Tropane alkaloids and terpenes synthase genes of *Datura stramonium* (Solanaceae)

Sabina Velázquez-Márquez, Iván M. De-la-Cruz, Rosalinda Tapia-López and Juan Núñez-Farfán

Laboratorio de Genética Ecológica y Evolución, Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad de México, DF, Mexico

ABSTRACT

Background. Plants have evolved physical–chemical defense to prevent/diminish damage by their enemies. Chemical defense involves the synthesis' pathways of specialized toxic, repellent, or anti-nutritive metabolites to herbivores. Molecular evolutionary studies have revealed the origin of new genes, acquisition and functional diversification along time in different plant lineages.

Methods. Using bioinformatic tools we analyze gene divergence of tropane alkaloids (TAs) and terpene synthases (TPSs) in *Datura stramonium* and other species of Solanaceae; compared gene and amino acids sequence of TAs and TPSs on genomes, cDNA and proteins sequences of Viridiplantae. We analyzed two recently assembled genomes of *D. stramonium* (Ticumán and Teotihuacán), transcriptomes of *Datura metel* and genomes of other Solanaceae. Hence, we analyzed variation of TAs and TPSs to infer genes involved in plant defense and plant responses before stress. We analyzed protein modeling and molecular docking to predict interactions between H6H and ligand; we translated the sequences (Teo19488, Tic8550 and Tic8549) obtained from the two genomes of *D. stramonium* by using Swiss-Model and Ramachandran plot and MolProbity structure validation of Teo19488 protein model.

Results. For TAs, we detected an expansion event in the tropinone reductase II (TRII) and the ratio synonymous/non-synonymous substitutions indicate positive selection. In contrast, a contraction event and negative selection was detected in tropinone reductase I (TRI). In Hy-oscyamine 6 b-hydroxylase (H6H), enzyme involved in the production of tropane alkaloids atropine and scopolamine, the synonymous/non-synonymous substitution ratio in its dominion indicates positive selection. For terpenes (TPS), we found 18 DsTPS in *D. stramomiun* and seven in *D. metel*; evolutionary analyses detected positive selection in TPS10.1 and TPS10.2 of *D. stramonium* and *D. metel*. Comparison of copies of TPSs in *D. stramonium* detected variation among them in the binding site. Duplication events and differentiation of TAs and TPSs of *D. stramonium*, as compared to other Solanaceae, suggest their possible involvement on adaptive evolution of defense to herbivores. Protein modeling and docking show that the three protein structures obtained of DsH6H from Teo19488, Tic-8550 and Tic8549 maintain the same interactions and the union site of 2OG-FeII_Oxy with the Hy-o ligand as in 6TTM of *D. metel*.

Conclusion. Our results indicate differences in the number of gene copies involved in the synthesis of tropane alkaloids, between the genomes of *D. stramonium* from two Mexican populations. More copies of genes related to the synthesis of tropane alkaloids

(TRI, TRII, H6H, PMT) are found in *D. stramonium* as compared to Viridiplantae. Likewise, for terpene synthases (TPS), TPS-10 is duplicated in *D. stramonium* and *D. metel*. Further studies should be directed to experimentally assess gain (overexpression) or loss (silencing) of function of duplicated genes.

INTRODUCTION

Plants are exposed to manifold environmental factors, biotic and abiotic, that affect their lifetime reproductive success. To cope with the different types of environmental stress, plants have evolved different mechanisms, ranging from morpho-anatomical, physiological, biochemical, molecular and epigenetic modifications, among others (*Lamalakshmi et al., 2017*). An ubiquitous defensive mechanism that protects plants from physical and biotic stresses is represented by the synthesis of specialized metabolites (SM; the so-called "secondary metabolites"). Thousands of chemical compounds found in plants promote protection against plants' natural enemies (i.e., pathogens, viruses and herbivores) (*Mithöfer & Boland, 2008*). More than 40,000 and 12,000 terpenoids and alkaloids, respectively, have been described in plants (*Zhou & Pichersky, 2020*).

Tropane alkaloids (TAs) and Terpene synthases (TPSs) are important specialized molecules of plants that help to protect them from herbivores and pathogens (*Kessler & Baldwin, 2002*). Tropane alkaloids are distinctive, but not exclusive, of Solanaceae (*Wink, 2003*; *Mithöfer & Boland, 2008*; *Pigatto et al., 2015*) and in *Datura stramonium,* implicated in plant defense to herbivores (*Shonle & Bergelson, 2000*; *Castillo et al., 2014*; *Miranda-Pérez et al., 2016*). Terpene synthases are enzymatic genes involved in the synthesis of volatile organic compounds (VOCs) in many organisms (*Picazo-Aragonés, Terrab & Balao, 2020*). Alkaloids and VOCs negatively affect herbivores. The concentration of VOCs may even rise before herbivores start to feed (*Heil, Lion & Boland, 2008*). Recent molecular studies in plants indicate modifications in the expression of SM as defense mechanisms (*Mithöfer & Boland, 2008*). Such modifications may result from the composition of the herbivore community, their abundance and dominance (*Koerner et al., 2018*). In this study we present a bioinformatic analysis of the changes in genes involved in the synthesis of tropane alkaloids and terpene synthases in Solanaceae, with particular attention to *Datura stramonium*. Gene modifications may include changes in the regulatory sequence or in the dominion, in the length of genes (including duplication), and post-transduction modifications (*Defoort, Van de Peer & Carretero-Paule, 2019*). Here, we analyze events of duplication in Solanaceae, with reference to whether duplicates are involved in modifications of biosynthetic pathways of TAs and TPSs.

Generally, copies of genes can have different fates. If different copies are retained, they may undergo sub-functionalization, dividing the original function. Alternatively, one copy may undergo neofunctionalization, acquiring a new function. However, the most frequent

outcome is pseudogenization, whereby one copy becomes non-functional (pseudogene) and tends to disappear due to accumulation of deleterious mutations. A pseudogene shows strong similarity to the parental copy, which encodes a particular protein, but has one or more alterations such as premature stop codons, mutations that cause phase changes, deletions and/or insertions that prevent a protein from being functional (*Ohno, 1970*). It has been reported that in some species the retention of pseudogenes for a long time only accumulates neutral mutations. However, several questions arise regarding the survival mode of these inactivated genes (*Ohno, 1970*; *Lynch & Conery, 2003*; *Conan & Wolfe, 2008*; Wang et al., 2012). Some retained functions maintain developmental homeostasis, for instance, the ability to rapid respond to a wide variety of environmental cues (Huot et al., 2014). One first defense response of plants involves the activation of signaling mechanisms (*De-la-Cruz, Velázquez-Márquez & Núñez Farfán, 2020*), and therefore protein turnover. It is estimated that more than 80% of the proteins are degraded through the proteosome pathway (*Maltsev et al., 2005*). In recent years it has been reported that the responses of specialized metabolites are grouped into several multigenic families. Therefore, we aim to elucidate the evolutionary process of two gene families, TAs (Tropane alkaloids) and TPSs (terpenes synthases) (*Karunanithi & Zerbe, 2019*; *De-la-Cruz et al., 2021*). Studies conducted in Solanaceae have found rapid responses to biotic and abiotic stimuli, in a matter of minutes, of genes involved in the synthesis of specialized metabolites (atropine, scopolamine, VOCs) (*Brille, Loreto & Baccelli, 2019*). To date, several groups have reported different TAs and TPSs and have found similar results (*De-la-Cruz, Velázquez-Márquez & Núñez Farfán, 2020*).

In this study we analyze the evolutionary history of genes involved in the synthesis of tropane alkaloids and terpene synthase in the Solanaceae family and specifically in draft genomes of two plants of *Datura stramonium* from populations of México (Ticumán and Teotihuacán). Overall, our bioinformatic analyses infer duplication of the genes of the main biosynthetic pathways leading to specialized metabolites.

