

RESEARCH

Open Access



Multiple variation patterns of terpene synthases in 26 maize genomes

Yang Sun^{1,2*}, Wenqing Xiao¹, Qing-nan Wang^{1,2,3}, Jing Wang^{1,2}, Xiang-dong Kong⁴, Wen-hui Ma¹, Si-xian Liu¹, Ping Ren¹, Li-na Xu^{3*} and Yong-Jun Zhang^{2*}

Abstract

Terpenoids are important compounds associated with the pest and herbivore resistance mechanisms of plants; consequently, it is essential to identify and explore terpene synthase (*TPS*) genes in maize. In the present study, we identified 31 *TPS* genes based on a pan-genome of 26 high-quality maize genomes containing 20 core genes (present in all 26 lines), seven dispensable genes (present in 2 to 23 lines), three near-core genes (present in 24 to 25 lines), and one private gene (present in only 1 line). Evaluation of k_a/k_s values of *TPS* in 26 varieties revealed that *TPS25* was subjected to positive selection in some varieties. Six *ZmTPS* had k_a/k_s values less than 1, indicating that they were subjected to purifying selection. In 26 genomes, significant differences were observed in *ZmTPS25* expression between genes affected by structural variation (SV) and those not affected by SV. In some varieties, SV altered the conserved structural domains resulting in a considerable number of atypical genes. The analysis of RNA-seq data of maize *Ostrinia furnacalis* feeding revealed 10 differentially expressed *ZmTPS*, 9 of which were core genes. However, many atypical genes for these responsive genes were identified in several genomes. These findings provide a novel resource for functional studies of *ZmTPS*.

Keywords Terpene synthases, Pangenome, Positive selection, Structural variation, Atypical, *Ostrinia furnacalis*

Introduction

Terpenoids are a class of secondary metabolites produced by plants. Terpenoids are vital to photosynthesis, regulation of plant growth and development, pollination, and plant resistance to biotic and abiotic stressors [1]. For example, carotenoids such as tetraterpenoids can absorb and transmit light energy, and certain essential plant hormones (such as abscisic acid, oleuropein lactone, and gibberellin) are terpene derivatives that can influence plant growth and development. Linalool and nerolidol produced by tea under low temperatures can improve cold tolerance [2]. Linalool [3], limonene [4], nerolidol [5], and γ -terpinene [6] can inhibit the activity of rice leaf blight fungus. In addition, terpenoids are involved in induced defence responses in plants; for example, (E)- β -farnesene repels aphids [7, 8], while sphagnum and linalool attract rice louse tassel wasps [9].

*Correspondence:

Yang Sun
2018259@ahnu.edu.cn
Li-na Xu
caasxln@163.com
Yong-Jun Zhang
yjzhang@ippcaas.cn

¹ Key Laboratory for Conservation and Use of Important Biological Resources of Anhui Province, Anhui Provincial Key Laboratory of Molecular Enzymology and Mechanism of Major Diseases, College of Life Sciences, Anhui Normal University, Wuhu 241000, Anhui, China

² State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, China

³ Institute of Plant Protection and Agro-products Safety, Anhui Academy of Agricultural Sciences, Hefei 230031, China

⁴ Institute of Bioinformatics, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, China



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Terpene synthase (TPS) catalyses the production of terpenoids, whose abundance is strongly linked to a large number of terpene synthase genes in plants [10]. With the exception of the moss *Physcomitrella patens*, which has only one functional *TPS* gene, the *TPS* gene family in plants is medium-sized. Other plant genomes contain approximately 20–152 *TPS* genes, although some have lost their functions over the course of their evolution [11].

Based on sequence characteristics, the *TPS* gene family is classified into three major classes and seven subfamilies, namely *TPS-a*, *-b*, *-c*, *-d*, *-e/f*, *-g*, and *-h*. Among them, genes in *TPS-a* are sesquiterpene synthases. *TPS-b* encodes monoterpene synthases exclusive to angiosperms, with a conserved R(R)X8W motif that functions in the initiation phase of the isomerisation cyclisation reaction or stabilises the protein through electrostatic interactions. The *TPS-g* subfamily is closely related to *TPS-b*, but its encoded protein lacks the conserved R(R)X8W motif and may produce acyclic, monocyclic, sesquiterpene, and diterpene products. *TPS-c* is found in terrestrial plants and is characterised by the ‘DxDD’ motif. *TPS-d* is gymnosperm-specific gene that encodes monoterpene, sesquiterpene, and diterpene synthases. The *TPS-e/f* subfamily is found mostly in vascular plants, encodes copalyl diphosphate synthase (CPS) and kaurene synthase (KS), and is responsible for gibberellin biosynthesis [10, 12, 13]. *TPS-h* is only present in the *selaginella tamariscina* which encode ‘DXDD’ and ‘DDXXD’ motifs [11].

The study of *ZmTPS* genes is crucial as terpenoids play vital roles in pest and herbivore resistance; however, the *TPS* gene family in maize has not yet been identified and analysed. Traditional gene family identification approaches generally involve a genome-wide search for conserved structural domains of the gene family, followed by the identification of the gene family members after confirmation of the conserved structural domains. However, a single reference genome-based approach to gene family identification cannot identify the gene family members that are missing from the reference genome but are present in other genomes. Hufford et al. published a maize pan-genome based on 26 high-quality genomes containing a large amount of presence-absence variation (PAV) and SV information, thus laying the foundation for gene family and functional studies [14].

The present study identified 31 *TPS* genes based on a pan-genome of 26 high-quality maize genomes, which included 20 core genes, seven dispensable genes, three near-core genes, and one private gene. The analysis of ka/ks values of *TPS* in 26 varieties revealed the effect of SV on gene structure, expression, and conserved structural domains. In some varieties, SVs impair conserved

ZmTPS structural domains resulting in a significant number of atypical genes. Ten *ZmTPSs* showed differential expression when the RNA-seq data from *Ostrinia furnacalis*-infested maize samples were analysed. These results provide a novel resource for functional studies of *ZmTPS* genes.

Materials and methods

Identification of maize *TPS* gene family

The 26 maize genomes were obtained from a study by Hufford et al. [14]. The hidden Markov model (HMM) profiles of the *TPS* N-terminal domain (PF01397) and *TPS*-C-terminal domain (PF03936) were retrieved from the Pfam database (<http://pfam.xfam.org/>). *TPS* domains were searched using HMMER 3.3.2, with a threshold of $e < 1E-5$. The *TPS* candidates were submitted to SMART (<http://smart.embl-heidelberg.de/>) [15] to confirm the existence of *TPS* N and *TPS* C terminal domains. Finally, among the 26 maize accessions, genes with a collinear relationship to *TPSs* were considered *TPS* members.

