



Retention and Molecular Evolution of Lipoxygenase Genes in Modern Rosid Plants

Zhu Chen¹, Danmei Chen¹, Wenyuan Chu¹, Dongyue Zhu¹, Hanwei Yan^{1,2} and Yan Xiang^{1,2*}

¹ Laboratory of Modern Biotechnology, Anhui Agricultural University, Hefei, China, ² Key Laboratory of Biomass Improvement and Conversion, Anhui Agriculture University, Hefei, China

OPEN ACCESS

Edited by:

Rosalba Giugno,
University of Verona, Italy

Reviewed by:

Matteo Benelli,
University of Trento, Italy
Giovanni Bacci,
University of Florence, Italy
Debjit Ray,

Sandia National Laboratories, USA

*Correspondence:

Yan Xiang
xiangyanahau@sina.com

Specialty section:

This article was submitted to
Bioinformatics and Computational
Biology,
a section of the journal
Frontiers in Genetics

Received: 06 June 2016

Accepted: 16 September 2016

Published: 30 September 2016

Citation:

Chen Z, Chen D, Chu W, Zhu D,
Yan H and Xiang Y (2016) Retention
and Molecular Evolution of
Lipoxygenase Genes in Modern Rosid
Plants. *Front. Genet.* 7:176.
doi: 10.3389/fgene.2016.00176

Whole-genome duplication events have occurred more than once in the genomes of some rosids and played a significant role over evolutionary time. Lipoxygenases (LOXs) are involved in many developmental and resistance processes in plants. Our study concerns the subject of the *LOX* gene family; we tracked the evolutionary process of ancestral *LOX* genes in four modern rosids. Here we show that some members of the *LOX* gene family in the *Arabidopsis* genome are likely to be lost during evolution, leading to a smaller size than that in *Populus*, *Vitis*, and *Carica*. Strong purifying selection acted as a critical role in almost all of the paralogous and orthologous genes. The structure of *LOX* genes in *Carica* and *Populus* are relatively stable, whereas *Vitis* and *Arabidopsis* have a difference. By searching conserved motifs of *LOX* genes, we found that each sub-family shared similar components. Research on intraspecies gene collinearity show that recent duplication holds an important position in *Populus* and *Arabidopsis*. Gene collinearity analysis within and between these four rosid plants revealed that all *LOX* genes in each modern rosid were the offspring from different ancestral genes. This study traces the evolution of *LOX* genes which have been differentially retained and expanded in rosid plants. Our results presented here may aid in the selection of special genes retained in the rosid plants for further analysis of biological function.

Keywords: lipoxygenase, purifying selection, gene duplication, syntenic chromosomal block, evolutionary history

INTRODUCTION

Whole-genome duplications (WGDs) bring a huge impact on genome sizes of many angiosperms and may have provided the genetic material for evolutionary novelties (Sémon and Wolfe, 2007; Jaillon et al., 2009). Duplication events are usually followed by gene loss (Bowers et al., 2003), nucleotide divergence (Bowers et al., 2003) and structural rearrangements (Hufton and Panopoulou, 2009). It has long been hypothesized that the ancient genome triplication event happened to a single common ancestor of *Arabidopsis*-*Populus*-*Vitis*-*Carica* and finally caused a paleohexaploid (Tang et al., 2008a). Other than that, the two recent paleopolyploidies that have affected *Arabidopsis* are β - and α - duplications. At- α was a recent event, and the At- β was an intermediate event (Barker et al., 2009). In *Populus*, there was a single genome-wide event. This duplication event was called the “salicoid” duplication event (P-duplication; Tuskan et al., 2006). *Vitis vinifera* and *Carica papaya* each have only γ -triplication event and no other polyploidies (Tang et al., 2008a).