MATERIALS AND METHODS

Selection and bioinformatic analyses: identification of TAs and TPs, genomics, transcripts and protein sequences

We obtained the sequences of nucleotide, transcripts and proteins of TAs and TPSs. Nine genomes (Table 1) of Solanaceae species were sourced for protein coding genes and CDS (Coding Sequence) genomes from the Sol Genomics Network (see links in Table 1): *Nicotiana tabacum* (*Edwards et al., 2017*), *Nicotiana sylvestris* (*Sierro et al., 2013*), *Nicotiana attenuata* (*Xu et al., 2017*), *Nicotiana tomentosiformis* (*Sierro et al., 2013*), *Solanum pimpinellifolium* (*Razali et al., 2018*), *Solanum lycopersicum* (*The Tomato Genome Consortium, 2012*), *Solanum pennellii* (*Bolger et al., 2014*), *Solanum tuberosum* (*The Potato Genome Sequencing Consortium, 2011*), *Capsicum annuum*, CM334 v1.55 (*Kim et al., 2014*). In addition, data of *Datura stramonium* were extracted from protein whereas TAs and TPSs' sequence were obtained from the genome of *D. stramonium* from the Mexican populations of Ticumán, Morelos (18°44′28.19″N,

Specie	Link	Link
Nicotiana tabacum	https://solgenomics.net/organism/Nicotiana_tabacum/genome	https://solgenomics.net/tools/blast/
Nicotiana sylvestris	https://solgenomics.net/organism/Nicotiana_sylvestris/genome	https://solgenomics.net/tools/blast/
Nicotiana attenuata	https://solgenomics.net/organism/Nicotiana_attenuata/genome	https://solgenomics.net/tools/blast/
Nicotiana tomentosiformis	https://solgenomics.net/organism/Nicotiana_tomentosiformis/genome	https://solgenomics.net/tools/blast/
Solanum pimpinellifolium	https://solgenomics.net/organism/Solanum_pimpinellifolium/genome	https://solgenomics.net/tools/blast/
Solanum lycopersicum	https://solgenomics.net/organism/Solanum_lycopersicum/genome	https://solgenomics.net/tools/blast/
Solanum pennellii	https://solgenomics.net/organism/Solanum_pinnelii/genome	https://solgenomics.net/tools/blast/
Solanum tuberosum	https://solgenomics.net/organism/Solanum_tuberosum/genome	https://solgenomics.net/tools/blast/
Capsicum annuum	https://solgenomics.net/organism/Capsicum_annuum/genome	https://solgenomics.net/tools/blast/

99°7′44.26″W), and Teotihuacán, Estado de México (19°41′16.58″N, 98°50′4.14″W) (https://github.com/icruz1989/Datura-stramonium-genome-project) (*De-la-Cruz et al., 2021*) and transcriptomes (https://medplantrnaseq.org/).

We downloaded the whole genome of *D. stramonium* to extract the TPSs genes and gene prediction was performed using AUGUSTUS (*Stanke et al., 2006*). All sequences were uploaded and mapped using as reference species of Solanaceae. A total of 366,000 genes were extracted, and gene redundancy was reduced with CDHIT (*Chen et al., 2016*). The selected TPSs genes were blasted against the NCBI, UNIPROT databases and the nine genomes of Solanaceae, using a threshold value of $1E^{-5}$. All TPSs were downloaded from Viridiplantae followed by a protein-protein blast; hits above 80% of similarity and 70% of cover were saved for further analyses. Meta-alignments of protein sequences were performed using T-coffee (*Wallace et al., 2006*), MUSCLE v3.8 (*Edgar, 2004*) and MAFF (https://mafft.cbrc.jp/alignment/server/) v6. (*Rozewick et al., 2019*).

Identification of orthogroups and inference of phylogenetic tree of TPa and TPSs

Construction of orthogroups, gene families, was performed according to *De-la-Cruz et al. (2021)* using Orthofinder v2.3.3 (*Emms & Kelly, 2015*; *Emms & Kelly, 2019*). The phylogenetic tree inference was accomplished using the program BEAST with the pair-wise deletion option, and reliability of the obtained phylogenetic tree was tested by bootstrapping with 1,000,000 MKKM chains. The model of substitution selected was JTT+G+F, with a branch support of 1 (*Drummond et al., 2002*). The trees were edited at the interactive tree of life (https://itol.embl.de) (*Letunic & Bork, 2019*).

We downloaded all orthologous sequences of TAs and TPSs of *D. stramonium*, *D. metel* and nine species of Solanaceae (see Table 1) to construct the orthogroups. Downloaded sequences of proteins and ORFs were extracted using TransDecoder and corroborated with GenScan. TPS genes were translated with GenScan and checked for the right direction and correct CDS; we corroborated this with the NCBI. The database used for BLAST was Uniprot; all available data for the analyzed plant species were downloaded. These amounts 847,544 sequences of TAs and TPSs. BLAST was carried out with a filtering E-value $< 1e^{-5}$ and only hits above 85% were selected (Data S1, S2). Orthogroups were

inferred using Orthfinder v 2.3.3. DIAMONS (*Emms & Kelly, 2015*; *Emms & Kelly, 2019*) while OrthoMCL was used to assess orthologs among species (*Chen et al., 2006*).

Protein modeling

Molecular modelling and protein-ligand docking was performed with GLIDE software (*Schrödinger Release 2021-1, 2021*). Simulations of interactions for each protein was done in a quadrat of 6LU/ processed with the same coordinates of the crystalized ligand as described in *Wallace & Zhang (2020)*.

The SWISS-MODEL (http://www.expasy.org/swissmod/SWISS-MODEL.html) (*Waterhouse et al., 2018*) was employed to model the tertiary structure of proteins, from the previously determined structure by the dominion DIOX_N (PF14226) (*Hagel & Facchini, 2010*) as a blueprint of modelling. Models were visualized and analyzed through Visual Molecular Dynamics (*Waterhouse et al., 2018*; *Bienert et al., 2017*)

Molecular docking

The in-silico docking (AutodockVina) of the PDB files (produced by the SWISS-Model) of Tic8550 (Ticumán) and Teo19488 (Teotihuacán) was performed in a simulation of molecular coupling with the native ligand of H6H, which is Hy-oscyamine (Hy-o). For docking, the protein was prepared with AutodockTools, assigning the parameters of energies and force fields of AutoDock4.0. Then, we obtained the results of the thermodynamically favored conformation for the Hy-o ligand using the algorithm VINA. The selected conformation had affinity coupling parameters of −7.3 and −7.5 kcl/mol for Teo19488 and Tic8550, respectively. As positive control, we performed the re-docking of the molecule of reference 6TTM (*Datura metel*), eliminating the ligand of the crystallized structure; the adopted conformation is similar to that obtained by the method of crystallization, and the coupling affinity for the reference model was −7.9 kcal/mol. The produced models are theoretically correct, and it is predicted that these have an oxidoreductase and use Hy-o as substrate.

Tropane alkaloids

Analyses detected variation in the number of TAs gene copies in *D. stramonium*: DsTRI, DsTRII, DsPMT and DsH6H. Compared to Viridiplantae *D. stramonium* has one or more copies of these genes suggesting a genome expansion of TAs.

We detected eight copies of tropinone reductase I (DsTRI) and four copies of DsTRII. TRI and TRII define two groups in Solanaceae (Fig. 1A). The gene DsTRII have had duplications events in *D. stramonium* (Fig. 1B). The first duplication occurs in Rosidae-Asteridae, followed by three duplications at the base of Solanaceae and one more in *D. stramonium* (Fig. 1B). The phylogeny and structure of 32 protein sequences (TRI) from different plant species show variability in the ADH_Short_C2 TRI dominion (Fig. 2); in *D. stramonium* (Tic23_dati33027) there is a duplication of this dominion but shorter (Fig. 2).

