGCAGGCGGCACCGGTAGA

GGCACCATCGGCTCCGTTGCTACGAGTTTCTCCGAGCTCGTTGCAGTGCGACCCG
GGGTACTTGATACCGAATAGTCCGGTTCTAGGTACTGGTGGTGTGGTTGGGTACG
AGTATTTGACAAACATCTTGTCGTCAAAAGTGTATGATGTAGCTTATGAGACGCC
TTTGCAGAAAGCGCCTAAACTGTCAGAAAGGCTGGGAGTTAATGTGTTTCTAAA
AAGAGAGGATCTTCAGCCCGTCTTCTCGTTCAAAATCAGAGGAGCTTACAATAT
GATGGCAAAACTACCAAAAGTGCAGTTGGAAAAAGGGGTTATATGCTCATCAG
CTGGAAACCATGCACAAGGTGTTGCATTATCTGCTCAGAGACTTGGCTGCGATG
CTGTGATTGCCATGCCTGTTACTACACCAGACATCAAATGGAAATCAGTTAAGA
GATTGGGTGCTACTGTTGTTCTAGTTGGGGACTCATATGATGAAGCTCAAGCATA
CGCCAAAGAGCGGGCTGAAGCTGAAGGCCGCACATTCATCCCTCCTTTTGATCA
CCCAGATGTCATCGTAGGGCAAGGTACAGTAGGAATGGAGATAAATCGCCAAC
TCAAAGATAACATTCATGCAATCTTTGTGCCTGTTGGAGGTGGGGGTCTTATAGC
TGGTATTGCTGCTTATTTGAAAAGGGTTGCCCCTGATATAAAGATTATTGGAGTT
GAACCACTTGATGCAAATGCATTGGCATTATCATTACAGCATGGTCAGAGAGTA
ATGCTGGACCAAGTTGGGGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTG
AAGAAACTTATCGTCTCTGCAAGGAATTGATAGATGGTGTAGTCCTCGTTGGTCG
TGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAGAAGAGAAGCATACT
TGAGCCTGCAGGTGCACTTGCTCTTGCTGGAGCTGAGGCATACTGCAAGTACTAT
GGCCTCAAGGATGAAAACGTCGTAGCAATTACTAGTGGAGCCAACATGAACTTT
GACAGACTTAGATTGGTAACTGAACTCGCAGATGTTGGTAGACAGCGGGAAGCT
GTTCTTGCTACTTTTATGCCAGAAAACCCAGGGAGCTTCAAAAAATTTGCTGAAA
TGGTAGGACCAATGAATATCACTGAATTCAAGTACAGATACAATTCTGATAAAG
AAAGAGCTCTAGTACTTTACAGTGTTGGTCTTCACACAAAATTAGAACTTGAAG
AAATGGTGGAGAGGATGAAATCAGCAGATCTGCAGACCATTAATCTTACAGAC
AACGACTTGGTCAAAGATCATCTAAGGCATTTGATGGGTGGTAGAACAAATGTT
CATAATGAGCTTTTGTGTCGATTCACTTTTCCCGAGAAACCCGGTGCTTTAATGA
AGTTTTTAGATGCTTTCAGCCCACGTTGGAATATAAGTTTGTTTCATTATCGTGCA
CAGGGAGATACTGGTGCAAATGTGCTAGTCGGGATCCAAGTTCCAAAGGATGA
GGTAGAAGAGTTCCAGGGACGAGCAGACAGTCTTGGTTATGAATATGCAGTGGA
GAGTCTCAATGAGGCATTCCAGCTCATAATGCAT

>Solanum_bulbocastanum_TD1

ATGGAGGTTYTTCGGTTTACCGGAGTGAAATCATTGAATTCCTGCGTCCGT
CCAGAATTCACGGCGATGAGTTCCGTCATTGTTCCTATTAACACTGTGAAAGTTT
CCGGAACGAGGAAAAGCAAAAGGAAGGCGTTTATTCGTGCTAAGGCGACGGAG
ATCTTGTCATCGCCAGCAACGGTAACGGAACCGTTGAAGTCGGAGCCGGTTGAG
GCACCGGAGGCTCCGGTTCCGTTGCTACGAGTTTCTCCGAGCTCGTTGCAGTGTG
AGCCGGGATATTTGATACCGAATACTCCGGTTTTAGGTACTGGTGGTGTGACCGG
GTATGAGTATTTGACAAACATCTTGTCATCGAAAGTTTATGATGTAGCTTATGAG
ACGCCTTTGCAGAAAGCGCCTAAGCTGTCAGAAAGATTGGGAATTAATGTATGG

CTTAAAAGAGAGGATCTTCAGCCGGTGTTTTCGTTCAAAATCAGAGGAGCTTAC
AATATGATGGCAAAACTCCCTAAAGAGCAGTTGGAAAAAGGGGTTATATGCTC
ATCAGCTGGAAACCATGCACAAGGTGTTGCATTATCTGCTCAGAGACTTGGCTG
CGATGCTGTGATTGTTATGCCTGTTACCACACCAGACATCAAATGGAAATCAGTT
AAGAGATTGGGTGCTACTGTTGTTCTTGTTGGGGACTCATATGATGAAGCTCAAG
CATATGCCAAAGAGCGGGCTGAAGCTGAAGGCCGCACATTCATCCCTCCTTTTG
ATCACCCAGATGTCATCGTAGGGCAAGGTACAGTAGGGATGGAGATAAATCGC
CAATTCAAAGATAACATTCATGCAATCTTTGTGCCTGTTGGAGGAGGGGGGCTT
ATAGCTGGTATTGCTGCTTATTTGAAAAGGGTTGCCCCAGATATAAAGGTTATTG
GAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGA
GAGTAATGCTGGACCAAGTTGGGGGTTTTGCAGATGGTGTAGCTGTTAAAGTGG
TTGGTGAAGAGACTTATCGTCTCTGCGAGGAATTAATAGATGGCGTAGTCCTAGT
TGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAGAAAAGGAG
CATACTAGAGCCTGCAGGTGCACTTGCTCTTGCTGGAGCCGAGGCATACTGCAA
GTATTATGGCCTCAAGGGTGAAAACGTAGTAGCAATAACTAGTGGAGCCAACAT
GAACTTTGACAGGCTTAGATTGGTAACAGAACTCGCAGATGTTGGTAGACAGCG
GGAAGCTGTTCTTGCTACTTTTATGCCAGAAGAACCAGGAAGCTTCAAAAAGTT
CGCTGAAATGGTAGGACCAATGAATATCACTGAATTCAAGTACAGATACAACTC
TGATAAAGAAAGAGCTCTTGTACTTTACAGTGTTGGACTTCACACAATATTAGA
ACTTGAAGGAATGGTGGAGAGGATGGAATCAGCAGATCTGCAAACCATTAATCT
TACAGACAATGACTTGGTCAAAGATCATCTTAGGCATTTGATGGGTGGTAGAAC
AAATGTTCATAATGAGCTTTTGTGTCGATTCACTTTTCCCGAGAAGCCTGGTGCTT
TGATGAAGTTTTTAGATGCTTTCAGCCCATGTTGGAATATAAGTTTGTTTCATTAT
CGTGCACAGGGAGATACTGGTGCAAATGTGCTAGTCGGGATCCAAGTTCCACCG
GATGAGGTAGTTGAGTTTGAGGGACGGGCAGACAATCTTGGTTATGAATATGCA
GTGGAGAGTCTCAATGAGGCATACCAGCTCATAATGCAT