Phylogenetic analysis and presence/absence variation of *ZmTPS* gene family

The protein sequences of the *TPSs* from *Arabidopsis* and maize were used for phylogenetic analysis. Multiple sequence alignments were performed using MAFFT v7.490 [16], and a phylogenetic tree was constructed using FastTree version 2.1.11 [17]. The final graphics were generated using iTOL v6 (<https://itol.embl.de/>) [18].

The PAV information of *TPSs* was obtained from Hufford et al. [14]. Heatmap was used to depict the presence/absence of each *TPS* in the 26 accessions using Rscript version 4.0.3 and the ComplexHeatmap package [19].

Ka/Ks calculation

The protein and coding sequence (CDS) sequences of *ZmTPS* genes in 26 maize genomes were obtained from the study by Hufford et al. [14]. *ZmTPS* sequences were compared, and Ka/Ks values were calculated using the KaKs Calculator [20]. The R packages ggridges and ggplot2 (v4.0.3) were used to create the Ridgeline plot of ka/ks values [21, 22]. The heatmap of the proportion of *ZmTPS* genes with Ka/Ks values greater than one was plotted using a ComplexHeatmap R (v4.0.3) [19].

Analysis of the expression of *TPSs* overlapped with SVs

The location of SV in each variety and the gene expression data in 26 accessions were obtained from Hufford et al. [14]. An in-house Perl script was used to determine whether the *TPSs* overlapped with SVs in each variety. If the *TPS* overlapped with SVs, the *TPS* expression data in this variety were considered as the expression data of genes with SVs; otherwise, they were

considered as the expression data of genes without SVs. The Pearson correlation coefficients were calculated between the presences of SVs overlapped with genes and gene expression level [23]. And *TPS*s with $p < 0.05$ and $|r| > 0.3$ were considered to have significantly altered expression levels due to SVs. Significant differences between atypical and typical *TPS* genes were determined by wilcox test.

Analysis of the gene structure of *ZmTPS*s

Genome annotation files in general feature format (GFF) were downloaded from <http://maize-pangenome.grame.ne.org>. MEME Suite v5.4.1 (<https://meme-suite.org/meme/tools/meme>) was used to analyse the protein sequences of *ZmTPS*, whose expression levels were significantly altered by SV, with the number of motifs set to 10. Subsequently, TBtools (v1.098761) was used to map gene structure using the files generated in the preceding step and the gff file [24].

Analysis of conserved structural domains of *TPS* under the influence of SV

The *ZmTPS* protein sequences of the most SV-overlapped genome (Ms71) and the reference genome (B73) were uploaded to MEME Suite v5.4.1 (<https://meme-suite.org/meme/tools/meme>), with the number of motifs set to 10, to obtain two weblogs of *TPS* protein sequences from the two genomes. The R package ComplexHeatmap was used to plot the presence of typical genes (containing completely conserved structural domains) and atypical genes (not containing completely conserved structural domains) for each *TPS* in the various maize genomes.

RNA-seq data analysis

The expression value (reads per kilobase of transcript, per million mapped reads; RPKM) and results of the differential expression analysis of control and maize samples fed upon by *Ostrinia furnacalis* were retrieved from the [supplementary tables](#) in the study by Guo et al. [25]. And other RNA-seq data about maize treated with oral secretions (OS) of *Mythimna separata* were also download [26, 27]. Low quality sequences were remove by fastp [28]. The clean reads were mapped to the reference genome by hisat 2 [29]. HTSeq was used to obtain the reads counts for each gene [30]. Differentially expressed genes were defined using a threshold of $|\log_2\text{FoldChange}| > 1$ and $\text{FDR} \leq 0.05$. Heatmaps were generated with \log_2 (RPKM + 1) values using R (v4.0.3) and the ComplexHeatmap package (v2.6.2) [19].

Result

Pan-genome-wide *ZmTPS* identification

In the present study, 32 *ZmTPS* genes were identified in the maize pan-genome, of which 20 were core genes, seven were dispensable genes, three were near-core genes, and one was a private gene. In the reference genome, 29 *TPS* were identified, three of which were atypical, and four *ZmTPS*s contained both typical and atypical genes. The [supplementary table](#) pan_gene_matrix_v3_cyverse.csv by Hufford et al. contained the sequence of the maize *ZmTPS* protein that was used to construct the evolutionary tree [14]. Based on the *OsTPS* categorisation information [31], we classified the *ZmTPS* (Fig. 1A) into four subgroups. *TPS-a* was the largest group, containing 18 *TPS* genes, followed by *TPS-c* and *TPS-ef*, each containing five *TPS* genes. *TPS-b* was the smallest group, with only three *TPS*. All the genes in *TPS-ef* were core genes. Figure 1B shows the presence or absence of *ZmTPS* genes other than the core genes in the 26 maize varieties. *TPS31* and *TPS28* were present in only one (CML333) and two (CML6 and Mo18W) varieties, respectively, suggesting that these two genes may be associated with traits specific to these varieties.

ZmTPS is subjected to different selection pressures among maize varieties

Analysis of K_a/K_s values can reveal the selection pressure on gene family members in different varieties forming. To explore the selection pressure on *ZmTPS* genes, we calculated the K_a/K_s value for each *TPS* gene based on the gene sequences in the 26 maize genomes (Fig. 2). Figure 2A depicts the K_a/K_s values of *TPS* in 26 maize cultivars. *ZmTPS28* was present in only two varieties and thus could only produce one K_a/K_s value, whereas *TPS31* was present in only one variety; thus, its K_a/K_s value could not be determined. All the K_a/K_s values of *TPS* genes peaked between 0 and 1, although the location of the peaks varied, with *ZmTPS11* having the highest K_a/K_s value. *TPS25* was positively selected for some varieties, as indicated by the K_a/K_s values, which ranged from 3 to 4. Six *ZmTPS* genes were found to have K_a/K_s values less than one, indicating that these genes were subjected to purifying selection. In the 26 maize genomes, *TPS25*, *TPS 27*, *TPS20*, *TPS19*, *TPS16*, *TPS2*, and *TPS5* had a large proportion of K_a/K_s values greater than one (Fig. 2B), indicating that these genes were under selection pressure during maize domestication.