In Hy-oscyamine- 6-ß-hydroxylase (*DsH6H*) in the Ticumán genome we detected two copies of the gene, Tic8550 and Tic8549 (Fig. 3A). Tic8550 has a tandem duplication of the DIOX-N dominion of 79 amino acids (Fig. 3) whereas Tic8549 has the dominion 2-oxaglutarate (65 amino acids) (Fig. 3A). In contrast, in the Teotihuacán genome, *DsH6H*
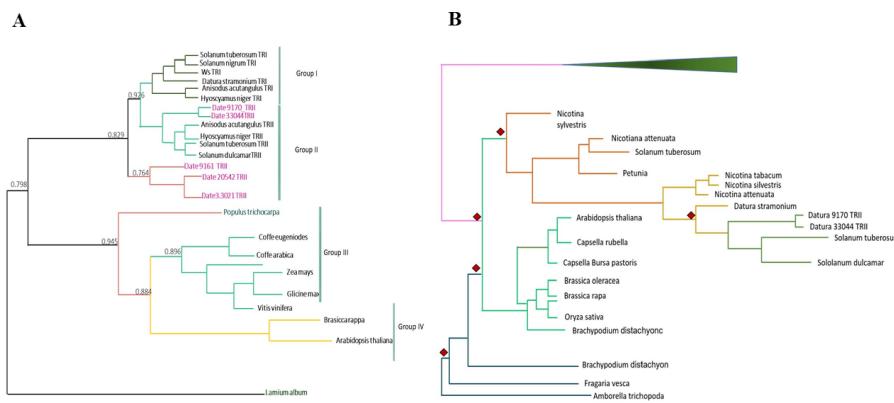
Figure 1 **Phylogenetic analysis of TRI and TRII protein sequences.** TR-I and TR-II define Group I and II, respectively, in Solanaceae (A). Substitution rates (ω) of *TR-II* (B). Three duplication events in the genome of *Datura stramonium* are indicated (diamonds).
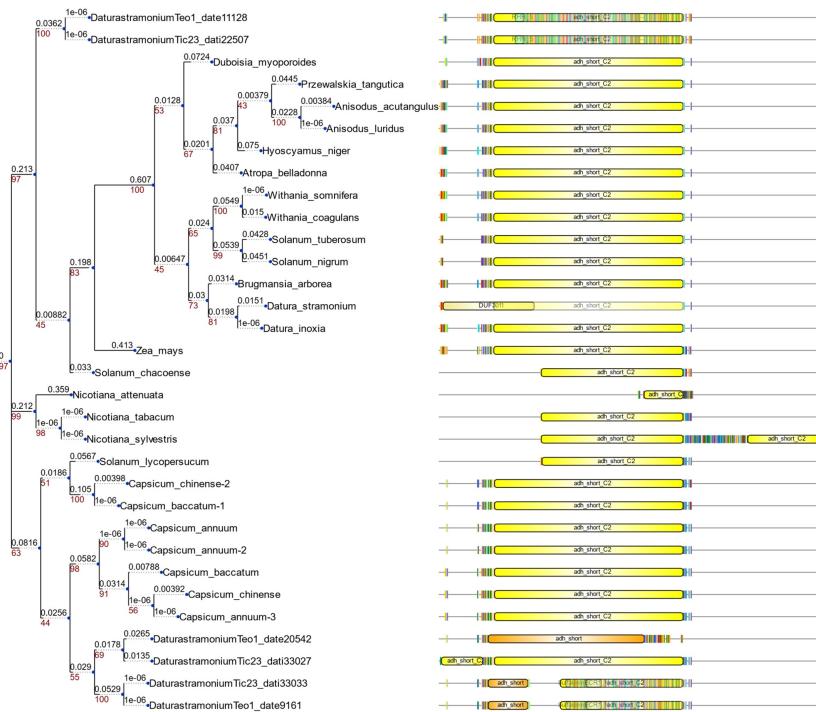
Figure 2 **Structure of domains of TRI.** Phylogenetic analysis and structure of domains of TRI gene in different species. Six copies are distributed in different clades. Tic23dati33027 has an extra domain adh_short_C.

(i.e., Teo19488) has only one copy (Fig. 3). The alignment of 29 sequences of H6H of different plant species indicate the variation in the duplicated DIOX-N dominion in Tic8550 (Fig. 3B).
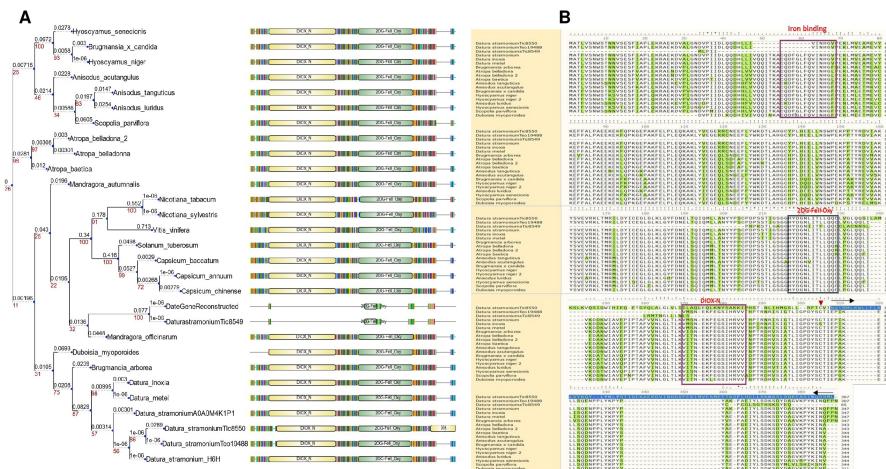
Figure 3 Phylogeny of gene H6H. Phylogeny of gene H6H in 19 species. In (A) we observe two copies of the gene are present in the two sequenced genomes of *D. stramonium* (Tic and Teo). In Tic8550 there is a tandem duplication of dominion DIOX N (PF14226) vs. Teo 19488; two copies of the gene are present in both genomes (segmental duplication) distributed in different clades. (B) Alignment of sequences of protein H6H with the conserved zones highlighted. In the terminal carboxyl, the DIOX-N dominion is duplicated in TIC8550.

Full-size ⬚ DOI: 10.7717/peerj.11466/fig-3
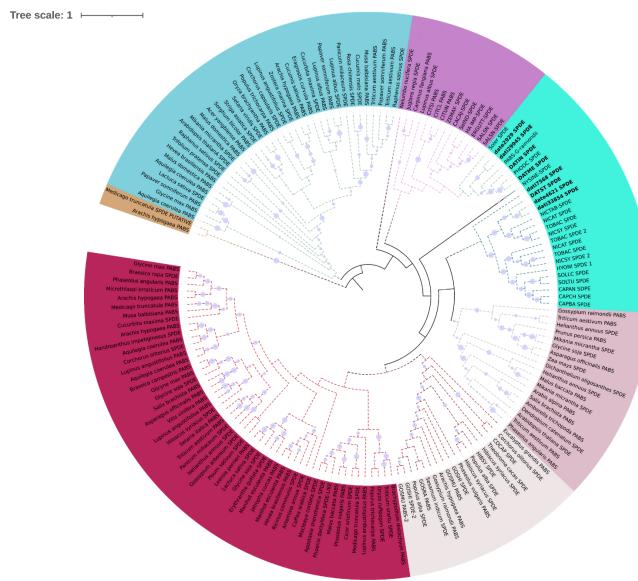
Figure 4 Phylogeny of pmt. Phylogeny of Putrescine N-methyltransferase (*pmt*). date = *Datura stramonium* Teotihuacán, dati = *Datura stramonium* Ticumán. The genomic analyses indicate an expansion of *pmt*, two copies in date and five in dati. The latter has an additional dominion of spermine synthase.

Full-size ⬚ DOI: 10.7717/peerj.11466/fig-4

In N-methyl putremescine transferase (DsPMT) we found variation in the number of gene copies in comparison to other Solanaceae (Fig. 4). In Tucumán's genome there are three copies while two copies were detected in Teotihuacán's genome (Fig. 4, Fig. S1).
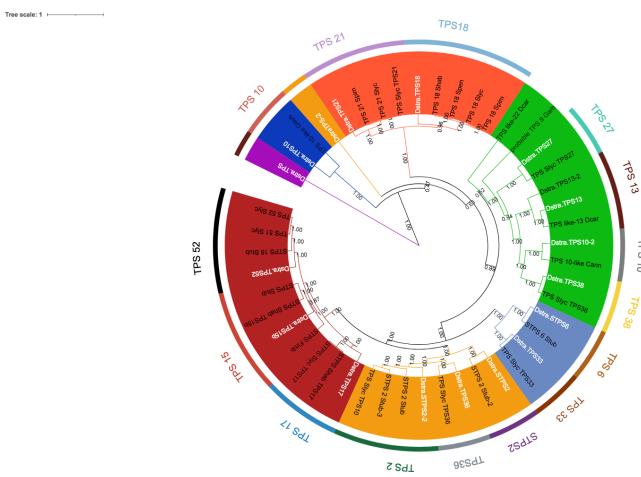
Figure 5 **Phylogeny of TPS.** 18 TPSs were found in *D. stramonium* (white letters) and Solanaceae and other angiosperms. Two copies of TPS-10 and TPS-13 in the subfamily TPSa. Tree obtained by Bayesian inference, with JTT+G+F evolutionary model. Most branches have a bootstrap support of 1.