>Solanum_pseudocapsicum_TD1

ATGGAGGTTCTTCGGTTTACAGCCGTGAAGTCATTGAATTCCTGCGGTCGT
CCCGAATTCACGGCGACGAGCTCCGTCATTGTTCCTATTAACACTGTGAAAGCTT
CCGGAACGAGGAAAAGCAAAAAGAAGGCGTTCATTCGTGCTAAGGCGACGGA
GATCCTGTCATCTCCAGCTAAGGTGACAGAGCCGTTGCAGGCGGCGCCGGTGGA
GGCACCGTCGGCTCCGGTTCCGTTGCTGCGAGTTTCTCCCAGCTCATTGCAGTGT
GAGCCGGGATGTTTGATACCGAATAGTCCGGTTCTAGGTAGTGGTGGTTTGATGG
GTTACGAGTATTTGACGAACATCTTGTCGTCAAAAGTGTATGATGTAGCTTATGA
GACGCCATTGCAGAAAGCGCCTAAGCTGTCAGAAAGATTGGGAGTTAATGTGTG
GCTCAAAAGAGAGGATCTCCAGCCGGTCTTCTCGTTCAAAATCAGAGGAGCTTA
CAATATGATGGCAAAACTCCCTAAAGAGCAGTTGGAAAAAGGGGTTATATGCTC
ATCAGCTGGAAACCATGCACAAGGTGTTGCATTATCTGCTCAGAGACTTGGCTG
CAATGCTGTGATTGCCATGCCTGTTACTACTCCAGACATCAAATGGAAATCAGTT

AAGAGATTGGGTGCTACTGTTGTTCTTGTTGGGGACTCATATGATGAGGCTCAAG
CATATGCCAAAGAGCGGGCTGAAGCTGAAGGCCGCACATTCATCCCTCCTTTTG
ATCACCCAGATGTCATCGCTGGTCAAGGTACAGTAGGAATGGAGATAAATCGCC
AACTCAAAGATAACATTCATGCAATCTTTGTGCCTGTTGGAGGAGGGGGGCTTA
TAGCTGGTATTGCTGCTTATTTGAAAAGGGTTGCCCCCGATATAAAGATTATTGG
AGTTGAACCACTTGATGCAAATGCGTTGGCGTTATCATTACAGCATGGTCAGAG
AGTAATGCTGGACCAAGTTGGGGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTT
GGTGAAGAAACTTATCGTATCTGCGAGGAACTAATAGATGGTGTAGTCCTAGTT
GGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAAAAAGAAGC
ATACTAGAGCCTGCAGGTGCACTTGCTCTCGCTGGAGCCGAGGCATACTGCAAG
TACTATGGCCTCAAGGGTGAAAACGTAGTAGCAATAACTAGTGGAGCCAACAT
GAACTTTGACAGACTTAGACTGGTTACGGAACTCGCAGATGTTGGTAGACAGAG
GGAAGCTGTTCTTGCTACTTTTATGCCAGAAGAGCCAGGGAGCTTCAAAAAGTT
CGCTGAAATGGTAGGACCAATGAATATCACTGAATTCAAGTACAGATACAATTC
TAATAAAGAAAGAGCTCTAGTACTTTACAGTGTTGGTCTTCACATGAAACTAGA
ACTTGAAGAAATGGTGGACAGGATGAAATCAGCAGATCTGAAAACCATTAATC
TTACAGACAATGACTTGGTCAAAGATCATCTAAGGCATTTGATGGGTGGTAGAA
CAAATGTTCATAATGAGCTTTTGTGTCGATTCACTTTTCCCGAGAAGCCCGGTGC
TTTGATGAAGTTTTTAGATGCTTTCAGCCCATGTTGGAATATAAGTTTGTTTCATT
ACCGTGCACAGGGAGACACTGGTGCAAATGTGCTAGTTGGGATCCAAGTTCCAC
AGGATGAGGTAGTTGAGTTCCAGAGACGGGCAGACAGTCTTGGTTATGAATATG
CAGTGGAGAGTCTCAATGAGGCATACCAGCTCATAATGCAT

>Solanum_arcanum_TD1

ATGGAGGTTCTGCGGTTTACCGCCGTGAAATCATTGAATTCATGCGTCCG
TCCAGAATTCACGGCGATGAGTTCCGTCATTGTTCCTATTAGCACTGTGAAAGTT
TCTGGAACGAGGAAAAGCAAAAAGAAGGCGTTTATTCGTGCTAAGGCGACGGA
GATCTTGTCATCGCCAGCTACGGTAACGGAACCGTTGAAGGCGGAGCCGGCAG
AGGCACCGGTTCCGTTGCTACGAGTTTCTCCGAGCTCGTTGCAGTGTGAGCCGGG
ATATTTGTTACCGAATAGTCCGGTTTTAGGTACGGGTGGTGTGACCGGGTATGAG
TATTTGACGAACATCTTGTCATCGAAAGTTTATGATGTAGCTTATGAGACGCCTT
TGCAGAAAGCTCCTAAGCTGTCAGAAAGATTGGGAGTTAATGTATGGCTAAAAA
GAGAGGATCTTCAGCCGGTGTTTTCGTTCAAAATCAGAGGAGCTTACAATATGA
TGGCAAAACTCCCGAAAGAGCAGTTGGAAAAAGGGGTTATATGCTCATCAGCT
GGAAACCATGCACAAGGTGTTGCATTATCTGCTCAGAGACTTGGCTGTGATGCT
GTCATTGTTATGCCTGTTACTACACCAGACATCAAATGGAAATCAGTAAAGAGA
TTGGGTGCTACTGTTGTTCTTGTTGGGGGCTCATATGATGAAGCTCAAGCATATG
CCAAAAAGCGGGCTGAATCTGAAGGCCGCACATTCATTCCTCCTTTTGATCACCC
AGATGTCATTGTAGGGCAGGGTACAGTAGGGATGGAGATAAATCGCCAACTCA
AAGATCACATTCATGCAATCTTTGTGCCTGTTGGAGGAGGGGGGCTTATAGCTGG