Expression and structure of *ZmTPS25* genes are affected by SV

A total of 478 SVs overlapped with 30 gene regions of *ZmTPS* as well as upstream and downstream 2-kb

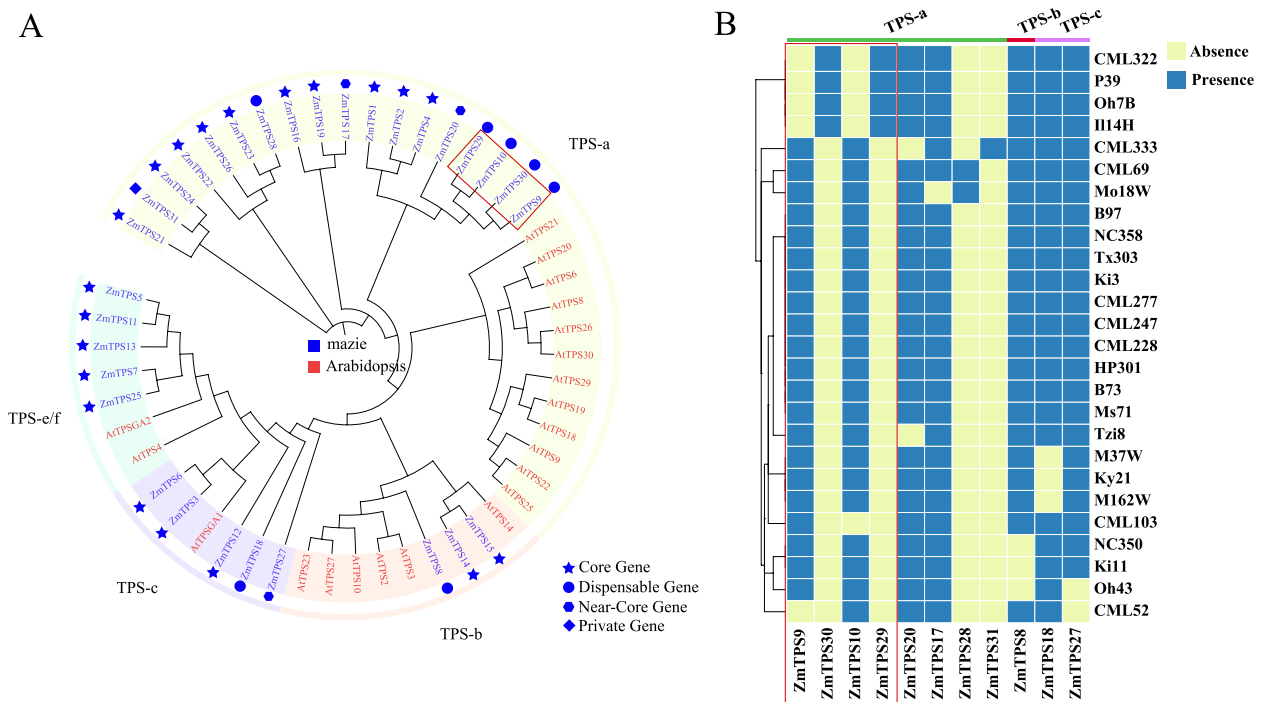


Fig. 1 **A** Phylogenetic tree of maize and Arabidopsis *TPS* genes. **B** Heatmap of the presence and absence of 11 *TPS*s in 26 maize varieties except for the core genes. The genes in red boxes represent possible mutually exclusive PAV genes

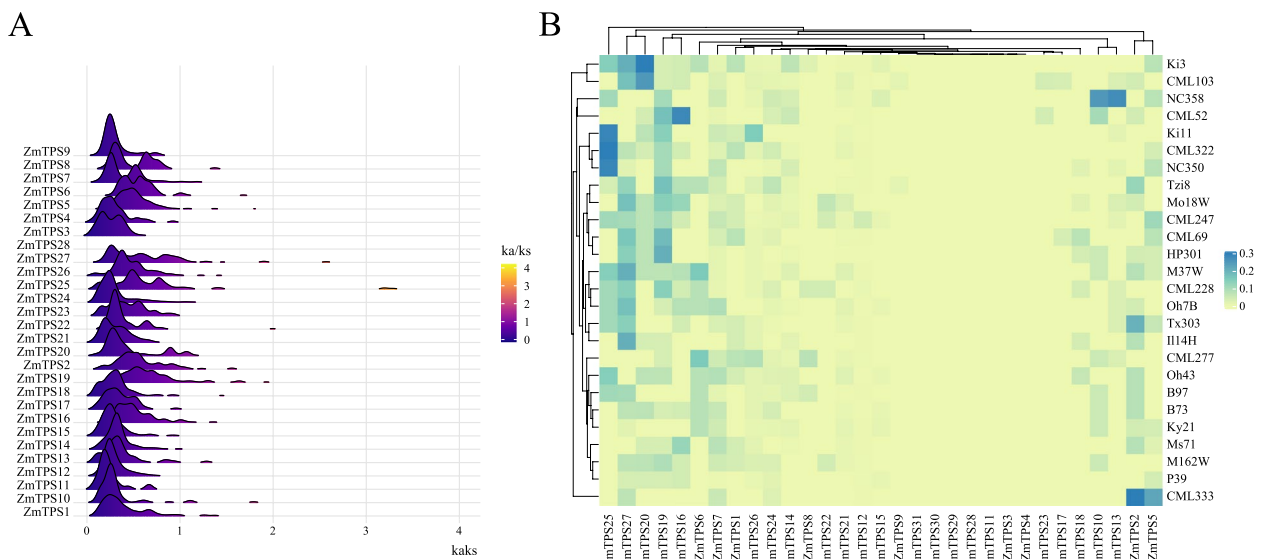


Fig. 2 **A** *Ka/Ks* values of *ZmTPS*. **A** Distribution of *Ka/Ks* values of *ZmTPS* in 26 maize varieties. **B** Heatmap of the frequency of occurrence of different maize varieties at each *TPS* with *Ka/Ks* ratio > 1

regions. Compared with the reference genome, SVs were characterised as deletions, insertions, inversions, and duplications (Fig. 3A). Pearson correlation coefficients were determined for the expression values of genes that overlapped with SV and genes that did not.

The results revealed a significant difference between genes with and without SV only for *ZmTPS25* ($p < 0.05$ and $|r| > 0.3$), indicating that SV significantly altered *ZmTPS25* expression (Fig. 3B).