Polyploidy has been and continues to have an extensive effect on the number or type of genes in plant evolution (Adams and Wendel, 2005). Analysis of the differential retention and expansion of ancestral genes in modern plants provide an informative and robust way to resolve relationships among many lineages (Rokas and Holland, 2000). In this study, we will take the Lipoxygenase gene family as an example and discuss the differential retention and expansion of ancestral genes in four rosids.

Lipoxygenases (LOXs) exist extensively within plants and animals (Brash, 1999). The best known function of these enzymes are to synthesize lipid mediators (Brash, 2015): as we know, leukotrienes and resolvins are in animals, jasmonates and short-chain aldehydes are in plants. LOXs catalyze polyenoic fatty acids PUFAs (Feussner and Kühn, 2000) like linoleic acid (LA), α -linolenic acid (α -LeA), or arachidonic acid, which have a (1Z, 4Z)-pentadiene moiety. According to their positional specificity of linoleic acid oxygenation, lipoxygenases have been divided into group 9-LOX and group 13-LOX (Hildebrand, 1989). LOXs contain a region rich in histidine residues, which was previously observed to be highly conserved in the primary structure of isozymes. This region contains a cluster of 5 His residues in the form of His-(X)4-His-(X)4-His-(X)17-His-(X)8-His (Shibata et al., 1987; Steczko et al., 1992; Boyington et al., 1993; Feussner and Wasternack, 2002).

Lipoxygenases involved in food-related applications during bread-making and production of the aroma are controlled by enzymes, which were found related to the formation of volatile compounds (Leenhardt et al., 2006). Studies have shown that extractable activities of enzymes are major factors that can affect the degrading efficiency of carotenoid pigments during the kneading step of bread-making in each of the three cultivated wheat species. Lipoxygenases also have a negative relationship with the color, off-flavor and antioxidant status of plant-based foods. Studies on soy-based foods have demonstrated that lipoxygenases are responsible for the off-flavor associated with biological components present in soybean (Leenhardt et al., 2006). So far, there is sufficient evidence to prove that lipoxygenase is the most crucial element in plant defense responses (Baysal and Demirdöven, 2007; Bannenberg et al., 2009). In recent years, one LOX gene in *Arabidopsis* (*AtLOX2*) was thought to function exclusively in jasmonates (JA) biosynthesis upon wounding (Van Loon et al., 2006). The study is backed up by recent findings in apple which showed that *MdLOX5* gene was more likely to be responsible for aphid tolerance or resistance (Vogt et al., 2013).

Lipoxygenase genes are chosen for their biological significance. In our research, taking lipoxygenases as an example, we studied the expansion of these genes in four species. Previous analysis showed that one or more paleopolyploidy events which had an impact on these four modern rosid genomes, fluctuate remarkably in size and arrangement. Our results trace the differential retention and expansion of the ancestral Lipoxygenases in *Arabidopsis thaliana*, *Populus trichocarpa*, *V. vinifera* and *C. papaya* and help facilitate the extrapolation of the evolutionary process.

MATERIALS AND METHODS

Ethics Statement

No specific permits were required for the described field studies. No specific permissions were required for these locations and activities. The location is not privately-owned or protected in any way and the field studies did not involve endangered or protected species.

Database Search and Sequence Retrieval

LOX genes were identified following the method described by (Podolyan et al., 2010; Umate, 2014; Chen et al., 2015). Protein and cDNA sequences of LOX genes in *Arabidopsis* were obtained from the *Arabidopsis* Information Resource (TAIR, <http://www.arabidopsis.org/>, release 10.0). Protein and cDNA sequences of *P. trichocarpa*, *V. vinifera*, and *C. papaya* were downloaded from PhytozomeV11.0 database. The respective genome sequence sites are as follows: *P. trichocarpa*, *V. vinifera*, and *C. papaya* (<http://genome.jgi.doe.gov/pages/dynamicOrganismDownload.jsf?organism=PhytozomeV11>).