TATTGCTGCTTATTTGAAAAGGGTTGCCCCAGATATAAAGATTATTGGAGTTGAA
CCACTTGATGCGAATGCATTGGCATTATCATTACATCATGGTCAGAGAGTAATGC
TGGACCAAGTTGGGGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAG
AGACTTATCGTCTCTGCGAGGAATTAATAGATGGCGTTGTCCTAGTTGGTCGTGA
TGCTATATGTGCATCTATAAAGGACATGTTTGAAGAGAAAAGGAGCATACTAGA
GCCTGCAGGTGCACTTGCTCTTGCTGGAGCCGAGGCATACTGCAAGTATTATGGC
CTCAAGGGTGAAGACGTAGTAGCAATAACTAGTGGAGCCAACATGAACTTTGA
CAGGCTTAGATTGGTAACAGAACTCGCAGATGTTGGTAGACAGCGGGAAGCTGT
TCTTGCTACTTTTATGCCAGAAGACCCAGGAAGCTTCAAAAAGTTCGCTGAAAT
GGTAGGACCAATGAATATCACTGAATTCAAGTACAGATACAACTCTGATAAAGA
AAGAGCTCTTGTACTTTACAGTGTTGGACTTCACACAATACTAGAACTTGAAGG
AATGGTGGAGAGGATGGAATCAGCAGATCTGCAAACCATTAATCTTACAGACA
ATGACTTGGTCAAAGATCATCTTAGGCATTTGATGGGTGGTAGAACAAATGTTC
ATAATGAGCTTTTGTGTCGATTCACTTTTCCTGAGAAGCCTGGTGCTTTGATGAAG
TTTTTAGATGCTTTCAGCCCACGTTGGAATATAAGTTTGTTTCATTATCGTGCACA
GGGAGATACTGGTGCAAATGTGCTAGTCGGGATCCAAGTTCCACAAGATGAGGT
AGTTGAGTTTGAGGGACGGGCAGACAGTCTTGGTTATGAATATGCAGTGGAGAG
TCTCAATGAGGCATACCAGCTCATAATGCAT

>Solanum_aviculare_TD1

ATGGAGGTTCTTCGTTTTACAGCCGTGAAGTCATTGAATTCCTGCGTCCGT
CCCGAATTCACGGCGACGAGCTTCGTCATTGTTCCTGTTAACACTGTGAAACCCT
CCGGAACAAGGAGAAGCAAAAGGAAGGCGTTCATTCGCGCTAAGGCGACGGA
GATCTTGTCATCTCCAGCTAAGGTAACGGAGCCGTTGCAGGCTGCGCCGGTAGR
GGCACCGTCGGCTCCGGTTCCTTTGCTACGAGTTTCTCCCAGCTCGTTGCAGTGT
GAACCGGGATATTTGATACCAAATAGTCCGGTTCTAGGTAGTGGTGGTTTGATCG
GGTACGAGTATTTGACAAACATCTTGTCGTCAAAAGTGTATGATGTAGCTTATGA
GACGCCTTTGCAAAAAGCGCCTAAGCTGTCAGAAAGATTGGGAGTTAATGTGTG
GCTCAAAAGAGAGGATCTTCAGCCCGTCTTCTCATTCAAAATCAGAGGAGCTTA
CAATATGATGGCAAAAATCCCTAAAGAGCAGTTGGAAAAAGGGGTTATATGCT
CATCAGCTGGAAACCATGCACAGGGTGTTGCATTATCCGCTCAGAGACTTGGCT
GCAATGCTGTGATTGCCATGCCTGTTACTACACCAGACATCAAATGGAAATCAG
TTAAGAGATTGGGTGCTACTGTTGTTCTTGTGGGGGACTCATATGATGAAGCTCA
AGCATATGCCAAAGAGCGGGCCGAAGCTGAAGGCCGCACATTCATCCCTCCTTT
TGATCACCCAGATGTCATCGTAGGGCAAGGTACAGTAGGGATGGAGATAAATC
GCCAACTCAAAGATAACATTCATGCAATCTTTGTGCCTGTTGGAGGAGGGGGCC
TTATAGCTGGTATCGCTGCTTATTTGAAAAGGGTTGCCCCTGATATAAAGATTAT
TGGAGTTGAACCACTTGATGCTAATGCGTTGGCATTATCATTACAGCATGGTCAG
AGAGTAATGCTGGACCAAGTTGGGGGTTTTGCAGATGGTGTAGCTGTTAAAGTG
GTTGGTGAAGAAACTTATCGTCTCTGCGAGGAATTAATAGATGGCGTAGTCCTA

GTTGGTCGTGATGCTATATGTGCGTCTATAAAGGACATGTTTGAAGAGAAAAGA
AGCATACTAGAGCCTGCAGGTGCACTTGCTCTTGCTGGAGCCGAGGCATACTGC
AAGTACTATGGCCTCAAGGGTGAAAACGTAGTAGCWATAACTAGTGGAGCCAA
YATKAACTTTGACAGACTTAGATTGGTAACTGAACTCGCAGATGTTGGTAGACA
GCGGGAAGCTGTTCTTGCTACTTTTATGCCAGAAGACCCAGGAAGCTTCAAAAA
GTTTGCTGAAATGGTAGGACCAATGAATATCACTGAATTCAAGTACAGATACAA
TTCTGATAAAGAAAGAGCTCTAGTACTTTACAGTGTTGGTCTTCACACCAAATTA
GAACTTGAAGAAATGGTGGAGAGGATGAAATCAGCAGATCTGCAAACCATTAA
TCTTACAGACAATGACTTGGTCAAAGATCATCTTAGGCATTTGATGGGTGGTAGA
ACAAATGTTCATGATGAGCTTTTGTGTCGATTCACTTTTCCTGAGAAGCCTGGTG
CTTTKATGAAGTTTTTAGATGCTTTCAGCCCACGTTGGAATATAAGTTTGTTTCAT
TATCGTGCACAGGGAGATACTGGTGCAAATGTGCTAGTCGGGATCCAAGTTCCA
CAGGATGAGGTAGTTGAGTTCCAGGGACGGGCAGAGAGCCTTGGTTATGAATAT
GCAGTGGAGAGCCTCAATGATGCATTCCAGCTCATAATGCAT