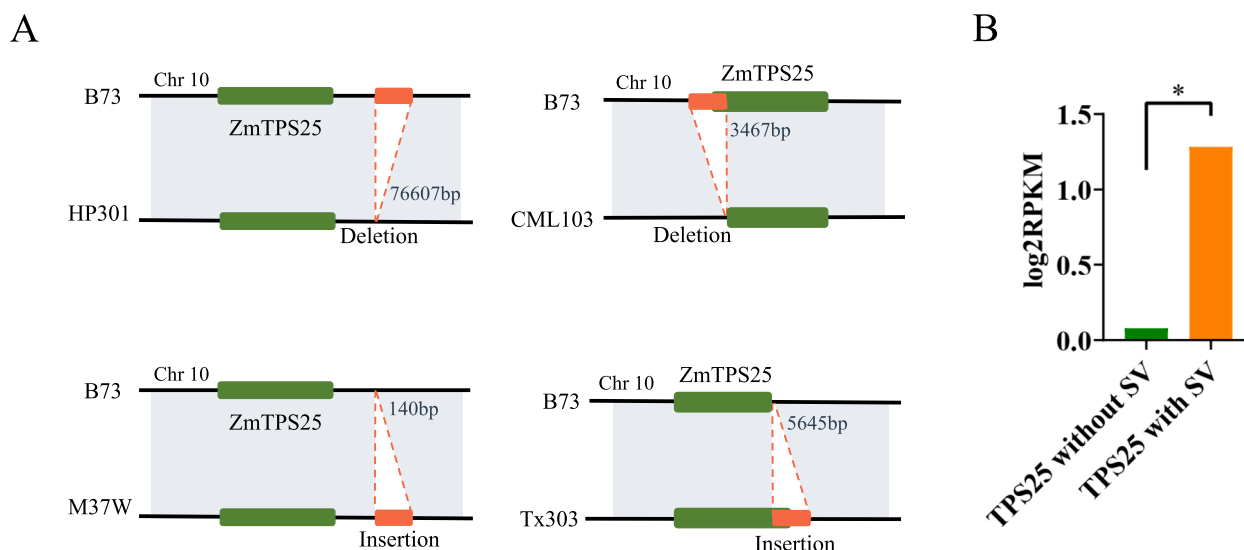


Fig. 3 Effect of SV on genes. **A** The effects of SV insertion and deletion on *ZmTPS25*. **B** The expression of *TPS25* was significantly affected by SVs

We analysed the gene structure of *ZmTPS25* in 26 maize genomes to explore whether SV affected the *TPS25* gene structure (Fig. 4). The domains of *ZmTPS25* in most maize genomes were consistent with the reference genome (B73). However, *ZmTPS25* structural domains in Mo18W, I114H, and CML277 were significantly altered, while one *ZmTPS25* gene in I114H had only one exon.

Atypical *ZmTPS* genes in the maize genomes

To explore the effects of SVs on the conserved structural domains of *ZmTPS* in different maize genomes, we selected Ms71, the variety with the highest number of *TPS* genes overlapping with SV, for conserved structural domain comparison with the reference (Fig. 5A). Four out of 10 motifs of *TPS* in Ms71 corresponded to the reference, but the remaining six did not. Furthermore, the amino acids in each set of mutually corresponding motifs did not perfectly match, indicating a strong effect of SV on the conserved structural domains of different *ZmTPS* genes. Because gene family identification is based on conserved structural domain searches, we calculated the number of typical (containing conserved structural domains) and atypical genes (without conserved structural domains) in each species (Fig. 5B). In most varieties, the atypical genes were *ZmTPS13*, *ZmTPS8*, and *ZmTPS20*. These findings imply that SVs influence *ZmTPS* conserved structural domains, leading to a significant number of atypical genes.

Atypical *TPS* genes were widely expressed in maize

To explore whether there is a correlation between the number and total expression dose of *ZmTPS* among

different varieties, we counted the number and total expression values of *ZmTPS* in each variety (Fig. 6). Each variety had almost the same number of *ZmTPS* genes, but the overall expression dose varied greatly, with a minimum total FPKM value of 7 (Oh43) and a maximum of 134.99 (NC358). In this study, we found that the absence of *ZmTPS* genes did not directly cause the total change in expression dose. For example, in CML52, the number of *TPS* genes was 25 and the total FPKM value was 77.88, whereas the number of *TPS* genes in CML69 was 28 and the total FPKM value was 36.72. In addition, some varieties had a small number of *TPS* genes that had low total *TPS* expression values; for example, there were only 25 *TPS* genes in Oh43, and the total FPKM value was 7.58. Pearson correlation analysis of the number of *TPS* and \log_2 RPKM values for each species revealed that $r=0.31$ and $p=0.12$ (Fig. S2), indicating that the number of *ZmTPS* genes and the total expression dose were not correlated.

In 26 maize genomes, *ZmTPS1*, *ZmTPS8*, *ZmTPS13*, *ZmTPS18*, and *ZmTPS2* had more than three atypical genes. Expression analysis showed that the expression of typical and atypical genes had different distribution patterns (Fig. 7). For example, the expression value distribution of *ZmTPS13* was significantly higher in maize varieties with typical structures than those with atypical structures. Atypical genes were widely expressed and distributed with high levels of expression in *ZmTPS8* and *ZmTPS20*, but this difference was not statistically significant.

The response of *ZmTPS* to the *Ostrinia furnacalis* feeding

The Asian corn borer (*Ostrinia furnacalis* Guenée) is an agricultural pest of several crops, mainly corn [32].



Fig. 4 Gene structure of *ZmTPS25* in 26 maize genomes

To analyse the response of *ZmTPS* genes to Asian corn borer feeding, we obtained differential expression data at different time points following Asian corn borer feeding from the study of Guo et al. [25]. All differentially expressed *ZmTPS* were upregulated in response to Asian

corn borer feeding. Among all differentially expressed *ZmTPS*, *ZmTPS13* and *ZmTPS16* had higher expression levels after corn borer feeding (Fig. 8A). The number of differentially expressed *TPS* increased with feeding time, reaching ten after 24 h, accounting for 31.25% of the total