Full-size 🖾 DOI: 10.7717/peerj.11466/fig-5

Terpene synthases

Eighteen TPSs genes were found in the genomes of *D. stramonium* (Fig. 5; Table S1) and seven in *Datura metel* (Fig. S2). The 18 *DsTPS* found in *D. stramonium* are distributed in four subfamilies, identified for other angiosperms: Nine in TPSa, 3 in TPSb, 2 in TPSc, 2 in TPSg and 2 TPS- unknown (Fig. S3, Tables S1 and S2) (*Huang et al., 2013*; *Huang et al., 2017*; *Falara et al., 2011*).

The domains of these TPSs, directly involved in the biosynthesis of terpenoids, show expansion events and positive selection in TPS10 (Fig. 6; Table S3). Changes in the dominions are: Terpene synthase, N-terminal domain (IPR001906), Terpenoid cyclases/protein prenyl transferase alpha-alpha toroid (IPR008930), Terpene synthase, metal-binding domain (IPR005630), Terpene cyclase-like 1, C-terminal (IPR034741). Analyses indicate that Solanaceae have the gene TPS10.1 whereas only *D. stramonium* and *D. metel* possess TPS10.2 (Fig. 7 and Tables S2 and S3); this possibly is a duplicate of TPS10.1_like.

Protein modeling and molecular docking

Superimposition of the reported protein structure of H6H from *D. metel* (PDB ID 6TTM) with the corresponding model obtained from *DsH6H* sequences for the two genomes of *D. stramonium* from two Mexican populations (Teo19488 and Tic8550), indicate similarities (Fig. 7A). The Teo19488 predicted structure match very well with the DmH6H structure (Fig. S2); similarly, Tic8550 matches well except for some residues that were not modeled (Fig. 7B, Fig. S3). This fragment of Tic8550 have a big terminal protein sequence (NQTMMLKWLLFLKLVYKQYLYKLYSTIGEKAEKDVINHGVPEKIMVEAMEFTKSFL-HCLLRKKKSLSQKEASIKQSSML) of 79 amino acids was not modeled (Fig. 7B, cf. Fig. 3B). Further, a NCBI protein-protein blast of this sequence aligns well to the region: 33–72 of DmH6H fragment; we analyzed this fragment separately as Tic8550_Fragment (Fig.
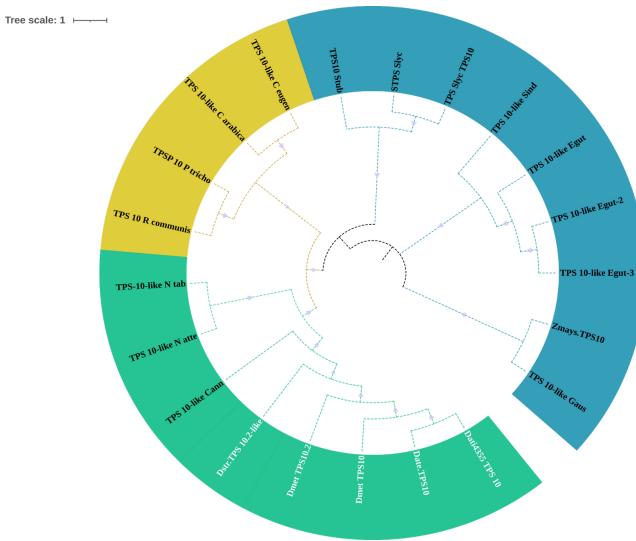
Figure 6 **Phylogeny of TPS-10.** Phylogeny of TPS-10 (green) of *Datura stramonium*, *D. metel* and other Solanaceae. In these species the TPS-10 is duplicated. Tree obtained by Bayesian inference, with JTT+G+F evolutionary model. Most branches have a support above 80%.

Full-size ☒ DOI: 10.7717/peerj.11466/fig-6
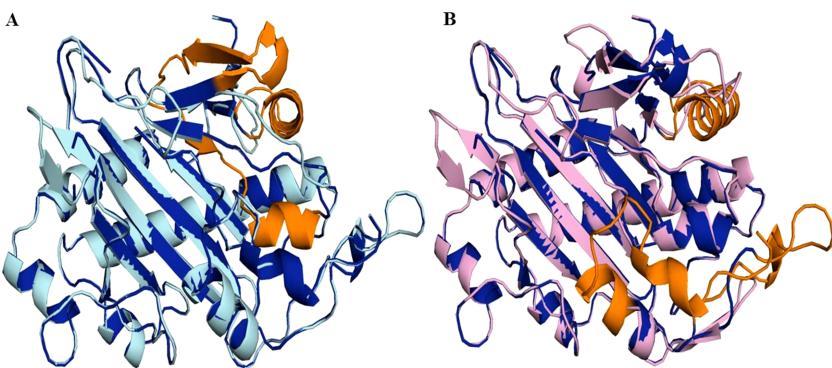
Figure 7 **Structure of 6TTM in *Datura metel*.** (A) Superimposition on the tertiary structure of H6H reported for *Datura metel* (PDB ID 6TTM) (in blue) of obtained models of the sequence for gene DsH6H *Datura stramonium* Teo19488 (Cyan) (A) and Tic8550 (Pink) (B). In orange color the folding present only in the reference structure 6TTM but not in *D. stramonium* are highlighted.

Full-size ☒ DOI: 10.7717/peerj.11466/fig-7

86). The Tic8549 sequence corresponds to a 65 amino acids sequence (MLPIIPRPKSTL-GAGGHYDGNIITFLQQDCLACNNSLLRMTNGLLLNLSYCFCGLSGTHSKGYEQ) that was modeled as Hy-oscyamine 6 beta-hydroxylase like fragment (Fig. 8). This fragment (Tic8549) aligns to DmH6H in one section of the binding pocket, where His-217 and Asp-219 are coordinated with Ni^{2+} (Fig. 8C). The presence of these amino acids in Tic8549 sequence is interesting since Ni^{2+} ion, a surrogate of the natively present Fe^{2+} ion, is coordinated by the side chains of His-217 (strand βII), Asp-219 (loop βII/ βIII) and His-274 (strand βVII) that forms a metal binding **His-X-Asp…His** motif, highly conserved in the oxoglutarate dependent oxygenases (ODD) family (*Kluza et al., 2020*).
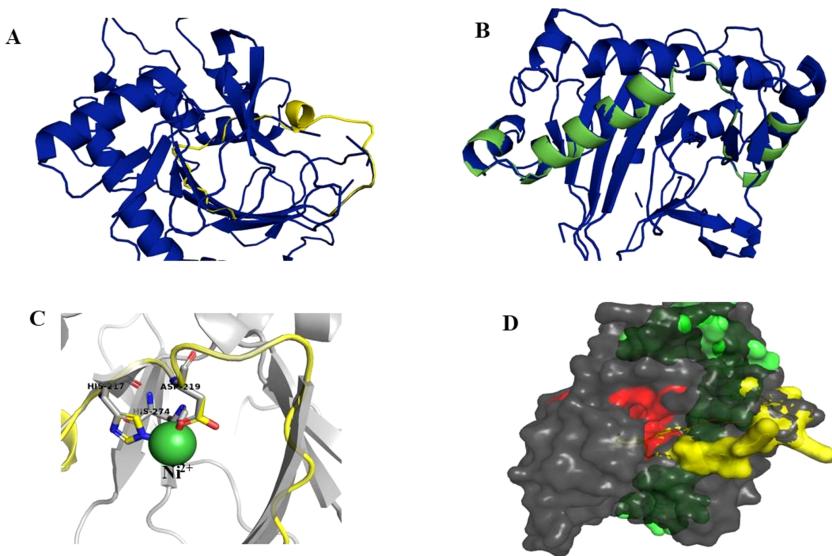
Figure 8 **Models of H6H.** Superimposition of reported structure of DmH6H (PDB ID 6TTM) (blue) and models obtained for the sequence Tic8549 (yellow) (A) and a fragment of the sequence Tic8550 (green) (B). Both sequences possess structural similarity with other sequences of the same protein. The fragment Tic8549 (yellow) aligns in the region where residues of the active site of H6H are located (gray) and involved in the coordination of the structure with metals (C). This site is highly conserved in the enzyme families ODD. (D) surface of fragments Tic8549 (Yellow) and Tic8550_Fragment (green) superimposed on H6H structure. In red are the residues that compose the active (binding) site.