>Solanum_dulcamara_TD1

ATGGAGGTTCTTCGTTTTACAGCCGTGAAGTCATTGAATTCCTGCGTCCAT
CCCGAATTCACGGCGACGAGCTTCGTCATTGTTCCTATTAACACTGTGAAAGCCT
CCGGAACGAGGAGAAGCAAAAGGAAGGGGTTCATTCGTGCTAAGGCGACGGA
GATCTTGTCATCTCCAGCTACGGTAAGGGAGCCGTTGCAGGCGGCGCCGGTAGA
GGCACCGTCGGCTCCGGTTCCGTTGCTACGAGTTTCTCCCAGCTCGTTGCAGTGT
GAGCCGGGATTTTTGATACCGAATAGTCCGGTTCTAGGTAGTGGTGGTTTGATCG
GGTACGAGTATTTGACAAACATCTTGTCGTCGAAAGTGTATGATGTAGCTTATGA
GACGCCTTTGCAGAAAGCGCCTAAGCTCTCAGAAAGATTGGGAGTTAATGTGTG
GCTCAAAAGAGAGGATCTTCAGCCCGTCTTCTCGTTCAAAATCAGAGGAGCTTA
CAATATGATGGCAAAACTCCCTAAAGAGCAGTTGGAAAAAGGGGTTATATGCTC
ATCAGCTGGAAACCATGCACAAGGTGTTGCACTATCCGCTCAGAGACTTGGCTG
CGATGCTGTGATTGCCATGCCTGTTACTACACCAGACATCAAATGGAAATCAGTT
AAGAGCTTGGGAGCTACTGTTGTTCTTGTAGGGGACTCATATGATGAAGCTCAA
GCATATGCCAAAGAGCGGGCCGAAGCTGAAGGCCGCACATTCATCCCTCCTTTT
GATCACCCAGATGTCATCGTAGGGCAAGGTACAGTAGGGATGGAGATAAATCG
CCAACTCAAAGATAGCATTCATGCAATCTTTGTGCCTGTTGGAGGAGGGGGGCT
TATAGCTGGTATTGCTGCTTATTTGAAAAGGGTTGCCCCTGATATAAAGATTATT
GGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAG
AGAATAATGCTGGACCAAGTTGGGGGGTTTGCAGATGGTGTAGCTGTTAAAGTG
GTTGGTGAAGAGACTTATCGTCTCTGCGAGGAATTAATAGATGGCGTAGTCCTA
GTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAGAAAAGA
AGCATATTGGAGCCTGCAGGTGCACTTGCTCTTGCTGGAGCCGAGGCATACTGC
AAATACTATGGCCTCAAGGGTGAAAATGTAGTAGCAATAACTAGTGGAGCCAA
CATGAACTTTGACAGGCTTAGATTGGTAACTGAACTCGCAGATGTTGGTAGACA

GCGAGAAGCTGTTCTTGCTACTTTTATGCCAGAAGACTCAGGAAGCTTCAAAAA
GTTCGCTGAAATGGTAGGACCAGTGAATATCACTGAATTCAAGTACAGATACAA
TTCTGATAAAGAAAGAGCTCTAGTACTTTACAGTGTTGGTCTTCACACAACATTA
GAACTTGAAGGAATGGTGGAGAGAATGGAATCAGCAGATCTGCAAACCATTAA
TCTTACAGACAATGACTTGGTCAAAGATCATCTTAGGCATTTGATGGGTGGTAGA
ACAAATGTTCATAATGAGCTTTTGTGTCGATTCATTTTTCCCGAGAAGCCTGGTG
CTTTGATGAAGTTTTTAGATGCCTTCAGCCCACGTTGGAATATAAGTTTGTTTCAT
TATCGTGCACAGGGAGATACTGGTGCAAATGTGCTAGTCGGGATCCAAGTTCTA
CAGGATGAGGTAGTTGAGTTCCATGAACGGGCAGACAGTCTTGGTTATGAATAT
GCAGTGGAGAGTCTCAATGAGGCATTCCAGCTCATAATGCAT

>Solanum_bulbocastanum_TD2

ATGGAGTTTCTTTGTTTAGCCCCAACACATAGTTTTTCCATTAACCCCAAA
TCCACAAAAAATATTTCTATTGACCGTACTAGTACCACGGGTCGTATCATGAAA
ATGTACCAAAATATGCGTGGCTCAACGATAAGACCATCGGCATTGCCGTTAAAA
ATGAAACGATTAATTTCTCCTTCGGATTCTTCAGCTCCAGTGGTAAGTGTTCCTGC
AAACTTGCCAAAAGTCACCCCTGGGGAGTTGATAGTGAATAATCCAACGGGGG
GCGATAAGGATGAATTGTTTCAATATTTGGTTGACATTTTGGCATCGCCAGTGTA
CGATGTAGCGATTGAATCCCCATTAGAGCGCGCAACGAAGCTTTCGACTAGGCT
TGGAGTTAACTTTTATCTCAAAAGAGAGGATAAGCAGCTTGTGCTTTCATTCAAG
CTTAGAGGAGCTTACAACATGATGAGTAGCCTCCCAAGGGAGGAATTAGATAA
AGGGGTTATAACAGCATCAGCTGGAAATCATGCACAAGGTGTTGCATTAGCTGC
TGAGAGACTTAAATGTTATGCTAAAATTGTCATGCCTACAACCACCCCACAAAT
CAAGGTTGATGCGGTTAGAGCCCTGGGAGGTGATGTACATCTTGAGGGTCAGAC
ATTTGATGAAGCTCAAGAATTTGCACTGAAAGAGAGTCAAGAAAAAGGGTACA
CATTTATCCCAGCATATGATGCTCCAGAAATCATCAAAGGCCAAGGTACGGTTG
GAACGGAGATAAACCGTCAACTCAAGGACATTTATGCGGTGTTTGTACCTGTTG
GTGGTGGGGGTCTAATATCTGGTGTTGCTTCCTATTTCAAACAAGTTGCCCCAAA
TACAAAAATTATTGGAGTAGAGCCAACTGGTGCAGCTTCAATGACATTGTCATT
GTATGAAGGAGAAATAATAAAATTGGAAAATGTTGATACTTTTGCTGATGGTGT
AGCTGTTGCACAAGTTGGTGAATACACTTTTGCAAAATGCCAAGAGTTAATTGAT
GGAATGGTACTTGTTAGAAATGATGGTATTAGTGCTGCAGTAAAGGATGTGTAT
GATGAAGGAAGGAACATATTAGAGACATCAGGTGCAGTTGCCATAGCTGGAGC
TACAGCCTATTGCGAATACTATAACATTAAAAATGAAAATATTGTAGCAATTGC
AAGTGGAGCTAATATGGATTTCAGCAAACTACACAAAGTTACAGAGTTAGCAG
GGTTAGGTGCAGAAACAGAAGCTTTGCTAGCTACTTTTATGATAGAACAACCAG
GAAGTTTCAAGACATTTGCTCAACTTGTGGCCTCTATGAATATTACTGAAGTGAC
ATACAGATTCACTTCTGAAAGAAAACATGCTCTAGTTTTGTACAGAGTTGATGTT
GATGAAAAATCAGATCTTGAAGAAATGATTAAGAAATTGAATTCATCAAATATG
AATACTTTTAATTTCTCACATAATGAATTGGTTTCAGAACATCTAAAACATTTGA

TTGGTGGCTCAGCAAATATCAGTGATGAAATTTTCGGTGAATTTACTTTACCTGA
GAAGGCTGGAACTTTAAGCACATTCTTAGAAGCTTTCTCCCCACGTTGGAATATA
ACTTTATGCCGTTACCGTGATCAGGGAGATATAAATGGAAATGTGTTAGTGGGA
TTCCAAGTCCCTCAAAGTGAGATGGATGAGTTCAAAAGTCAAGCTGATGGACTT
GGCTATCCATATGACTTGACAATTTCAATGAGGCTTTTAACATTG