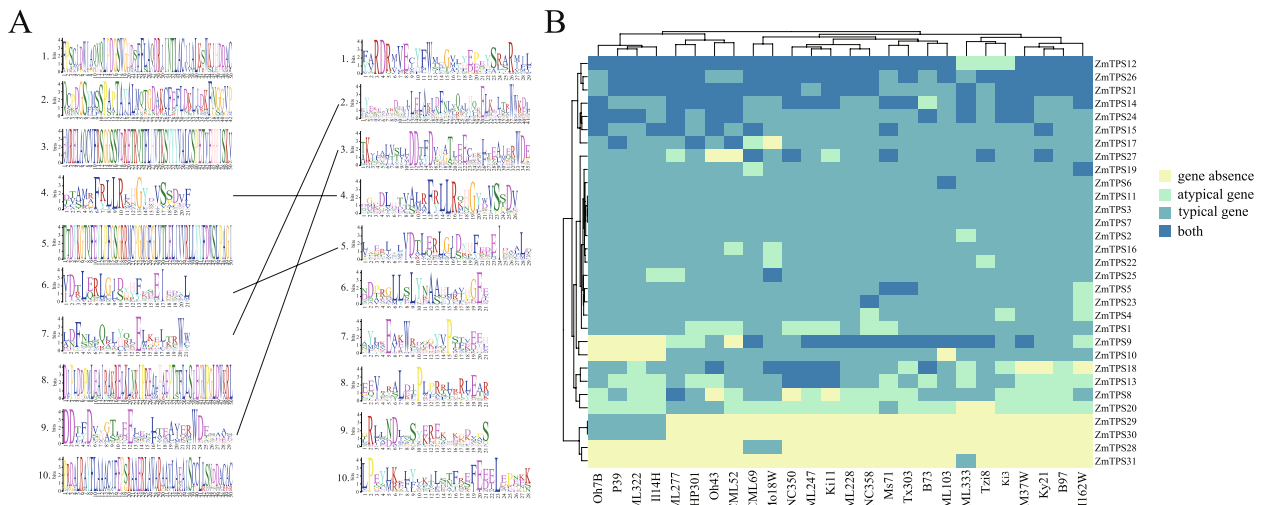


Fig. 5 The conserved structural domains of *TPS* were affected by SVs. **A** The weblogs of the Ms71 *ZmTPS* and the reference genome are shown on the left and right, respectively. The weblogs connected by the lines indicate that they are corresponding. Weblogs are arranged in the order of E-value. **B** Heatmap showing the typical or atypical nature of each *TPS* gene in each variety. 'both' represents the presence of both typical and atypical genes

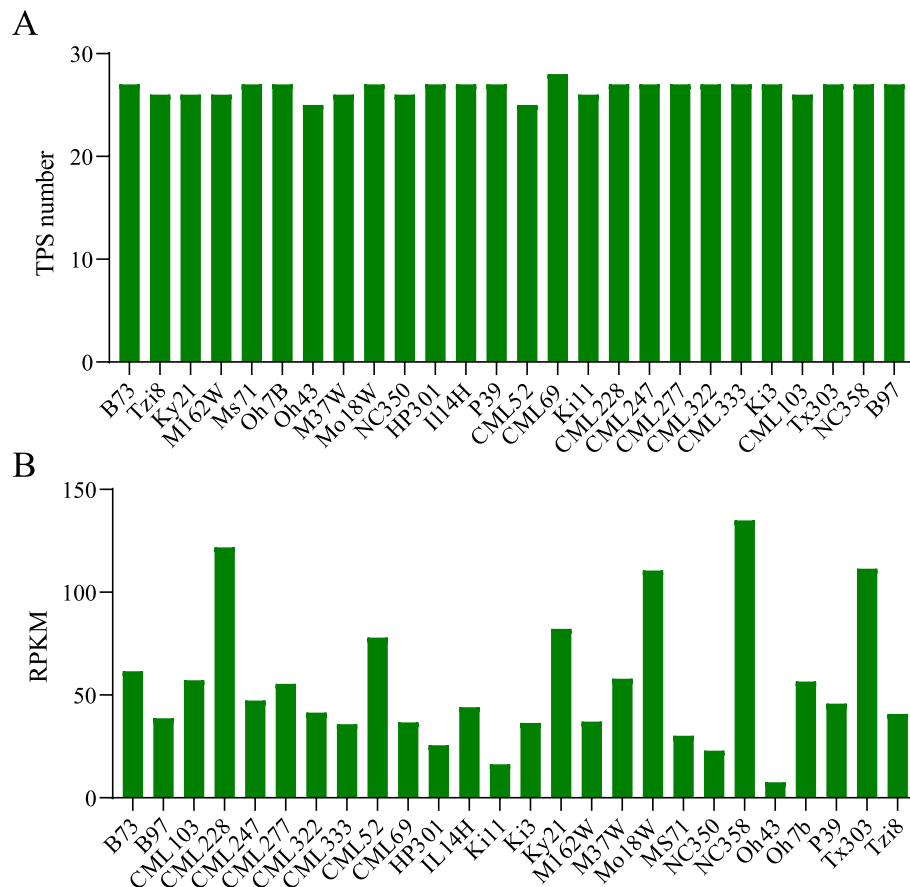


Fig. 6 Number of *ZmTPS* (A) and total expressed dose (B) in 26 maize genomes

number of *ZmTPS*, indicating that the *ZmTPS* gene family plays a vital role in response to corn borer infestation (Fig. 8B).

Discussion

The maize pan-genome consists of 26 high-quality genomes containing the genes absent in the reference genome. Compared to reference genome-based gene family analysis, pan-genome-based analysis is more comprehensive and can identify non-reference *ZmTPS* genes. We identified three non-reference *TPS* genes and 29 *TPS* genes in the reference genome based on the maize pan-genome. More importantly, even if a gene is present in the reference genome, it might not be present in other maize genomes. Only 20 of the 32 *ZmTPS* genes were present in all the varieties. This phenomenon was observed in different species. For example, a recent pan-genomic *TPS* gene analysis in rice found that one *OsTPS* gene was missing from the rice reference genome [33]. However, although the *ZmTPS* numbers differed, the varieties with fewer *ZmTPS* genes did not show low total *TPS* gene expression. This could be

due to many reasons, including the fact that some gene family members have similar functions and can compensate for one another. Consequently, when one or more genes are absent, expression of the spare genes increases to preserve the robustness of plant life activities. Some investigations have identified homologous genes with mutually compensatory mechanisms, such as the ability of paralogous *CLVI* homologues to be upregulated in the absence of *CLVI* gene activity, thus acting as a compensatory mechanism to replace *CLVI* in stem cell endocytosis [34]. *ZmTPS10*, *ZmTPS9*, *ZmTPS29*, and *ZmTPS30* were closely related in the phylogenetic tree, and none of them were core genes; however, the PAV heatmap (Fig. 1B) shows that none of them are simultaneously missing in any of the varieties. The mutually exclusive absence suggests that these four genes are functionally redundant with each other but will not be absent simultaneously to ensure the functioning of these genes.