Full-size 🖼 DOI: 10.7717/peerj.11466/fig-8

Protein models of H6H for *D. metel* and *D. stramonium* (Teo19488 and Tic8550) predict more interaction between residues in Tic8550 than on *D. metel* and Teo19488 (Fig. 9). It has recently been reported that the binding pocket (BP) of DmH6H is mainly formed by hydrophobic amino acids. It has been demonstrated that in the crystalized H6H the phenyl ring of Hy-o is bounded in an aromatic cage formed by Phe-103, Tyr-295, Tyr-319, Phe-322 and Tyr-326, of which the most prominent bound is with Tyr-326 which forms CH-π hydrogen bonds with the phenyl ring of Hy-o in an edge-to-face bidentate manner (*Kluza et al., 2020*) (Figs. 9A, 9D). Our results shown that the predicted model for Teo19488 have the same amino acids forming the aromatic cage composed by Phe-88, Tyr-255, Tyr-279, Phe-282 and Tyr-286 and the predicted pose of Hy-o obtained by the docking in silico shows that the phenyl ring is inside of this aromatic cage (Fig. 9B). Nevertheless, the main interaction with Tyr-286 was not detected (Fig. 9E). This could be because the cutoff distance was set 4 Å and the predicted pose of this residue is beyond this cutoff distance. The predicted structure of Tic8550 has different amino acids arrangement inside the binding pocket. This predicted structure does not contain the aromatic cage, instead it is predicted that the phenyl ring of Hy-o inside of a cavity formed by the non-polar amino acids Leu-308, Leu-92, Ile-179, Met-181, Met-292 and the polar uncharged amino acids Asn-206 and Asn-289 (Fig. 9C). Also, it is predicted that the phenyl ring has several non-polar interactions with these side chain amino acids (Fig. 9F). The Asn-206 forms a
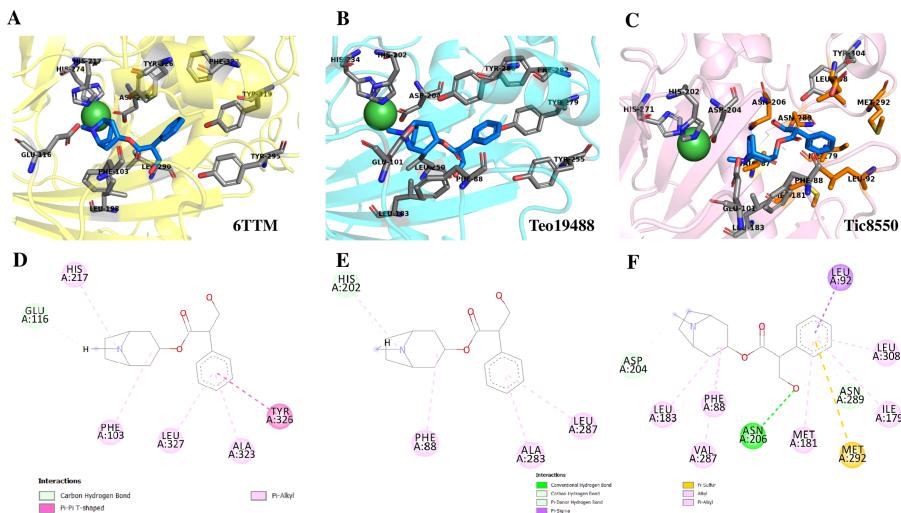
Figure 9 Structure of 6TTM. (A) Structure of 6TTM in *Datura metel* (yellow). The ligand (blue) inside the BP is surrounded by amino acids (gray). (B) Structure modelled for *Datura stramonium* Teo_19488 (cyan) and (C) *D. stramonium* Tic_8550. The interactions between the predicted residues of the BP and the ligand Hy-o are illustrated in D, E and F.

conventional hydrogen bond with the hydroxyl group from Hy-o which may stabilize the structure in the BP.

DISCUSSION

Gene duplicates product of molecular evolution are the raw material for evolutionary innovation (*Defoort, Van de Peer & Carretero-Paule, 2019*). Recent findings have uncovered the interactions and contrasting functions of genes that may help plants to confront environmental stresses in nature and serve as important agronomic characteristics. Hence, to trace the origin, molecular mechanisms, evolutionary fate and function of gene duplicates is a main goal (*Panchy, Lehti-Shiu & Shiu, 2016*; *Soltis & Soltis, 2016*; *Van de Peer, Mizrachi & Marchal, 2017*). The phylogenetic analyses of genes involved in the synthesis of tropane alkaloids (*DsPMT, DsTRI, DsTRII* and *DsH6H*) here presented, indicate that two genes (DsTRII and DsH6H) have expanded in *D. stramonium*, presenting five and two copies, respectively. Previous studies have shown that tropinone reductases are involved in the bifurcation of the synthesis pathways in Solanaceae (*Dräger et al., 1992*; *Nakajima, Hashimoto & Yamada, 1993*; *Birgit, 2006*; *Kanehisa & Sato, 2019*). These genes are responsive to environmental stimuli and activate the signaling pathways (*Campos, Kang & Howe, 2014*). The presence of these genes in the Solanaceae, as revealed by phylogenetic analyses, suggests that these duplicates in *D. stramonium* may contribute to increase the production of tropane alkaloids. The absence of TRII in *Zea mays* and *A. thaliana* suggest that these species never had this gene or lost it during their evolution. However, the gene is also absent in species of other clades (e.g., *Brassica rapa, Glycine max, Vitis vinifera,* etc.; Fig. 1). Another possibility is that TRII is not related to the synthesis of tropane alkaloids in these species (*Miller, Arteca & Pell, 1999*).

The gene DsPMT of *D. stramonium* found in the population of Ticumán, México, had an extra domain of spermine-synthase in comparison with its homologous gene found in the same species but from the population of Teotihuacán, México (*De-la-Cruz et al., 2021*). We named this gene as *Ds* PMT7568ti (Fig. 4), Spermine gene is a potent plant defense activator, with protective effects of broad-spectrum (*Seifi & Shelp, 2019*). Reports indicate that overexpression of spermidine synthase enhances tolerance to multiple environmental stresses, including the attack by herbivores and pathogens (*Kasukabe et al., 2004*; *Seifi et al., 2019*). In addition, PMT is a key enzyme in the catalysis of N-methylputrescine from putrescine and S-adenosyl-L-methionine and triggers the production of hygrine and other different tropane alkaloids (*Kanehisa & Sato, 2019*).

In contrast to other plants, in Solanaceae (*Atropa belladona*, *Hy-osciamus niger* and *Datura stramonium*), TRII is duplicated and under positive selection. In this line, we hypothesize that TRII may have acquired a different functional importance after its duplication in *D. stramonium* and hence, it has retained multiple copies of the gene. Yet, this hypothesis warrants further analysis. The phylogeny indicates an independent history, for instance, from Brassicaceae.

Perhaps TRs in Brassicales have other functions, different to the production of alkaloids, but using the same pathways. For instance, they may have conferred plasticity to deal with environmental stresses. Plant species with small genome size may have lost considerable fractions of their genome. Thus, enzymes with similar function may have been replaced by TRII.

Of the four subfamilies of TPSs found in the *D. stramonium*'s genome, three have been reported in other angiosperms (TPSa, TPSb and TPSg). TPS10 is member of the subfamily *a* and its role in defense against pathogens and herbivores has been dilucidated in *Zea mays* (*Huang et al., 2017*; *Singh & Sharma, 2015*; *Tholl, 2006*). We detected positive selection on TPS; positive selection on TPS10.1 and TPS10.2 was detected in *D. stramonium* and *D. metel.* We speculate that the presence of two copies of this gene can contribute to plant defense against plants' natural enemies (*Köllner, Gershenzon & Degenhardt, 2009*; *Köllner, Degenhardt & Gershenzon, 2020*).