Local Blast searching was performed using *Arabidopsis* LOX proteins as queries for the identification of LOX genes from *Carica* and *Vitis*, and then using the resulting poplar and grapevine sequences as secondary queries. To obtain good gene models of *CpLOX* genes, all-against-all nucleotide sequence similarity searches were performed between the gene models and EST sequences using BLASTN software (Supplementary Table 1). Besides, we also worked on the multiple sequence alignment with *CpLOX* proteins and 70 experimentally verified gene models (Supplementary Tables 2, 3). All of the obtained genes were further manually analyzed to confirm the presence of the LOX domain (PF00305) and PLAT/LH2 (polycystin-1, lipoxygenase, α -toxin domain, or the lipoxygenase homology) domain (PF01477) in the Pfam HMM database (<http://pfam.sanger.ac.uk/>) (Finn et al., 2006) and InterPro (European Bioinformatics Institute) (<http://www.ebi.ac.uk/interpro/scan.html>) (Supplementary Table 4; Mulder et al., 2007). Redundant sequences with different identification numbers and the same chromosome locus were removed.

Phylogenetic Trees Construction

Complete protein sequences of LOX in the four plant species were aligned with the aid of ClustalW (Larkin et al., 2007). The phylogenetic tree was constructed by MEGA version 6.0 software with the minimum evolution (ME) method (Tamura et al., 2013). Bootstrap analysis with 1000 replicates was performed to calculate the reliability of the ME tree. To confirm the robustness of the ME tree, we also constructed other phylogenetic trees by using the Neighbor-Joining (NJ) method.

Exon–Intron Structural Analysis and Identification of Conserved Motifs

The exon–intron structure positions of LOX genes were generated online using the online program Gene Structure Display Server (GSDS; <http://gsds.cbi.pku.edu.cn/>; Guo et al.,

2007) by alignment of the cDNAs with their corresponding genomic DNA sequences. To identify the conserved motifs of the *LOX* genes in four rosid plants, the structural motif annotation was employed using the MEME (Multiple Em for Motif Elicitation, Version 4.11.1) program (Bailey et al., 2006) with the following parameters: the maximum number of motifs was set at 20, and the optimum motif widths were set between six and 200 residues. Structural motif annotation was provided by using the Batch search tool in Pfam program.

Identification of Paralogs and Orthologs

Paralogs and orthologs were identified by using the same procedure described in Blanc and Wolfe (2004). This method was performed by running a BLASTN (Altschul et al., 1997) for all nucleotide sequences for each species. A pair of matching sequences were defined as pairs of paralogs when the identity was more than 40% and the alignment covered over 300 bp. To identify putative orthologs between two species, for example A and B, each sequence from species A was searched against all sequences from species B using BLASTN. At the same time, each sequence from species B was searched against all sequences from species A. The two sequences were defined as orthologs whose reciprocal best hits were each within ≥ 300 bp of the two sequences aligned.

ω and Ks Analysis

First, pairwise protein sequence alignment was performed using MUSCLE (Edgar, 2004). Then, used in conjunction with protein alignments, CDS sequence, and an in-house PERL script, the input file format of KaKs_Calculator2.0 could be got. Finally, the input file were converted into computation of Ks (synonymous substitution rate) and Ka (non-synonymous substitution rate) values using KaKs_Calculator2.0 (Wang et al., 2010). To further assess whether positive selection acts upon specific sites, the gene pairs for all the paralogs and orthologs were used to calculate the ω , where $\omega = Ka/Ks$.