In the present study, SV was found to affect the expression of *ZmTPS25*, which warrants further investigation. Previous studies showed that the promoter region of maize bZIP68 differs from the ancestral species *Z. mays*

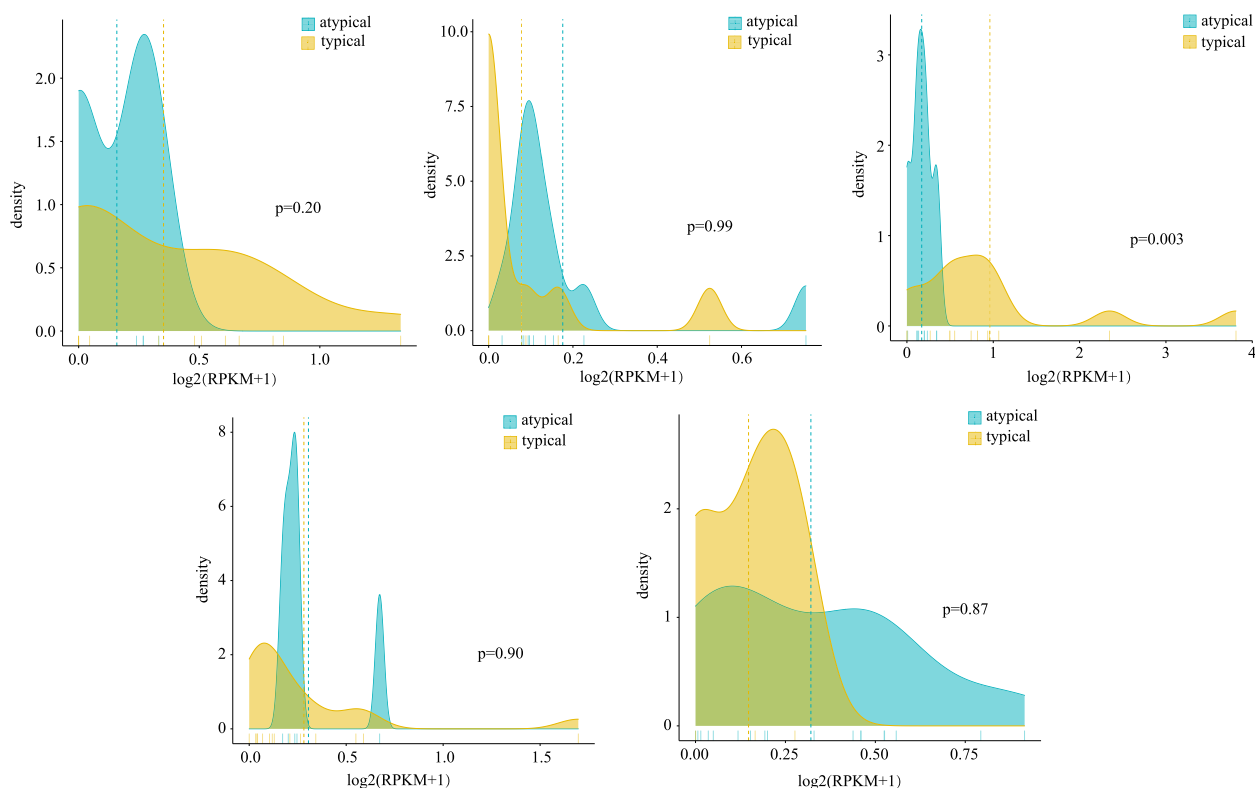


Fig. 7 Expression density distribution of typical and atypical genes in 26 maize genomes. In order, *ZmTPS1*, *ZmTPS8*, *ZmTPS13*, *ZmTPS18* and *ZmTPS20*

ssp. parviglumis and *Z. mays ssp. mexicana*. A 358-bp fragment insertion at the -972 (B73) locus upregulated the expression of bZIP68 in maize, leading to a reduction in cold tolerance [35, 36]. The effect of SV on the promoter or downstream regions of the *ZmTPS25* gene is shown in Fig. 3A, suggesting that structural variation in gene regions can also affect gene expression, as reported in a tomato pan-SV study [37]. SVs affected both *ZmTPS* expression as well as the gene sequence and structure. This may result in the absence of some domains, which would prevent the identification of gene family members using conventional bioinformatics techniques. However, gene family members lacking these regions may still have vital biological functions. For example, Zhao et al. identified a *Ptr* gene on rice chromosome 12 that encoded an unusual broad-spectrum disease resistance protein and contained a two-bp deletion resulting in a truncation of the encoded protein [38]. While maize pan-genomes were constructed based on multiple high-quality genomes, atypical gene family members arising due to SV affecting multiple genomes can be detected by a collinear relationship between genomes. SVs may also alter the number of gene exons. For example, *ZmTPS25* showed variations in the number of exons in different

varieties. These genes with a reduced number of exons may belong to atypical gene families but are still functional. For example, Zheng et al. identified an atypical *LEA* gene in *Ipomoea pes-caprae* L. with only one exon that confers salt/drought and oxidative stress tolerance [39]. Therefore, this study identified several atypical *TPS* genes in maize, providing additional resources that can aid in better understanding the function of *TPS* in maize.

A total of 10 *TPS* genes were differentially expressed in maize in response to *Ostrinia furnacalis* feeding, of which only *ZmTPS17* was a near-core gene, and the others were core genes. This suggests that *ZmTPS*, which is conserved in maize with minimal gene absences, may be implicated in the response of plants to insect feeding. *ZmTPS17*, a gene encoding monoterpene synthase, has been found to be upregulated in other studies on insect feeding on maize [40], indicating that it may be an important insect feeding response gene. *ZmTPS14*, a gene of the *TPS-b* class involved in the synthesis of sesquiterpenoids but not in the same subgroup as *ZmTPS17*, was induced after feeding by the Asian corn borer *Ostrinia furnacalis* and the larvae of the herbivorous beetle *Holotrichia parallela* [41]. *TPS15*, a gene adjacent to *TPS14* in the phylogenetic tree, catalyses the production of