In this line, duplicated genes involved in the specialized metabolism and plant defense often show differential restrictions (constraints, trade-offs, limitations) either indicating positive selection or differential selection (*Wang et al., 2016*). Apparently, TRII displays genic redundancy since it possesses different copies and polymorphic sites. However, this may not be the case since this process is only present in genes of primary metabolism or development (*Wang et al., 2016*; *Maltsev et al., 2005*) which are constitutive and display low variation. The gene H6H is involved in the last step in the production of scopolamine and atropine in one of the pathways for the production of tropane alkaloids (*Kanehisa & Sato, 2019*). Here, we detected one duplicated in tandem of this gene in *D. stramonium* (Ticumán has two domains of DIOX_N (PF14226), although each copy belongs to two different gene families (OG0028637 and OG0043057; cf. Fig. 3). Similarly, we found 18 TPSs in *D. stramonium* that likely are involved in plant defense against herbivores like in other plants species (e.g., Bharat & Sharma 2015). Lineages in the Solanaceae have high values of the ratio Ka/Ks (Table 2). Thus, positive selection may have acted on TPS10.1

Table 2 Selection test for tropane alkaloids (TAs) and terpene syntahses (TPSs) genes of *Datura stramonium*.

Genes	Chi-Square	Proportion of sites selected	ω
PMT	5.43432	0.1389	2.5778
TRI	0	NA	
TRII	31.5335	0.09994	10.6356
H6H	15.66146	0.15136	6.469
TPS10_1	22.75614	0.07163	1.704
TPs10_2	16.16614	0.014589	2.295
TPS14	4.67532	0.098812	3.7809
TPS21	1.43981	NA	
GGPS1	3.89631	0.0899	2.5988

and TPS10.2 of *D. stramonium*, not recently and the selection effect on allelic diversity may be masked by the accumulation of neutral mutation.

On the other hand, results of protein simulation and molecular docking suggest that the structure and interactions of this protein (H6H) in Tic8550 might have a different structure to that of the known H6H; we speculate if this may be related to enhancing the efficiency of alkaloids' synthesis.

It is interesting to note that the Tic8549 possess only the fragment 2OG-FeII_Oxy (PF03171) whereas Tic8550 possess an extra dominion (DIOX_N) which aligns structurally to a highly conserved region of H6H proteins (Figs. S2 and S3) (and associated to a coordination of binding metal Ni^{+2} (Fig. S5)). Further, it is interesting noting that a particular gene Tic8549 is associated to this fragment DIOX_N (PF14226).

Finally, the number of copies and differentiation of TAs and TPSs of *D. stramonium*, as compared to other solanaceous species, suggest their involvement on adaptive evolution of chemical defense against herbivores. Further studies should be directed to experimentally assess gain (overexpression) or loss (silencing) of function of duplicated genes.

CONCLUSION

Our results indicate differences in the number of gene copies involved in the synthesis of tropane alkaloids, between the genomes of *Datura stramonium* from the two Mexican populations (Ticumán and Teotihuacán). Furthermore, more copies of genes related to the synthesis of tropane alkaloids (TRI, TRII, H6H, PMT-10) are found in *D. stramonium* as compared to Viridiplantae. Likewise, for terpene synthases (TPS), TPS-10 is duplicated in *D. stramonium* and *D. metel*. These results point future experimental studies of gain/loss of duplicated genes.

Links to data on Figshare (Data S3).

ACKNOWLEDGEMENTS

ADDITIONAL INFORMATION AND DECLARATIONS

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Sabina Velázquez-Márquez conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored drafts of the paper, and approved the final draft.
- Ivan M. De-la-Cruz performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Rosalinda Tapia-López performed the experiments, prepared figures and/or tables, and approved the final draft.
- Juan Núñez-Farfán conceived and designed the experiments, performed the experiments, co-authored drafts of the paper, and approved the final draft.

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

Whole genome of two Mexican plants (Teo and Tic; reported at *Scientific Reports* https://doi.org/10.1038/s41598-020-79194-1) of *Datura stramonium* is available at GitHub: https://github.com/icruz1989/Datura-stramonium-genome-project.

The genome assemblies, Illumina and PacBio raw sequences from the two plants of *D. stramonium*, are available at DDBJ/ENA/GenBank, BioProject PRJNA622882: Teotihuacan assembly, acc. JAAWWX000000000, Ticumán assembly acc. JAAWWY000000000.

The Illumina and PacBio sequences for the Ticumán genome are available at SRR11474700 and SRR11474698, respectively. The Illumina and PacBio sequences for the Teotihuacán genome are available at: SRR11474701 and SRR11474699, respectively.

The two transcriptomes of *Datura stramonium* (Datura_stramonium_1.tgz and Datura_stramonium_2.tgz) and one of *Datura metel* (Datura_metel.tgz) are available at: https://medplantrnaseq.org/.

Data Availability

The following information was supplied regarding data availability:

The TPS proteins of *Datura stramonium* and D. metel sequence are available in the Supplemental Files.

The data is available at:

Velázquez, Sabina; Fárfan, Juan Núñez (2021): TPS gene analysis. figshare. Dataset. https://doi.org/10.6084/m9.figshare.14343356.v3

Velázquez, Sabina; Fárfan, Juan Núñez (2021): TPS genome blast. figshare. Dataset. https://doi.org/10.6084/m9.figshare.14343347.v2

Velázquez, Sabina; Fárfan, Juan Núñez (2021): Augustus analysis. figshare. Dataset. https://doi.org/10.6084/m9.figshare.14343320.v2

Velázquez, Sabina; Fárfan, Juan Núñez; De la Cruz, Ivan; Tapia-López, Rosalinda (2021): *Datura metel* transcriptome analysis. figshare. Dataset. https://doi.org/10.6084/m9.figshare.14343317.v2

Velázquez, Sabina; Fárfan, Juan Núñez (2021): Transcriptome analysis D. stramonium and D. metel. figshare. Dataset. https://doi.org/10.6084/m9.figshare.14343302.v2

Velázquez, Sabina; Fárfan, Juan Núñez (2021): Transcriptome assembly D. metel. figshare. Dataset. https://doi.org/10.6084/m9.figshare.14343296.v2

Velázquez, Sabina; Fárfan, Juan Núñez (2021): TPS Genes and Tropane Alkaloids. figshare. Dataset. https://doi.org/10.6084/m9.figshare.14343284.v2

Velázquez, Sabina; Fárfan, Juan Núñez (2021): TPS10. figshare. Dataset. https://doi.org/10.6084/m9.figshare.14343278.v1

Velázquez, Sabina; Fárfan, Juan Núñez (2021): Tropinone reductase. figshare. Dataset. https://doi.org/10.6084/m9.figshare.14343272.v3

Velázquez, Sabina; Fárfan, Juan Núñez (2021): PMT Tropane alkaloid. figshare. Dataset. https://doi.org/10.6084/m9.figshare.14343269.v4

Velázquez, Sabina; Fárfan, Juan Núñez (2021): H6H Tropane alkaloid. figshare. Dataset. https://doi.org/10.6084/m9.figshare.14343263.v8.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.11466#supplemental-information.

REFERENCES

Bienert S, Waterhouse A, De Beer TAP, Tauriello G, Studer G, Bordoli L, Schwede T. 2017. The SWISS-MODEL repository—new features and functionality. *Nucleic Acids Research* **45(D1)**:D313–D319 DOI 10.1093/nar/gkw1132.

Birgit D. 2006. Tropinone reductases, enzymes at the branch point of tropane alkaloid metabolism. *Phytochemistry* **67(4)**:327–337 DOI 10.1016/j.phytochem.2005.12.001.

Bolger A, Scossa F, Bolger ME, Lanz C, Maumus F, Tohge T, Quesneville H, Saleh Alseekh S, Sørensen I, Lichtenstein G, Fich EA, Conte M, Keller H, Schneeberger K, Schwacke R, Ofner I, Vrebalov J, Xu Y, Osorio S, Alves Aflitos S, Schijlen E, Jiménez-Goméz JM, Ryngajllo M, Kimura S, Kumar R, Koenig D, Headland LR, Maloof JN, Sinha N, Van Ham RCH, Klein Lankhorst RM, Mao L, Vogel A, Arsova B, Panstruga R, Fei Z, Rose JKC, Zamir D, Carrari F, Giovannoni JJ, Weigel D, Usadel B, Fernie AR. 2014. The genome of the stress-tolerant wild tomato species *Solanum pennellii*. *Nature Genetics* **46**:1034–1038 DOI 10.1038/ng.3046.

Brille F, Loreto F, Baccelli I. 2019. Exploiting plant Volatile Organic Compounds (VOCs) in agriculture to improve sustainable defense strategies and productivity of crops. *Frontiers in Plant Science* **10**:264 DOI 10.3389/fpls.2019.00264.