>Solanum_melongena_TD2

ATGGAATTTCTTTGTTTAGCCCCATCACATAATTTTACCATCAATCCCAAA
TTCACCCCACTCTCAATCAATCGCCTTGCTATCACTCTCCGTGAAATGAAATTGT
ATCGTAATAACATGCGTTGTACTAGGGTAAAACCATCGGCATTGCCATTAAAAT
CGTCAGGATTAGTTTACCCATCGGATTCTTCTGCCCCCTTGAGGGTGAATGAACC
TCTACCCGACCTCGAAAAATACAAGCCTGGAGAGTTGATAGAAAATCATCGAAT
TCCTTATGATCCAGATGAATTAAGGCAATATTTGCTTGATATATTGGCATCGCGA
GTCTATGATGTCGCGATCGAGTCCCCGTTGGAACGTGCAACAAAGATTTCAGAT
AAGCTTGACGTAAATTTTTTCATCAAAAGAGAGGATAGACAGCCTGTGTTTTCAT
TCAAGATTAGAGGAGCTTACAACATGATGAGCAGCCTCTCACAGGAAGTGTTAG
ATAAAGGGGTTGTAACAGCATCAGCTGGAAATCATGCACAAGGTGTTGCAGTAA
GTGCTAAAAGACTTAAATGTCAAGCTACAATTTGCATGCCTAAAACTACCCCAC
AAATCAAGGTTGACGCAGTTAGAACCCTAGGTGGTGAATTTGTTACAGTAGAAC
TTGAGGGTAACACATTTGATGAAGCTCAAGCTTATGCCCTGCAATTGGTTCAAGA
AAAAGGGTACAAATACATCCCACCATTTGATGATCCTGGAGTCATCAAAGGCCA
AGGTACAATTGGAGTGGAAATAAACCGCCAGCTTAAGGACATCCATGCAATATT
TATACCTGTTGGTGGTGGTGGTCTAATATCTGGTGTTGCTGCTTTTTTCAAACAAG
TAGCCCCAAATACAAAAATTATTGGGGTTGAACCATATGGTGCAGCTTCAATGA
CTGTGTCATTACTCGAAGGtAAAAGAATAAAATTGGACAATGTTGATACTTTTGC
TGATGGTGTTGCTGTTGCACTTGTTGGTGAATACAATTTTCAAAAATGCCATGAA
TTAATTGATGGAATGGTACTCGTCCATCGTGATGGTATTAGTGCTGGACTAAAGG
ATGTGTATGATGAAGGGAGGAACATATTAGAGACATCAGGTGCACTTGCCATAG
CTGGAGCTCAAGCATATTGTAAATTTTATGGCATTAAAAATGAAAATATTGTAGC
AATTGCAAGTGGAGCAAATATGGATTTCAGCAAACTAAAATTAGTTGCAGAATT
AGCAGAGATTGGTGCAGGAAAAGAAGCTTTGTTAGCTACTTTTATGCCAGAACA
AGTAGGAAGCTTCCACAAATTCATTAAATTGTTGGGCTCTTTCAATATTACTGAA
TTCACATACAGATATAACTCTGACGGAAAACAAGCTGTAGTTCTATACAGTGTT
GATATTGATCCTGAGCATCCAGAAAAAATTAAAGAAATAATTGATAAGATGAAT
TCAGAAGGATTCACAACTGTTGATCTCTCACATCATGATTTGGCCAAGGAACATC
TAAGGCATTTGGTTGCTGGTGGGGCGTCAAATCCTAGTGATGAGATTATCTGTCA
ATTTATTTTCCCTGAAATTGCTGGAGCGCTAAAAAGATTCTTAGATGCTTTCTCCC
CTCGTTGGAATATAACCTTGTTCCGCTATCGTGAACAGGGAGAAATAGATGCAA
GTGTGTTGGTGGGATTCCAAGTTCCACAAAGTGAGATGGAAGAGTTCAAAAAAC

AAGCTAATAAACTTGGTTATCCATATGCATTTGAGAGTCTTGATGAGGCTCAAAA
GCTTATAAGGAATGAA

>Jaltomata_procumbens_TD1*

TCATTGAATTCCTCCATCCGTCCCGAATTCACGCCGACGAAGACCTCCGT
CATTGTTCCCGTTAATACTGTGAAATGCAAAAAGAAGGCGTTTATTCGTGCTAAG
GCGGCGGAGATCTTGTCATCTCCAGCTAAAGCAACGGAGCCGTTAACGGCGGCA
CCGGTAGAGCCGCCGTCAGCTCCGTTGCTACGAGTTTCTCCGAGCTCGTTGCAGT
GTGAACCGGGACATTTGATACCGAATAGTCCGGTTTTAGGTAGTGGTGGTAAGA
TCGGGTACGAGTATTTAACAAACATCTTGTCGTCAAAAGTGTATGATGTAGCTTA
TGAGACGCCTTTACAGAACGCGCCTAAACTGTCAGAAAGGCTGGGAGTTAAAGT
GTGGCTCAAAAGAGAGGATCTTCAGCCTGTTTTCTCGTTCAAAAATCAGAGGAG
CTTACAATATGATGGCAAAACTCCCTAAAGAGCAGTTGGAAAAAGGGGTTATAT
GCTCATCAGCTGGAAACCATGCACAAGGTGTTGCATTATCTGCTCAGAGACTTG
GCTGCGATGCTGTGATTGCCATGCCTGTTACTACACCAGACATCAAATGGAAAT
CGGTTAAGAGATTGGGTGCTACTGTTGTTCTAGTGGGGGACTCATATGATGAAGC
TCAAGCATACGCCAAAGAGCGGGCCAAAGCTGAAGGCCGCACATTCATCCCTC
CTTTTGATCACCCAGATGTCATCGTAGGGCAAGGTACAGTGGGTATGGAAATAA
ATCGCCAACTCAAAGATAACATTCATGCAATCTTTGTGCCTGTTGGAGGAGGGG
GTCTTATAGCTGGTATTGCTGCTTATTTGAAAAGGGTTGCCCCTGATATAAAGAT
TATTGGAGTTGAACCACTTGATGCAAATGCATTGGCATTATCATTACAGCATGGT
CAGAGAGTAATGCTGGACCAAGTTGGGGGTTTTGCAGATGGTGTAGCTGTTAAA
GTGGTTGGTGAAGAAACTTATCGTATCTGCGAGGAACTAATAGATGGTGTAGTC
CTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAGAAA
AGAAGCATACTAGAGCCTGCAGGTGCACTTGCTCTCGCTGGAGCCGAGGCATAC
TGCAAGTACTATGGCCTCAAGGGTGAAAACGTAGTAGCAATAACTAGTGGAGC
CAACATGAACTTTGACAGACTTAGACTGGTTACGGAACTCGCAGATGTTGGTAG
ACAGAGGGAAGCTGTTCTTGCTACTTTTATGCCAGAAGAGCCAGGGAGCTTCAA
AAAGTTCGCTGAAATGGTAGGACCAATGAATATCACTGAATTCAAGTACAGATA
CAATTCTGATAAAGAAAGAGCTCTAGTACTTTACAGTGTTGGTCTTCACATGAAA
CTAGAACTTGAAGAAATGGTGGACAGGATGAAATCAGCAGATCTGAAAACCAT
TAATCTTACAGACAATGACTTGGTCAAAGATCATCTAAGGCATTTGATGGGTGGT
AGAACAAATGTTCATAATGAGCTTTTGTGTCGATTCACTTTTCCCGAGAAGCCCG
GTGCTTTGATGAAGTTTTTAGATGCTTTCAGCCCATGTTGGAATATAAGTTTGTTT
CATTACCGTGCACAGGGAGACACTGGTGCAAATGTGCTAGTTGGGATCCAAGTT
CCACAGGATGAGGTAGTTGAGTTCCAGAGACGGGCAGACAGTCTTGGTTATGAA
TATGCAGTGGAGAGTCTCAATGAGGCATACCAGCTCATAATGCATTGAAAAAGG
TATTGCCTGTTATGGCTTGTTCTGGGATCTCGTTGCAAAGAAAATAACGTGGTAT
TCGCCGACT