Intraspecies and Interspecies Microsynteny Analysis

Microsynteny analysis within four species was detected by MicroSyn software (Cai et al., 2011). At the beginning of the generated MSY file step, three property files are needed: the gene list file, the CDS file and the gene identifier file. The microsynteny graphic file was provided by loading the three files. Then MicroSyn creates a homologous relationship among all genes. Finally, the microsynteny graphic file was provided by the software. In order to analyze duplication of *LOX* genes of the four rosid plants we researched the expansion of *LOX* genes through segmental or whole-genome duplications (S/WGD) for *LOX* genes in each species by using the plant genome duplication database (PGDD; <http://chibba.agtec.uga.edu/duplication/>; Lee et al., 2013). In order to determine whether the *LOX* genes of the four rosid plants arose from a large-scale duplication event (duplicated blocks derived from whole-genome or segmental duplication) or tandem duplication, genome-wide analysis was undertaken to examine whether *LOX* genes occurred within duplicated blocks. We researched the expansion of *LOX* genes

through segmental or whole-genome duplications (S/WGD) for *LOX* genes in each species by using the plant genome duplication database (PGDD; <http://chibba.agtec.uga.edu/duplication/>). The *LOX* genes duplicated through S/WGDs were inferred on the basis of gene collinearity on syntenic blocks. First, from the PGDD, we download the file containing collinear block information within and between the four rosids. Then, all download blocks information were imported into MySQL, and the *LOX* gene ID were used as the query to perform a search in these species. The *LOX* genes duplicated through S/WGDs were identified in this way. In this analysis, the counterparts of a particular *LOX* gene on an SCB may have been retained as LOXs, subfunctionalized into non LOXs (indicated by an “N” before the first letter in the code name). *LOX* genes expanded through tandem duplication (TD) were inferred following the method that (1) belong to *LOX* gene family, (2) are located within 60 kb each other (data comes from Phytozome), and (3) are separated by five or fewer gene loci (non LOXs). Syntenic blocks between species were identified using the MCscanX (Wang et al., 2012) software with default parameters. To categorize the expansion of the *LOX* gene families, the positions of the *LOX* genes in the blocks were *A. thaliana*, *P. trichocarpa*, *V. vinifera*, and *C. papaya*. Circos software was used to draw the syntenic diagram (Krzywinski et al., 2009).

RESULTS

LOX Genes in Four Modern Rosids

Based on the previous studies, we obtained 6 and 20 putative *LOX* genes from the *Arabidopsis*, and poplar, respectively. In a recently published report, a total of 18 *LOX* genes were identified in *Vitis*. By removing pseudogenes, 13 *LOX* genes were identified in the *Vitis* genome. In this study, by removing pseudogenes we further filtered five additional *LOX* genes in *Vitis* and changed the total member into 13. To identify *LOX* in *Carica*, we performed a search against the genome database with BlastP using At*LOX* protein sequences as queries. Finally, 11 *LOX* genes were identified in *Carica*. The detail information of each *LOX* genes are listed in **Table 1**.

To date, four studied rosids have been suggested to possess paleohexaploidy in a common ancestor (Jaillon et al., 2007). Based on previous results, the multiplicity ratio for an ancestral gene comparison in the genomes of four species should be 4:2:1:1. But in our research results, the number of *LOX* genes in *Arabidopsis* is far fewer than that estimated for other plant species. Previous studies suggest that the *V. vinifera* genome is by far the closest to the ancestral arrangement such that the ancestral gene order can be deduced from this species with no difficulty. The ratio of *LOX* genes for the four species is 0.5: 1.5: 0.85: 1 when the number of *LOX* genes in *V. vinifera* is used as a benchmark. In addition to *A. thaliana*, this result is basically in line with the expected current ratios of *LOX*s.

LOX Paralogs and Orthologs

We detected 33.3% (2/6, *Arabidopsis*), 72.7% (8/11, *Carica*), 95% (19/20, *Populus*), and 76.9% (10/13, *Vitis*) *LOX* genes in

TABLE 1 | Detailed information about the LOX gene family in rosid plants.