Campos ML, Kang JH, Howe GA. 2014. Jasmonate-triggered plant immunity. *Journal of Chemical Ecology* **40(7)**:657–675 DOI 10.1007/s10886-014-0468-3.

Castillo G, Cruz LL, Tapia-López R, Olmedo-Vicente E, Carmona D, Anaya-Lang AL, Fornoni J, Andraca-Gómez G, Valverde PL, Núñez Farfán J. 2014. Selection mosaic exerted by specialist and generalist herbivores on chemical and physical defense of *Datura stramonium*. *PLOS ONE* **9(7)**:e102478 DOI 10.1371/journal.pone.0102478.

Chen F, Mackey AJ, Stoeckert CJ, Roos D. 2006. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Research* **34**:363–368 DOI 10.1093/nar/gkj123.

Chen Q, Wan Y, Lei Y, Zobel J, Verspoor K. 2016. Evaluation of CD-HIT for constructing non-redundant databases. In: *IEEE international conference on bioinformatics and biomedicine (BIBM)*. 703–706 DOI 10.1109/BIBM.2016.7822604.

Conan GC, Wolfe KH. 2008. Turning a hobby into a job: how duplicated genes find new functions. *Nature Reviews Genetics* **9**:938–950 DOI 10.1038/nrg2482.

De-la-Cruz IM, Hallab A, Olivares-Pinto U, Tapia-López R, Velázquez-Márquez S, Piñero D, Oyama K, Usadel B, Núñez Farfán J. 2021. Genomic signatures of the evolution of defense against its natural enemies in the poisonous and medicinal plant *Datura stramonium* (Solanaceae). *Scientific Reports* **11(1)**:1–19 DOI 10.1038/s41598-020-79194-1.

De-la-Cruz IM, Velázquez-Márquez S, Núñez Farfán J. 2020. What do we know about the genetic basis of plant defensive responses to herbivores? A minireview. In: Núñez-Farfán J, Valverde P, eds. *Evolutionary ecology of plant-herbivore interaction*. Springer DOI 10.1007/978-3-030-46012-9_16.

Defoort J, Van de Peer Y, Carretero-Paule L. 2019. The Evolution of gene duplicates in angiosperms and the impact of protein–protein interactions and the mechanism of duplication. *Genome Biology and Evolution* **11(4)**:2292–2305 DOI 10.1093/gbe/evz156.

Dräger B, Portsteffen A, Schaal A, McCabe P, Peerless A. 1992. Levels of tropinone-reductase activities influence the spectrum of tropane esters found in transformed root cultures of *Datura stramonium (L.)*. *Planta* **188(4)**:81–586 DOI 10.1007/BF00197052.

Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W. 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* **161**:1307–1132 DOI 10.1093/genetics/161.3.1307.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32(5)**:1792–1797 DOI 10.1093/nar/gkh340.

Edwards KD, Fernandez-Pozo N, Drake-Stowe K, Humphry M, Evans AD, Bombarely A, Allen F, Hurts R, White B, Kernodle SP, Bromley J, Sanchez-Tamburrino J, Lewis R, Mueller L. 2017. A reference genome for *Nicotiana tabacum* enables map-based cloning of homologous loci implicated in nitrogen utilization efficiency. *BMC Genomics* **8**:448 DOI 10.1186/s12864-017-3791-6.

Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology* **16**:157 DOI 10.1186/s13059-015-0721-2.

Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology* **20**:238 DOI 10.1186/s13059-019-1832-y.

Falara V, Akhtar TA, Thuong TH, Spyropoulou EA, Bleeker PM, Schauvinhold I, Matsuba Y, Bonini ME, Schilmiller AL, Last RL, Schuurink RC, Pichersky E. 2011. The tomato terpene synthase gene family. *Plant Physiology 1* **57**:770–789 DOI 10.1104/pp.111.179648.

Hagel JM, Facchini PJ. 2010. Dioxygenases catalyze the O-demethylation steps of morphine biosynthesis in opium poppy. *Nature Chemical Biology* **6**:273 DOI 10.1038/nchembio.317.

Heil M, Lion U, Boland W. 2008. Defense-inducing volatiles: in search of the active motif. *Journal of Chemical Ecology* **34(5)**:601–604 DOI 10.1007/s10886-008-9464-9.

Huang AC, Kautsar SA, Hong YJ, Medema MH, Bond AD, Tantillo DJ, Osbourn A. 2017. Unearthing a sesterterpene biosynthetic repertoire in the Brassicaceae through genome mining reveals convergent evolution. *Proceedings of the National Academy of Sciences of the United States of America* **114**:E6005–E6014 DOI 10.1073/pnas.1705567114.

Huang X, Xiao Y, Köllner TG, Zhang W, Wu J, Wu J, Zhang Y. 2013. Identification and characterization of (E)-β-caryophyllene synthase and α/β-pinene synthase potentially involved in constitutive and herbivore-induced terpene formation in cotton. *Plant Physiology and Biochemistry* **73**:302–308 DOI 10.1016/j.plaphy.2013.10.017.

Kanehisa M, Sato Y. 2019. KEGG Mapper for inferring cellular functions from protein sequences. *Protein Science* **29**:28–35 DOI 10.1002/pro.3711.

Kasukabe Y, He L, Nada K, Misawa S, Ihara I, Tachibana S. 2004. Overexpression of spermidine synthase enhances tolerance to multiple environmental stresses and up-regulates the expression of various stress-regulated genes in transgenic *Arabidopsis thaliana*. *Plant & Cell Physiology* **45**:712–722 DOI 10.1093/pcp/pch083.

Kessler A, Baldwin I. 2002. Plant responses to insect herbivory: the emerging molecular analysis. *Annual Review of Plant Biology* **53**:99–328 DOI 10.1126/science.291.5511.2141.

Kim S, Park M, Yeom SI, Kim M, Lee JM, Lee HA, Choi D. 2014. Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. *Nature genetics* **46(3)**:270–278 DOI 10.1038/ng.2877.

Kluza A, Wojdyla Z, Mrugala B, Kurpiewska K, Porebski PJ, Niedzialkowska E, Borowski T. 2020. Regioselectivity of Hy-oscyamine 6 β-hydroxylase-catalysed hydroxylation as revealed by high-resolution structural information and QM/MM calculations. *Dalton Transactions* **49**:4454–4469 DOI 10.1039/d0dt00302f.

Koerner SE, Smith Md, Burkepile DE, Hanan NP, Avolio ML, Collins SL, Knapp AK, Lemoine NP, Forrestel EJ, Eby S, Thompson DI, Aguado-Santacruz G, Anderson JP, Anderson TM, Angassa A, Bagchi S, Bakker ES, Bastin G, Baur LE, Beard KH, Beever EA, Bohlen PJ, Boughton EH, Canestro D, Cesa A, Chaneton E, Cheng J, D'Antonio CM, Deleglise C, Dembélé F, Dorrough J, Eldridge DJ, Fernandez-Going S, Fern[a]ndez-Lugo S, Fraser LH, Freedman B, García-Salgado G, Goheen , Liang Guo JR, Husheer S, Karembé M, Knops MH, Kraaij T, Kulmatiski A, Minna-Maarit K, Lezama F, Loucougaray G, Loydi A, Milchunas DG, Milton SJ, Morgan JW, Moxham C, Nehring KC, Olff H, Palmer TM, Rebollo S, Riginos C, Risch AC, Rueda M, Sankaran M, Sasaki T, Schoenecker KA, Schultz NL, Schütz M, Schwabe A, Siebert F, Smit K, Stahlheber KA, Storm C, Strong DJ, Su J, Tiruvaimozhi YV, Tyler C, Val J, Vandegehuchte ML, Veblen KE, Vermeire LT, Ward D, Wu J, Young TP, Yu Q, Zelikova TJ. 2018. Change in dominance determines herbivore effects on plant biodiversity. *Nature Ecology and Evolution* **2**:1925–1932 DOI 10.1038/s41559-018-0696-y.

Köllner TG, Degenhardt J, Gershenzon J. 2020. The product specificities of maize terpene synthases TPS4 and TPS10 are determined both by active site amino acids and residues adjacent to the active site. *Plants* **9(5)**:552 DOI 10.3390/plants9050552.