>Lycium_barbarum_TD1*

ACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCGAATTC
ACTAGTGATTATAACCCTTTGCCAACGACCTCCTCAGCCGCCGGGAAAAAAAAA
TCCACGTGTCCGGCTACCACCAATTCCATAACTTACACATTTAATCTCCATTACA
CGTGAAAAATGGAGGTTCTTCATTTCACCGCCGTTAAACCATTAAACTCCTACCT
CCGTCCCGAATTCTCGCCGTCAACCTCCGTCATTATCCCCATTAACACCGTTAAA
ACCTCCGTAACCAAAAAAAAGACATCTATTCGTGCTAAGGCAACGGAGATCTTA
TCATCTCCAACTTTTACCGTAACGGAACCGTTACCGGCGTCACCGGTAACGGCA
CCTTCGGCTCCGTTATTACGAGTTTCTTCGAGCGAGTTGCAGTGTGAACCGGGTT
ACTTGATACCGAATACTCCGGTTTTAGGTACTGGTGGTGTGGTTGGATACGAGTA
TTTGACGAGCATTTTGTCATCGAAAGTGTATGATGYAGCTTATGAGACGCCTTTA
CAGAAAGCGCCTAAGCTGTCAGAAAGGCTTGGAGTTAATGTGTGGCTTAAAAGA
GAGGATCTTCAGCCTGTCTTCTCGTTCAAAATCAGAGGAGCTTACAATATGATGG
CAAAACTGCCTAGGGAGCAGTTGGAAAGAGGGGTTATATGCTCATCAGCTGGA
AACCATGCACAAGGTGTTGCATTATCCGCTCAGAGACTCGGCTGTGATGCTGTG
ATTGCCATGCCTGTTACTACACCAGACATTAAATGGAAATCAGTTAAGAGATTG
GGTGCCACTGTTGTTCTAGTGGGGGACTCATACGATGAAGCTCAAGCATACGCC
AAAGAGCGGGCCGAAGCTGAAGGCCGCACATTCATCCCTCCTTTTGATCACCCA
GATGTCATCATAGGGCAAGGTACAGTAGGAATGGAGATAAATCGCCAACTCAA
AGAAGACATTCATGCAATCTTTGTGCCTRTTGGAGGAGGAGGTCTTATAGCTGGT
ATTGCYGCTTATTTGAAGAGGGTTGCCCCTGATATAAAGATTATTGGAGTTGAAC
CCCTTGATGCAAATGCAATGGCATTATCATTGCAGCATGGTCAGAGAGTAATGC
TGGACCAAGTTGGGGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGNGAAT
CGAATTCCCGCGGCCGCCATGGCGGCCGGGAGCATGCGACNTCGGGCCCA

>Solanum_dulcamara_TD2*

TAGGGCGAATTGGGCCCGACGTCCCATGCTCCCGGCCGCCATGGCGGCC
GCGGGAATTCGATTAGGAGGAATAGCTAAAGGGGTTATAACAGCATCAGCTGG
AAATCATGCACAAGGTGTTGCAGTTAGTGCTGCAACACTTAAATGTCAGGCTAC
AATTGTCATGCCTGAGACCACCCCAGAAATCAAGGTTGAGGCGGTTAGATCCTT
AGGTGGTGATTATGTCACGATAAAACCTGTGGGTCAGACATTTGATGAAGCTCA
AGCAGAGGCCCTGAGGTTGGTTGAAGAAAATGGGTACATATTCGTCTCACCATT
CGATGATCCAGGAGTCATCAAAGGCCAAGGTACAGTTGGAATGGAGATAAACC
ATCAACTTAAGGACATCTATGCCGTGTTTGTGCCTGTTGGTGGTGGTGGTCTAAT
AGCTGGTGTTGCTGCTTATTTCAAACAAGTTGCCCCAAATACAAAAATTATTGGA
GTTGAACCATATGGTGCAGCTTCAATGACATATTCATTATATAAAGGTGAAAGA
ATAAAATTGGAAAATGTTGATACTTTTGCTGATGGTGTAGCTGTTGCACAAGTTG
GTGTATACAATTTTCAAAAATGTCAAGAGTTAATTGATGGAATGGTACTTGTTCA
TCGTGATGATATTAGTGCTGCAATAAAGGATGTGTATGATGAAGGAAGGAACAT
ATTAGAGACATCAGGTGCAGTTTCCATAGCTGGAGCTCAAGCATATTGCAAATA
TTATAACATCCAAAATGAAAATATTGTAACAATTGCAAGTGGAGCAAATATGGA

TTTCAGCAAACTAAAATTAGTTGCAGACTATGCAGCGATTGGTGCAGGGGAAGA
AGCTTTGTTAGCTACTGTTATGCCAGAAGAACAAGGAAGCTTCTTGAAATTCATT
CAAATTTTGGGCTCTTTCAATATTACTGAATTCACATACAGATATAGTGCTGAAA
GAAAAAAGGCTGTAGCTCTATACAGTGTTGATATTGATGAAAATACGGATATTA
AAGCAATAATTGAGAAGTTGAATTCAGAAGGATTCCATACTGTTGATCTCTCAC
ATCATGATGTGGCCAATGAACACCTAAGGCATTTGGTGGCTGGTGGCGCATCAA
ATCCTAGTGATGAGATTCTCTGTCAATTTATTTTCCCTGAAATTGCTGGAGCTCTA
AGAAAATTCTTATATGCTTTCTCCCCTCGTTGGAATATAACTTTGTTCCGTTACCG
TGAACAGGGAGATATAAATGCAAGTGCGTTAGTGGGATTCCAAGTTCCAAAAA
GTGAGATGGATGAATTCCAAAATCAAGCTAATAATCTTGGTTATGCATATAATTT
TGAGAGTCTTGATGAGGCTTATAAGCTTATAAGGAGTGATTGAAGAAAAAAAAT
AGTGTTCTAAGTTGCAATTATTGCAAATCCTACTTTTTTTTCTTTAAAAGAATAA
CCGTCTTGTACGAATCAGTTGTTAAGAAATAATTGATGAATTAACTTGTTGTAAG
TTAATTTCATCTTGTTGTATAAATAATAATGCATGTATCCAATGTAATGGCAATC
ACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGC
GTTGGA