Species	Gene name	Gene ID	Chr.	Location coordinates(5'-3')	Protein length(a.a.)	ORF length(bp)
<i>Carica papaya</i>	<i>CpLOX1</i>	evm.TU.supercontig_8.58	supercontig_8	393,002-397,307	797	2394
	<i>CpLOX2</i>	evm.TU.supercontig_17.119	supercontig_17	1,496,985-1,501,648	816	2451
	<i>CpLOX3</i>	evm.TU.supercontig_25.128	supercontig_25	1,317,447-1,322,293	925	2778
	<i>CpLOX4</i>	evm.TU.supercontig_32.35	supercontig_32	482,778-486,631	854	2565
	<i>CpLOX5</i>	evm.TU.supercontig_32.64	supercontig_32	770,894-774,557	855	2568
	<i>CpLOX6</i>	evm.TU.supercontig_43.30	supercontig_43	308,246-311,725	922	2769
	<i>CpLOX7</i>	evm.TU.supercontig_48.63	supercontig_48	351,329-356,020	867	2604
	<i>CpLOX8</i>	evm.TU.supercontig_58.126	supercontig_58	1,216,903-1,220,688	849	2550
	<i>CpLOX9</i>	evm.TU.supercontig_58.127	supercontig_58	1,233,011-1,236,967	849	2550
	<i>CpLOX10</i>	evm.TU.supercontig_458.2	supercontig_458	18,363-22,072	788	2367
	<i>CpLOX11</i>	evm.TU.supercontig_458.4	supercontig_458	24,745-28,879	917	3754
<i>Arabidopsis thaliana</i>	<i>AtLOX1</i>	AT1G55020	1	20,525,708-20,530,273	859	2580
	<i>AtLOX2</i>	AT3G45140	3	16,525,410-16,529,352	896	2691
	<i>AtLOX3</i>	AT1G17420	1	5,977,411-5,981,480	919	2760
	<i>AtLOX4</i>	AT1G67560	1	25,319,899-25,324,264	917	2754
	<i>AtLOX5</i>	AT3G22400	3	7,926,879-7,931,351	886	2661
	<i>AtLOX6</i>	AT1G72520	1	27,308,515-27,312,754	926	2781
<i>Vitis vinifera</i>	<i>VvLOX1</i>	GSVVT01010359001	1	19,772,666-19,777,638	920	2763
	<i>VvLOX2</i>	GSVVT01017943001	5	4,934,967-4,939,395	751	2256
	<i>VvLOX3</i>	GSVVT01025342001	6	1,774,659-1,781,744	817	2454
	<i>VvLOX4</i>	GSVVT01025340001	6	1,853,936-1,868,694	872	2619
	<i>VvLOX5</i>	GSVVT01025339001	6	1,875,239-1,882,842	901	2706
	<i>VvLOX6</i>	GSVVT01025328001	6	1,988,366-1,989,826	335	1008
	<i>VvLOX7</i>	GSVVT01005730001	7	13,887,191-13,893,238	641	1926
	<i>VvLOX8</i>	GSVVT01016738001	9	811,736-816,741	927	2784
	<i>VvLOX9</i>	GSVVT01032029001	13	23,366,475-23,371,929	866	2601
	<i>VvLOX10</i>	GSVVT01000083001	14	3,311,501-3,315,829	738	2217
	<i>VvLOX11</i>	GSVVT01000084001	14	3,315,947-3,324,623	900	2703
	<i>VvLOX12</i>	GSVVT01003798001	chr7_random	201,678-209,816	619	1860
	<i>VvLOX13</i>	GSVVT01005215001	Un	19,276,130-19,281,690	533	1602
<i>Populus trichocarpa</i>	<i>PtLOX1</i>	Potri.001G015300	1	1,076,313-1,081,197	898	2697
	<i>PtLOX2</i>	Potri.001G015400	1	1,090,420-1,098,069	902	2709
	<i>PtLOX3</i>	Potri.001G015500	1	1,105,670-1,110,895	898	2697
	<i>PtLOX4</i>	Potri.001G015600	1	1,118,168-1,123,930	898	2697
	<i>PtLOX5</i>	Potri.001G167700	1	14,106,872-14,112,847	923	2772
	<i>PtLOX6</i>	Potri.003G067600	3	9,576,888-9,583,048	925	2778
	<i>PtLOX7</i>	Potri.005G032400	5	2,425,802-2,431,106	866	2601
	<i>PtLOX8</i>	Potri.005G032600	5	2,435,033-2,439,658	796	2391
	<i>PtLOX9</i>	Potri.005G032700	5	2,451,619-2,456,194	866	2601
	<i>PtLOX10</i>	Potri.005G032800	5	2,462,946-2,469,256	863	2592
	<i>PtLOX11</i>	Potri.008G151500	8	10,276,751-10,281,394	880	2643
	<i>PtLOX12</i>	Potri.008G178000	8	12,146,645-12,151,320	927	2784
	<i>PtLOX13</i>	Potri.009G022400	9	3,421,114-3,425,183	901	2706
	<i>PtLOX14</i>	Potri.010G057100	10	8,651,258-8,655,916	926	2781
	<i>PtLOX15</i>	Potri.010G089500	10	11,305,668-11,310,367	881	2646
	<i>PtLOX16</i>	Potri.013G022000	13	1,454,479-1,459,136	871	2616
	<i>PtLOX17</i>	Potri.013G022100	13	1,461,474-1,466,287	862	2589
	<i>PtLOX18</i>	Potri.014G018200	14	1,725,218-1,731,715	860	2583
	<i>PtLOX19</i>	Potri.014G177200	14	14,542,431-14,547,953	860	2583
	<i>PtLOX20</i>	Potri.017G046200	17	3,854,572-3,860,007	898	2697