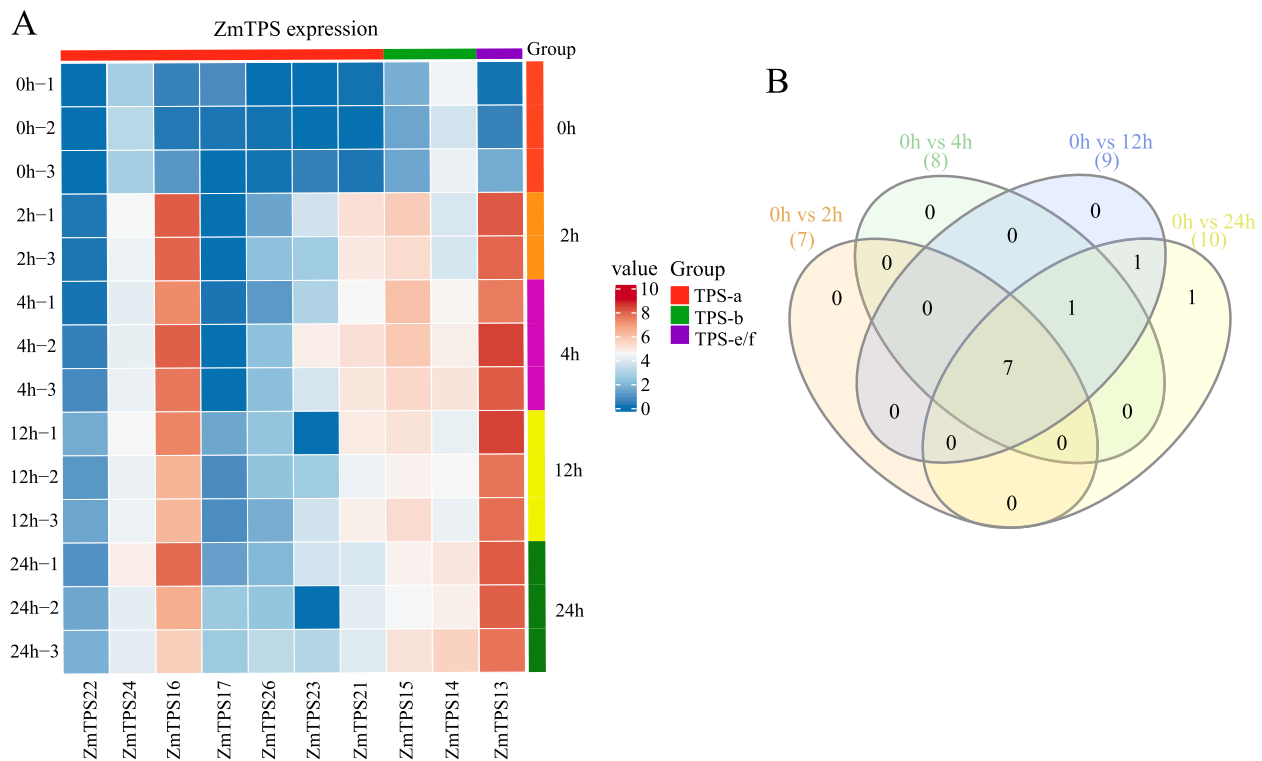


Fig. 8 Response of *ZmTPS* to *Ostrinia furnacalis* feeding. **A** Heatmap of differentially expressed *ZmTPS* expression in different samples. The value in the graph was $\log_2(\text{FPKM}+1)$ value. **B** Venn Diagrams of overlapping and specific differentially expressed *ZmTPS* genes for different comparisons

linalool, (E)-nerolidol, and (E, E)-geranylinalool, the expression of which has been shown to be induced by mechanical damage to maize seedlings combined with the oral secretion of caterpillars [42], suggesting that *TPS15* is also involved. These results indicate that *TPS* genes are critical for insect resistance in maize and are entirely conserved, except for *ZmTPS17*, which was absent in Mo18W. However, we found that some varieties have atypical genes, such as the novel *TPS14* copies in the genomes of M162W, M37W, and Ki3, which have both typical and atypical genes. These atypical *ZmTPS* genes warrant further investigation.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-023-09137-3>.

Additional file 1: Fig. S1. The ratio of SV to typical and atypical genes overlap. **Fig. S2.** Scatter plot of *ZmTPS* gene number and total expressed dose correlation analysis. **Fig. S3.** The differentially expressed *ZmTPS* genes in other studies. **Table S1.** The *ZmTPS* names and their corresponding gene names in multiple maize genomes. **Table S2.** The atypical *ZmTPS* genes in maize genomes.

Acknowledgements

This work was supported by the National Key R&D Program of China .

Authors' contributions

YS, XDK, and YJZ conceived and designed the experiments. YS, WQX, JW, QNW and LNX contributed to paper writing. SXL, WHM, PR and JW conducted the experiment. JW and PR contributed to the data analysis. The authors read and approved the final manuscript.

Funding

The National Key R&D Program of China (2021YFD1400700).

Availability of data and materials

The RNA-seq data were obtained from the NCBI SRA database (PRJNA531363) and Genome Sequence Archive of BIG Data Center (PRJCA006850). The maize genomes were downloaded from maize pangenome database (<http://maize-pangenome.gramene.org>). Data supporting the findings of this work are available within the paper and its Supplementary Information files. The datasets generated and analyzed during the current study are available from the corresponding author upon request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing financial interests.