>Jaltomata_procumbens_TD2*

AGGAGGAATTAGCTAAAGGGGTTATAACAGCATCAGCTGGAAATCATGC
ACAAGGTGTTGCATTAGGTGCTCAAAGACTTAATTGTACTGCTAAAATTGTCATG
CCTGTTACCACCCCTGTGATCAAGAGAGAAGCAGTTGAAAACTTGGGTGGTATA
GTAGTTCTTGAGGGTGACACATTTGATGAAGCTCAAGCACATGCACTCACTTTGT
CTCAAGAAGAAGGGCTCACATTCATCCCACCATTCGATCATCCAGATGTCATCA
AAGGCCAAGGTACAATTGGAACGGAGATAAACCGTCAACTTAAGGACATACAT
GCAGTATTTGTGCCTGTTGGTGGTGGTGGTCTAATAGCTGGTGTTGCTGCCTTTTT
CAAGAAAGTTGCACCAAATACAAAGATTATTGGTGTTGAACCATTTGGTGCAAG
TTCAATGACATTGTCATTGTACAATGGTGAAAGAATAAAGTTGGAACAAGTTGA
TAATTTTGCAGATGGTGTTGCTGTTGCACTTGTTGGTGAACACACTTTTAGAATTT
GTCAAGATTTGATTGATGGTATGGTACTTGTTGGTCGTGATGCTATTAGTGCTGCA
GTAAAGAATGTTTATGAAGAAGGAAGGAACATATTAGAGACATCAGGTGCACT
TGCCATAGCTGGAGCAGAAGCATATTGCAATTACTATAAGATTAAAGATGAAAA
TATTGTAGCAATTGCAAGTGGAGCAAATATGGACTTTAGCAAACTAAAATTAGT
AGTGGAGTTAGCAGATATTGGTGCACAAACAGAAGCTCTATTAGCTACTTTTATG
CCAGAAGAACCAGGAAGCTTCAAGAAATTCTGTCAACTTGTGGGCTCTTTGAAT
ATTACTGAATTCAGATACAGATTTAATTCTGGAAGAAAACGAGCTCTAGTTCTAT
ACAGTGTTGATTTTTCAACAAAATTGGATCTTGAAGCAATGTTGGAGAGAATGA
AATCATCAGAAATGAATACTGTTAATCTCTCAAATAATGATTTGGTCAAAGAAC
ATCTAAGGCATTTGGTGGGTGGCACATCAAATCCTAGTAATGAGATTTTCTGTCA
ATTTATTTTTCCTGAGAAGCCTGGAGCTTTAAGAAAATTCTTAGATGCTTTTTGCC
CTCGTTGGAATATAACTTTGTTCCATTATCGCGAACAGGGAGAACTTGATGCAAG

TGTGTTGGTGGGATTCCAAGTTCCAAAAAGTGAGATGGAAGAATTTCAAAATCA
AGCTAACAATCTTGGTTATTCATATGAAATTGAGAGTCTCAATGAGGCTTATAAT
CTTATAATTGAGTGAAGAAAAAAAAATGGTGTTCTAAGTTGCAATTATTGCAAA
TCCGAAAAAAAAAATAACCGTGTTGTACGGATCAGCTT

>Lycium_barbarum_TD2*

CTATAGGGCGAATTGGGCCCGACGTCGCATGCTCCCGGCCGCCATGGCG
GCCGCGGGAATTCGATTCCCTTCAACTTAGTGCAAGTGATATCAGGAAAATAAA
AATTTATCTATCTACTATCATAGTGAGTAAAGATTTCGATCAATTATAGTACAAA
AATATGGTTGCTCTTTGTCTAGCTCCAACACATAATTTTACCAGCAACCCCAAAT
CTACAACAATTAGTCGAACCGCGGCAATTCCTGCTGGCCGCAATGACACATTTA
AAATATCGACTACTGGCCGTAACATGAAAATGGGTGGCTCAAGGGTGAGACCA
TCGGCATTGCCGTTGAAGATGTCACAATTAATTTCTCTATCTGACTCTTCAGCACC
GGTGGTGAGTGCTCCTGCATCCTTGCAGGACATCGAGGCTGGAAAATTGATAGA
GAATAATCCAACGGGTGGCGATAGTGAAGAATTGTTCCAGTACTTGCTTGAAAT
ATTGTCTTCGAGAGTGTACGATGTTGCTATCGAGTCCCCGTTGCAGCATGCTACA
AAGCTTTCAAAGAGGCTTGGAGTTAACCTTTTTATCAAAAGAGAGGATATGCAG
TCCGTGTTTTCGTTCAAGCTTAGAGGAGCTTACAACATGATGTCCAAACTTCCAA
GGAGCAATTAGAAAGAGGGGTTATAACAGCATCAGCTGGAAATCACGCACAAG
GAGTAGCATTAGGTGCTCAGAGACTTCAATGTACTGCTAAGATTGTCATGCCTGT
TACCACCCCTGTGATCAAGAGGGAAGCAGTTGAAAACTTAGGTGGTACAATAGT
ACTTGAGGGTGACACATTTGATGAAGCTCAAACATATGCATTGAAGTTGGCTGA
AGAGGAAGGTCTCACATTCATCCCACCATTCGATCATCCAGATGTCATCATAGG
TCAAGGTACAATTGGAACAGAGATAAACCGTCAACTTAAGGATATACATGCAGT
ATTTGTGCCTGTTGGTGGTGGGGGTCTATTAGCTGGTGTTGCTGCCTTTTTCAAAA
AAGTTGCCCCTAACACAAAGATTATTGGAGTTGAGCCATATGGTAATCACTAGT
GAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGA
TGCATAGCTTGAGTATNCTATA

>Solanum_pseudocapsicum_TD2*

ATGGAATTTCTTTGTTTAGCCCCATCATCACATAATTTTACCACCAGCCCC
AAACTTGCCCCATTTTCTACGAATCGCATTATTACTACCACCACTCTCCGTGACA
TGAAAATGTATCGTAACATGCGTGGCTCAAGGGTAAAAACATTGGCAAAGCCAT
TAAAATTATCAAACTTATTTTATCCTTCGAATTTGGCGAGTGACCCTTCATGCATC
GTGCCTGAGCCCGGAGAGTTGGTACCACATCAGCCAACCGGGGGCGATAGAGA
AGAGTTGTTACAATATTTATTTGACATGTTGGCATCGCCAGTCTACGATGTCGCG
ATCGAATCCCCGTTGGAAAAGGCAGAGAAGATTTCGACAAAGCTTGGAGTAAA
TTTTTTTATCAAAAGAGAGGATAAGCAGAAAGTATTTTCATTCAAGCTTAGAGG
AGCTTACAACATGATTAGCAAACTCCCAGAGGAGCAATTAAAGAAAGGGATTA
TAACAGCATCAGCTGGAAATCATGCACAAGGTGTTGCATTAAGTTCTTATAAAA
AATGTCCAGCTACAATTGTCATGCCTACTACCACCCCACAAATCAAGATTGATG