Köllner TG, Gershenzon J, Degenhardt J. 2009. Molecular and biochemical evolution of maize terpene synthase 10, an enzyme of indirect defense. *Phytochemistry* **70**:1139–1145 DOI 10.1016/j.phytochem.2009.06.011.

Karunanithi PS, Zerbe P. 2019. Terpene synthases as metabolic gatekeepers in the evolution of plant terpenoid chemical diversity. *Frontiers in Plant Science* **10**:1166 DOI 10.3389/fpls.2019.01166.

Lamalakshmi E, Kumar S, Basanta T, Susheel S, Beemrote A, Chingakham P, Chongtham SK, Chongtham HS, Yumlembam R, Haribhushan A, Prakash N, Wani SH. 2017. Adaptation strategies and defense mechanisms of plants during environmental stress. In: Ghorbanpour M, Varma A, eds. *Medicinal plants and environmental challenges.* Springer, 359–413 DOI 10.1007/978-3-319-68717-9_20.

Letunic I, Bork P. 2019. Interactive Tree of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research* **47**:256–259 DOI 10.1093/nar/gkz239.

Lynch M, Conery JS. 2003. The origins of genome complexity. *Science* **5649(302)**:1401–1409 DOI 10.1126/science.1089370.

Maltsev N, Glass EM, Ovchinnikova G, Gu Z. 2005. Molecular mechanisms involved in robustness of yeast central metabolism against null mutation. *Journal of Biochemistry* **34(5)**:601–604 DOI 10.1093/jb/mvi017.

Miller JD, Arteca RN, Pell EJ. 1999. Senescence-associated gene expression during ozone-induced leaf senescence in Arabidopsis. *Plant Physiology* **120**:1015–1024 DOI 10.1104/pp.120.4.1015.

Miranda-Pérez A, Castillo G, Hernández-Cumplido J, Valverde PL, Borbolla M, Cruz LL, Tapia-López R, Fornoni J, Flores-Ortiz CM, Núñez Farfán J. 2016. Natural selection drives chemical resistance of *Datura stramonium*. *PeerJ* **4**:e1898 DOI 10.7717/peerj.1898.

Mithöfer A, Boland W. 2008. Plant defense against herbivores: chemical aspects. *Annual Review of Plant Biology* **63**:431–450 DOI 10.1146/annurev-arplant-042110-103854.

Nakajima K, Hashimoto T, Yamada Y. 1993. Two tropinone reductases with different stereo specificities are short-chain dehydrogenases evolved from a common ancestor. *Proceedings of the National Academic of Science of the United States of America* **90**:9591–9595 DOI 10.1073/pnas.90.20.9591.

Ohno S. 1970. *Evolution by gene duplication.* New York: Springer-Verlag Press.

Panchy N, Lehti-Shiu M, Shiu SH. 2016. Evolution of gene duplication in plants. *Plant Physiology* **171(4)**:2294–2316 DOI 10.1104/pp.16.0052.

Picazo-Aragonés J, Terrab A, Balao F. 2020. Plant volatile organic compounds evolution transcriptional regulation, epigenetics and polyploidy. *International Journal of Molecular Sciences* **21(23)**:8956 DOI 10.3390/ijms21238956.

Pigatto AS, Blanco CC, Mentz LA, Soares GL. 2015. Tropane alkaloids and calystegines as chemotaxonomic markers in the Solanaceae. *Anais de Academia Brasileira de Ciencias* **87(4)**:2139–2149 DOI 10.1590/0001-3765201520140231.

Razali R, Bougouffa S, Morton M, Lightfoot DJ, Alam I, Essack M, Arold S, Kamau A, Schmöckel S, Pailles Y, Shaid M, Michel C, Al-Babili S, Shwen Y, Tester M, Bajic , Negrap S. 2018. The genome sequence of the wild Tomato *Solanum pimpinellifolium* provides insights into salinity tolerance. *Frontiers in Plant Science* **9**:1402 DOI 10.3389/fpls.2018.01402.

Rozewick J, Songling L, Amada KM, Standley DM, Kazutaka K. 2019. MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Research* **47**:W5–W10 DOI 10.1093/nar/gkz342.

Schrödinger Release 2021-1. 2021. Desmond molecular dynamics system. D.E. Shaw Research, New York, NY, 2021. In: *Maestro-Desmond interoperability tools.* Schrödinger, New York, NY.

Seifi HS, Shelp BJ. 2019. Spermine differentially refines plant defense responses against biotic and abiotic stresses. *Frontiers in Plant Science* **10**:117 DOI 10.3389/fpls.2019.00117.

Seifi HS, Zarei A, Hsiang T, Shelp BJ. 2019. Spermine is a potent plant defense activator against gray mold disease on *Solanum lycopersicum*, Phaseolus vulgaris, and *Arabidopsis thaliana*. *Phytopathology* **109(8)**:1367–1377 DOI 10.1094/PHYTO-12-18-0470-R.

Shonle I, Bergelson J. 2000. Evolutionary ecology of the tropane alkaloids of *Datura stramonium L. (Solanaceae)*. *Evolution* **54(3)**:778–788 DOI 10.1111/j.0014-3820.2000.tb00079.x.

Sierro N, Battey JN, Quadi S, Boved L, Goeofert S, Bakaher N, Ivanov N. 2013. Reference genomes and transcriptomes of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*. *Genome Biology* **14**:R60 DOI 10.1186/gb-2013-14-6-r60.

Singh B, Sharma RA. 2015. Plant terpenes: defense responses, phylogenetic analysis, regulation and clinical applications. *Biotechnology* **5(2)**:129–151 DOI 10.1007/s13205-014-0220-2.

Soltis PS, Soltis DE. 2016. Ancient WGD events as drivers of key innovations in angiosperms. *Current Opinion in Plant Biology* **30**:159–165 DOI 10.1016/j.pbi.2016.03.015.

Stanke M, Keller O, Gunduz I, Hayes A, Waak S. Morgenstern B. 2006. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Research* **34(2)**:W435:439 DOI 10.1093/nar/gkl200.

The Potato Genome Sequencing Consortium. 2011. Genome sequence and analysis of the tuber crop potato. *Nature* **475**:189–195 DOI 10.1038/nature10158.

The Tomato Genome Consortium. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **485**:635–641 DOI 10.1038/nature11119.

Tholl D. 2006. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Current Opinion in Plant Biology* **9(3)**:297–304 DOI 10.1016/j.pbi.2006.03.014.

Van de Peer Y, Mizrachi E, Marchal K. 2017. The evolutionary significance of polyploidy. *Nature Review Genetics* **18(7)**:411–424 DOI 10.1038/nrg.2017.26.

Wallace KB, Zhang Y. 2020. Virtual Screening of Human Class-A GPCRs Using Ligand Profiles Built on Multiple Ligand-Receptor Interactions. *Journal of Molecular Biology* **432(17)**:4875–4890 DOI 10.1016/j.jmb.2020.07.003.

Wallace IM, Sullivan O, Higgins DG, Notredame C. 2006. M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Research* **34(6)**:1692–1699 DOI 10.1093/nar/gkl091.

Waterhouse A, Bertoni MB, Studer S, Tauriello G, Gumienny R, Heer FT, Beer TAde, Rempfer C, Bordoli L, Lepore R, Schwede T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research* **46(W1)**:W296–W303 DOI 10.1093/nar/gky427.

Wang L, Bei X, Gao J, Lin Y, Yan Y, Hu Y. 2016. The similar and different evolutionary trends of MATE family occurred between rice and Arabidopsis thaliana. *BMC plant biology* **16(1)**:1–19 DOI 10.1186/s12870-016-0895-0.

Wink M. 2003. Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry* **64**:3–19 DOI 10.1016/s0031-9422(03)00300-5.

Xu S, Brockmöller T, Navarro-Quezada A, Kuhl H, Gase K, Ling Z, Zhou W, Kreitzer C, Stanke M, Tang H, Lyons E, Pandey P, Pandey SP, Timmermann B, Gaquerel E, Baldwin IT. 2017. Wild tobacco genomes reveal the evolution of nicotine biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **114**:6133LP–LP6138 DOI 10.1073/pnas.1700073114.

Zhou F, Pichersky E. 2020. The complete functional characterization of the terpene synthase family in tomato. *New Phytologist* **226(5)**:1341–1360 DOI 10.1111/nph.16431.