Received: 28 September 2022 Accepted: 13 January 2023

Published online: 27 January 2023

References

- Loreto F, Dicke M, Schnitzler JP, Turlings TCJ. Plant volatiles and the environment. *Plant Cell Environ.* 2014;37:1905–8.
- Zhao M, Wang L, Wang J, Jin J, Zhang N, Lei L, et al. Induction of priming by cold stress via inducible volatile cues in neighboring tea plants. *J Integr Plant Biol.* 2020;62:1461–8.
- Taniguchi S, Hosokawa-Shinonaga Y, Tamaoki D, Yamada S, Akimitsu K, Gomi K. Jasmonate induction of the monoterpene linalool confers resistance to rice bacterial blight and its biosynthesis is regulated by JAZ protein in rice. *Plant Cell Environ.* 2014;37:451–61.
- Lee GW, Chung MS, Kang M, Chung BY, Lee S. Direct suppression of a rice bacterial blight (*Xanthomonas oryzae* pv. *Oryzae*) by monoterpene (S)-limonene. *Protoplasma.* 2016;253:683–90.
- Kiryu M, Hamanaka M, Yoshitomi K, Mochizuki S, Akimitsu K, Gomi K. Rice terpene synthase 18 (OsTPS18) encodes a sesquiterpene synthase that produces an antibacterial (E)-nerolidol against a bacterial pathogen of rice. *J Gen Plant Pathol.* 2018;84:221–9.
- Yoshitomi K, Taniguchi S, Tanaka K, Uji Y, Akimitsu K, Gomi K. Rice terpene synthase 24 (OsTPS24) encodes a jasmonate-responsive monoterpene synthase that produces an antibacterial γ -terpinene against rice pathogen. *J Plant Physiol.* 2016;191:120–6.
- Gao L, Zhang X, Zhou F, Chen H, Lin Y. Expression of a peppermint (E)- β -Farnesene synthase gene in Rice has significant repelling effect on bird cherry-oat aphid (*Rhopalosiphum padi*). *Plant Mol Biol Report.* 2015;33:1967–74.
- Wang X, Gao Y, Chen Z, Li J, Huang J, Cao J, et al. (E)- β -farnesene synthase gene affects aphid behavior in transgenic *Medicago sativa*. *Pest Manag Sci.* 2019;75:622–31.
- Xiao Y, Wang Q, Erb M, Turlings TCJ, Ge L, Hu L, et al. Specific herbivore-induced volatiles defend plants and determine insect community composition in the field. *Ecol Lett.* 2012;15:1130–9.
- Degenhardt J, Köllner TG, Gershenzon J. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry.* 2009;70:1621–37.
- Chen F, Tholl D, Bohlmann J, Pichersky E. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J.* 2011;66:212–29.
- Whittington DA, Wise ML, Urbansky M, Coates RM, Croteau RB, Christianson DW. Bornyl diphosphate synthase: structure and strategy for carbocation manipulation by a terpenoid cyclase. *Proc Natl Acad Sci U S A.* 2002;99:15375–80.
- Christianson DW. Structural and chemical biology of terpenoid cyclases[J]. *Chem Rev.* 2017;117(17):11570–648.
- Hufford MB, Seetharam AS, Woodhouse MR, Chougule KM, Coletta D, Tittes S, et al. De novo assembly, annotation, and comparative analysis of 26 diverse maize genomes. *Science (80-).* 2021;373:655–62.
- Letunic I, Khedkar S, Bork P. SMART: recent updates, new developments and status in 2020. *Nucleic Acids Res.* 2021;49:D458–60.
- Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30:772–80.
- Price MN, Dehal PS, Arkin AP. FastTree 2 - approximately maximum-likelihood trees for large alignments. *PLoS One.* 2010;5:e9490.
- Letunic I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 2021;49:W293–6.
- Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics.* 2016;32:2847–9.
- Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genom Proteom Bioinform.* 2010;8:77–80.
- Wilke CO. Ggrridges: ridgeline plots in 'ggplot2'; 2018. p. 1.
- Wickham H, Chang W. Package 'ggplot2'; 2016. p. 1–189.
- Qin P, Lu H, Du H, Wang H, Chen W, Chen Z, et al. Pan-genome analysis of 33 genetically diverse rice accessions reveals hidden genomic variations. *Cell.* 2021;184:3542–3558.e16.
- Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant.* 2020;13:1194–202.
- Guo J, Qi J, He K, Wu J, Bai S, Zhang T, et al. The Asian corn borer *Ostrinia furnacalis* feeding increases the direct and indirect defence of mid-whorl stage commercial maize in the field. *Plant Biotechnol J.* 2019;17:88–102.
- Ma C, Li R, Sun Y, Zhang M, Li S, Xu Y, et al. ZmMYC2s play important roles in maize responses to simulated herbivory and jasmonate. *J Integr Plant Biol.* 2022;00:1–18.
- Zhang C, Li J, Li S, Ma C, Liu H, Wang L, et al. ZmMPK6 and ethylene signalling negatively regulate the accumulation of anti-insect metabolites DIMBOA and DIMBOA-Glc in maize inbred line A188. *New Phytol.* 2021;229:2273–87.
- Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 2018;34:i884–90.
- Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol.* 2019;37:907–15.
- Putri GH, Anders S, Pyl PT, Pimanda JE, Zanini F. Analysing high-throughput sequencing data in python with HTSeq 2.0. *Bioinformatics.* 2022;38:2943–5.
- Aubourg S, Lecharny A, Bohlmann J. Genomic analysis of the terpenoid synthase (AtTPS) gene family of *Arabidopsis thaliana*. *Mol Gen Genomics.* 2002;267:730–45.
- Jacobs DS, Bastian A. Predator–prey interactions: co-evolution between bats and their prey. In: *Co-evolution*; 2016. p. 1–16.
- Sun Y, Zhang P, Kou D, Han Y, Fang J, Ni J, et al. Terpene synthases in Rice Pan-genome and their responses to Chilo suppressalis larvae infesting. *Front Plant Sci.* 2022;13:1–11.
- Rodriguez-Leal D, Xu C, Kwon CT, Soyars C, Demesa-Arevalo E, Man J, et al. Evolution of buffering in a genetic circuit controlling plant stem cell proliferation. *Nat Genet.* 2019;51:786–92.
- Li Z, Fu D, Wang X, Zeng R, Zhang X, Tian J, et al. Natural variation in the bZIP68 promoter modulates cold tolerance and was targeted during maize domestication. *Plant Cell.* 2022;5:11.
- Hufford MB, Xu X, Van Heerwaarden J, Pyhäjärvi T, Chia JM, Cartwright RA, et al. Comparative population genomics of maize domestication and improvement. *Nat Genet.* 2012;44:808–11.
- Alonge M, Wang X, Benoit M, Soyk S, Pereira L, Zhang L, et al. Major impacts of widespread structural variation on gene expression and crop improvement in tomato. *Cell.* 2020;182:145–161.e23.
- Zhao H, Wang X, Jia Y, Minkenber B, Wheatley M, Fan J, et al. The rice blast resistance gene Ptr encodes an atypical protein required for broad-spectrum disease resistance. *Nat Commun.* 2018;9:1–12.
- Zheng J, Su H, Lin R, Zhang H, Xia K, Jian S, et al. Isolation and characterization of an atypical LEA gene (IpLEA) from *Ipomoea pes-caprae* conferring salt/drought and oxidative stress tolerance. *Sci Rep.* 2019;9:1–21.
- Rodriguez VM, Padilla G, Malvar RA, Kallenbach M, Santiago R, Butrón A. Maize stem response to long-term attack by *Sesamia nonagrioides*. *Front Plant Sci.* 2018;9:1–11.
- Pan Y, Zhao S-w, Tang X-l, Wang S, Wan X, Zhang X-x, et al. Transcriptome analysis of maize reveals potential key genes involved in the response to belowground herbivore H. parallela larvae feeding. *Genome.* 2020;63:1–12.
- Wang D, Li S, Liu J, et al. Study on the function of Terpene synthase gene tps2 and its promoter functional segment in *Zea mays*. *Biotechnol Bull.* 2020;36:1.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