CAGTTAGAAGCTTGAATCCTGACGCTGATGTACCATTAAAGGGGGAAAATTTTG
ATGANGCTCTCAATTGCATTGGATATGGCTGACAAGCAAGGGCTCACATTCATC
CCACCATATGATCATCCAGAAATCATCAAAGGCCAAGGTACGATTGGAGTGGA
GATAAACCGTCAACTTAAGGACATTTATGCAGTATTTGTACCTGTTGGTGGTGGG
GGTCTAATATCTGGTATTGCTGCATATATGAAACAAATAGCCCCAAAGACAAAA
ATTATTGGAGTTGAANCCATATGGTGCAGCTTCAATGACTTTGGCATTATGNTAN
GGTAAGCCAGAAAAGTTGGAAAGGTGGATAATTTTGCTGATGTGTAGCTGTGGG
AAANGGCTGGGGAGTACACTTTTAAAATATGTAATGAGTTAATTGATGGAATGG
TACTCATGCATCGTGAGGGTATTAGTGCTGGATTAAAGGATGTGTATGATCAAG
GAAGGGAAATATTAGAGACATCAGGTGCACTTGCCATAGCTGGAGCTGAAGCA
TATTGCAACTATTATAACATCAAAGATAAAAATATTGTTGCAATTGCAAGTGGA
GCAAATATGGATTTCAGCAAATTAAAATTAGTTGCAGAGTTGGCCCCAATTGGT
GCAGGGACAGAAGCTTTGTTAGCTACTACTGTGCCAGAAAAATCAGGAAGCTTC
CTCGAATTCACTAAAGTGTTGGGCTCTTTCAGTATTACTGAATTCACATACAGAT
ATAGTGATCAAAGAAAACATGCTGTATTTCTATACAGTATTGATCTAACTGCTAA
AGATCATCCTGATGTTGACAATGTAATTGCTAAGATGAATTCAGCAGGATTCGA
CACTGTTGATATCTCACATCATGATTTGGCCAAAGAACATCTAAGGCATTTGGTT
GCTGGTGGAGTATCAAGTGTTAGTGATGAAATTTTCGTTCAATTTGTTTTCCCTGA
GATTCAAGGAGCTTTGAACAAATTCTTAAATGCATTCTGTCCTCGTTGGAATATA
ACTTTGTTCCGTTATCGCAGCCAGAGAGAAATTGATGCAAGTGTTTTGGTGGGAT
TCCAAGTTCCACAAAGTGAGATGGATGAATTCAAAAATCAAGCCAATAAACTTG
GTTATTCATATGCATTTGAGAGTCTTGATGAGGCAGCTAAGCTAATAGTTGAGTG
ACCAAACAAAAAATAAATGATATTCTAAGTTGCAATTATTGAAAATCCTACTTTT
TTTGCTTTAAAAGAATAATCGTGTTGTACAAGTCAGGTGTTAAGAAATAATTGAT
GAATTAACTTGTTGGAAGTTAATTTCATCTTGTTGTACGAACAATAATGCATGTA
TCCAATGTAATG

References

Bridges, C.B., 1936. THE BAR "GENE" A DUPLICATION. *Science*, 83(2148), pp.210–211.

Cameron, D., 1952. Inheritance in Nicotiana tabacum. XXIV. Intraspecific differences in chromosome structure. *Genetics*, 37(3), p.288.

Chen, H. et al., 2005. Jasmonate-inducible plant enzymes degrade essential amino acids in the herbivore midgut. *Proceedings of the National Academy of Sciences of the United States of America*, 102(52), pp.19237–19242.

Chen, H. et al., 2007. Stability of Plant Defense Proteins in the Gut of Insect Herbivores. *PLANT PHYSIOLOGY*, 143(4), pp.1954–1967.

Fang, S. et al., 2009. Molecular evolution and functional diversification of fatty acid desaturases after recurrent gene duplication in Drosophila. *Molecular Biology and Evolution*, 26(7), pp.1447–1456.

Fitch, W., 1971. Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Biology*.

Force, A. et al., 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, 151(4), pp.1531–1545.

Gonzales-Vigil, E. et al., 2011. Adaptive evolution of threonine deaminase in plant defense against insect herbivores. *Proceedings of the National Academy of Sciences*.

Hughes, A.L., 1994. The evolution of functionally novel proteins after gene duplication. *Proc. R. Soc. Lond. B*

Innan, H. & Kondrashov, F., 2010. The evolution of gene duplications: classifying and distinguishing between models. *Nature Reviews Genetics*, 11(2), pp.97–108.

Kang, J.H. et al., 2006. Silencing Threonine Deaminase and JAR4 in Nicotiana attenuata Impairs Jasmonic Acid-Isoleucine-Mediated Defenses against Manduca sexta. *THE PLANT CELL ONLINE*, 18(11), pp.3303–3320.

Khersonsky, O. & Tawfik, D. S., 2010. Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective. *Annual Review of Biochemistry*, *79*(1), 471–505. doi:10.1146/annurev-biochem-030409-143718

Long, M. et al., 2003. The origin of new genes: glimpses from the young and old. *Nature Reviews Genetics*, 4(11), pp.865–875.

Marais, Des, D.L. & Rausher, M.D., 2008. Escape from adaptive conflict after duplication in an anthocyanin pathway gene. *Nature*, 454(7205), pp.762–765.

Mojzsis, S.J. et al., 1996. Evidence for life on Earth before 3,800 million years ago. *Nature*, 384(6604), pp.55–59.

Ohno, S., 1970. *Evolution by gene duplication* 1st ed, Springer-Verlag.

Olmstead, R.G. et al., 2008. A molecular phylogeny of the Solanaceae. *TAXON,* 57 (4), 1159–1181

Rodriguez, F., Wu, F., Ané, C., Tanksley, S., & Spooner, D. M. (2009). Do potatoes and tomatoes have a single evolutionary history, and what proportion of the genome supports this history? *BMC Evolutionary Biology*, *9*(1), 191. doi:10.1186/1471-2148-9-191

Sayers, E.W. et al., 2012. Database resources of the National Center for Biotechnology Information. *Nucleic acids research*, 40(Database issue), pp.D13–25.

Swofford, D. L., 2003, PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.

Tamura, K. et al., 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*, 28(10), pp.2731–2739.

Umbarger, H. E., 1956. Evidence for a negative-feedback mechanism in the biosynthesis of isoleucine. *Science*, *123*(3202), 848.

Yang, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, 24(8), pp.1586–1591.

Yang, Z., 2012. PAML 4: phylogenetic analysis by maximum likelihood (Manual). Version 4.6. pp. 31.

Yang, Z. et al., 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution,* 22:1107-1118.