each species involved in paralogous duplication (Supplementary Table 5). Thus, over half of the *LOX*s were closely bound up with intra-specific duplication in *Carica*, *Vitis*, and *Populus*. By contrast, there was just one pair of *LOX* paralogs in *Arabidopsis*, although this species has been expanded by three rounds of whole-genome duplication. The higher ratio in *Populus* reflects the preferential gene retention after multiple rounds of WGD. Our results show that *Populus* and *Vitis* shared the most orthologous pairs, of up to 26 pairs of orthologous *LOX*s. We only got one pair of orthologous *LOX*s between *Arabidopsis* and *Vitis*. Two pairs of orthologous *LOX*s were detected between *Arabidopsis* and *Populus*. After comparing *Carica* and other three species we found no orthologous *LOX*s between them.

In order to better understand the evolutionary constraints acting among the four rosoid species, we measured the Ka/Ks ratios for these pairs of *LOX* paralogs and orthologs. The ratio of non-synonymous substitutions per non-synonymous site (Ka) vs. the synonymous substitutions per synonymous site (Ks) is an indicator of the history of selection (Yang and Bielawski, 2000). If Ka/Ks < 1, it suggests that the gene is undergoing purifying selection. When Ka/Ks > 1, it means there is accelerated devolution with positive selection, and Ka/Ks = 1 suggests neutral selection. A summary of Ka/Ks for *LOX* paralogous and orthologous pairs is shown in Supplementary Table 5. The resulting pairwise comparison data showed the Ka/Ks values of only one Vv paralogous pair larger than 1. The relatively higher Ka/Ks ratio of VvLOX6/11 suggests that they may have experienced relatively rapid evolution following duplication. There were two *Vitis* pairs (VvLOX6/10 and VvLOX7/12) and one *Populus* pair (PtLOX18/19) that were larger than 0.5 but less than 1, while all of the remaining Ka/Ks ratios were less than 0.5, suggesting that the *LOX* family has mainly undergone strong purifying selection and these *LOX* genes are slowly evolving at the protein level. Our calculations shows that all orthologous *LOX*s between species were less than 1.

Expansion and Structural Characteristics of the *LOX* Genes in Four Rosid Plants

To investigate the extent of the expansion of the *LOX* genes in rosoid plants, we performed a joint phylogenetic analysis with MEGA using the ME method (Figure 1) and the NJ method (Supplementary Figure 1). The ME and NJ trees show identical topologies. As mentioned above, plant lipoxygenases are clustered into two groups (9-*LOX* and 13-*LOX*). In our study, a total of 50 genes formed two distinct clades and are in agreement with the previously studied results (Brash, 1999; Figure 1). As shown in Figure 1, 9-*LOX* consisted of 20 *LOX* genes from four modern rosids; two from *Arabidopsis*, eight from *Populus*, six from *Carica*, and four from *Vitis*. This clade is composed of four sub-clades, one of which includes purely six *Populus* *LOX* genes. Yet there is another sub-clade simply containing four *Carica* *LOX* genes. The rest includes *LOX* genes from two or more species. Paralogous groups in this clade are CpLOX2/4/5/7 and CpLOX8/9 from *Carica*; PtLOX7/8/9/10/16/17 and PtLOX11/15 from *Populus*; VvLOX6/10/11 from *Vitis*. Beyond that there are three orthologous pairs shared by the four modern rosids. The remaining *LOX*s are placed in the 13-*LOX*s group. Paralogous groups in this clade were A3-A6, from *Arabidopsis*; PtLOX1/2,

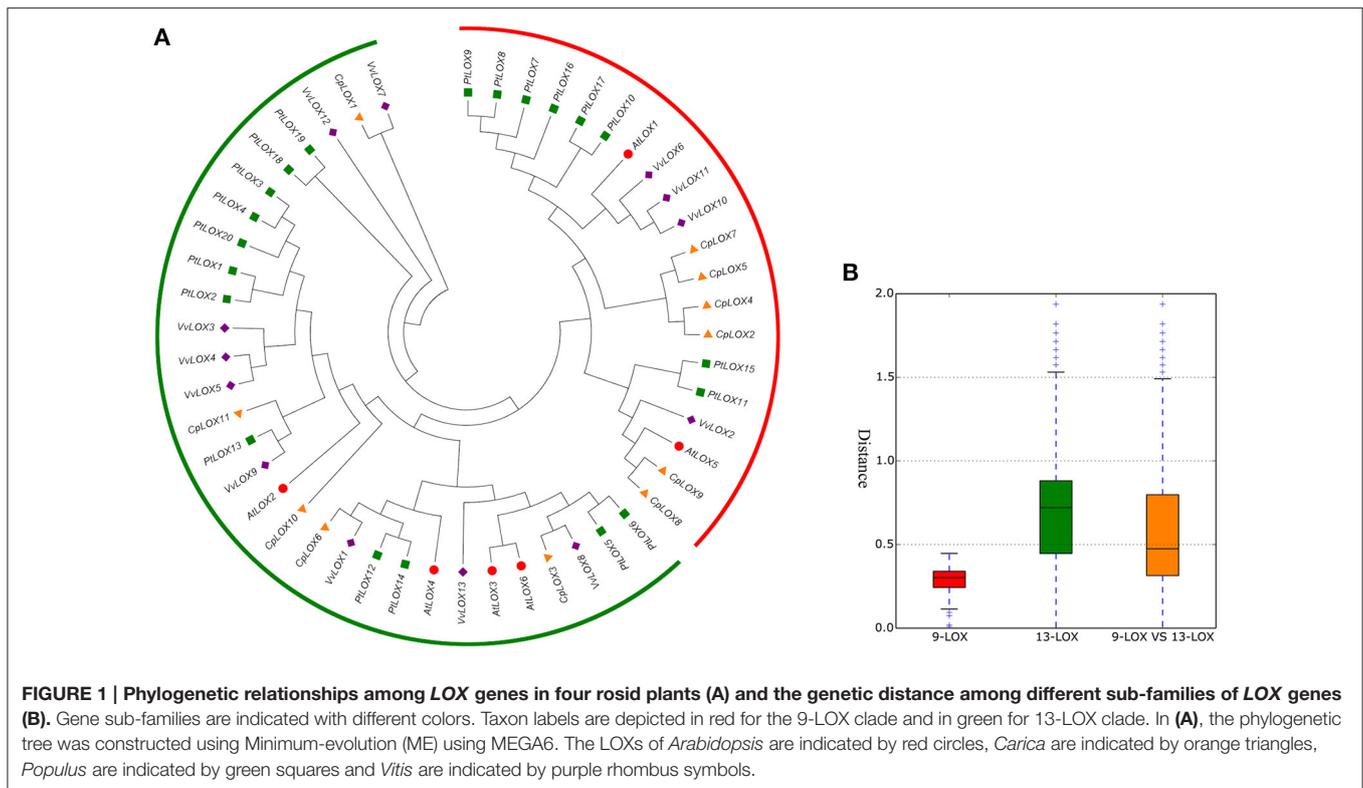
PtLOX3/4/20, PtLOX17/26, and PtLOX10/12 from *Populus*; VvLOX3/4/5, VvLOX4/9 and VvLOX7/12 from *Vitis*. In addition, this clade contained only one paralogous pair from *Carica*, CpLOX10/11. In this group, there are seven orthologous pairs shared by the four modern rosids. Besides, the genetic distances among the two *LOX* sub-families were studied and the result showed that the genetic distance of 9-*LOX* genes was smaller than 13-*LOX* genes, indicating that 9-*LOX* genes are more closely related to each other.

For a better understanding of the structural diversity of *LOX* genes, using the structures of *LOX* genes we generated the exon-intron architecture of each *LOX* gene in four rosoid plants (Figure 2). Overall, the structures of *LOX* genes in *Carica* and *Populus* were conserved. But some changes take place in the *AtLOX* and *VvLOX* genes. The detailed structural analysis of the exon/intron are presented in Figure 3. Of the four species surveyed, *Carica* and *Populus* *LOX* genes are in a similar position with eight or nine exons, and the number of exons in *Arabidopsis* range from six to nine. *VvLOX* genes are much more dramatic, VvLOX9/10/11 have the highest number of exons at 11, but in the same species VvLOX6 also has the least number of exons at five. We further analyzed the exon/intron structure of the *LOX* orthologous and paralogous gene pairs discussed previously. The results showed that majority of these gene pairs have different exon numbers. Among paralogous gene pairs, the structure rationality changes obviously in VvLOX6/10. Simultaneously, by comparing the orthologous pairs, we found that all the differences come from *Populus* and *Vitis*.

We also studied the conserved motifs of *LOX* genes because of its particularity and the importance to the diversified functions of *LOX* genes. Therefore, we used the MEME web server to find the relatively conserved motifs which are shared with the 50 *LOX* proteins. In total, 20 distinct conserved motifs were found (Figure 3, Supplementary Table 6), and the relevant information is shown in Supplementary Table 6. Each of the putative motifs is well commented by searching in Pfam database. In detail, motifs 1, 2, 3, 5, 6, 7, 8, 10, 11, 12, 14, 15, and 19 are associated with the Lipoxygenase domain; motif 9 is found to encode the PLAT domain; motif 4 is thought to be involved in forcing the PLAT and Lipoxygenase domain. However, the other motifs have no functional annotation. As illustrated in Figure 3, most *LOX* members belong to the same sub-family and are alike in motif compositions, suggesting that a lot of similarity may have many overlapping parts from a functional perspective. Motif 5 is widely presented in all fifty *LOX* proteins. Motif 15 and motif 20 are unique to the proteins in the 9-*LOX* clade. The former is considered for all of the components of the Lipoxygenase domain. Even though the function of motif 20 is still unknown, we still think that these motifs might be important to the functions of unique *LOX* proteins due to their specificity. To some extent, these specific motifs may help us to understand the functional divergence of *LOX* genes during evolutionary history.

Expansion Manners of *LOX* Genes within Four Rosid Plants

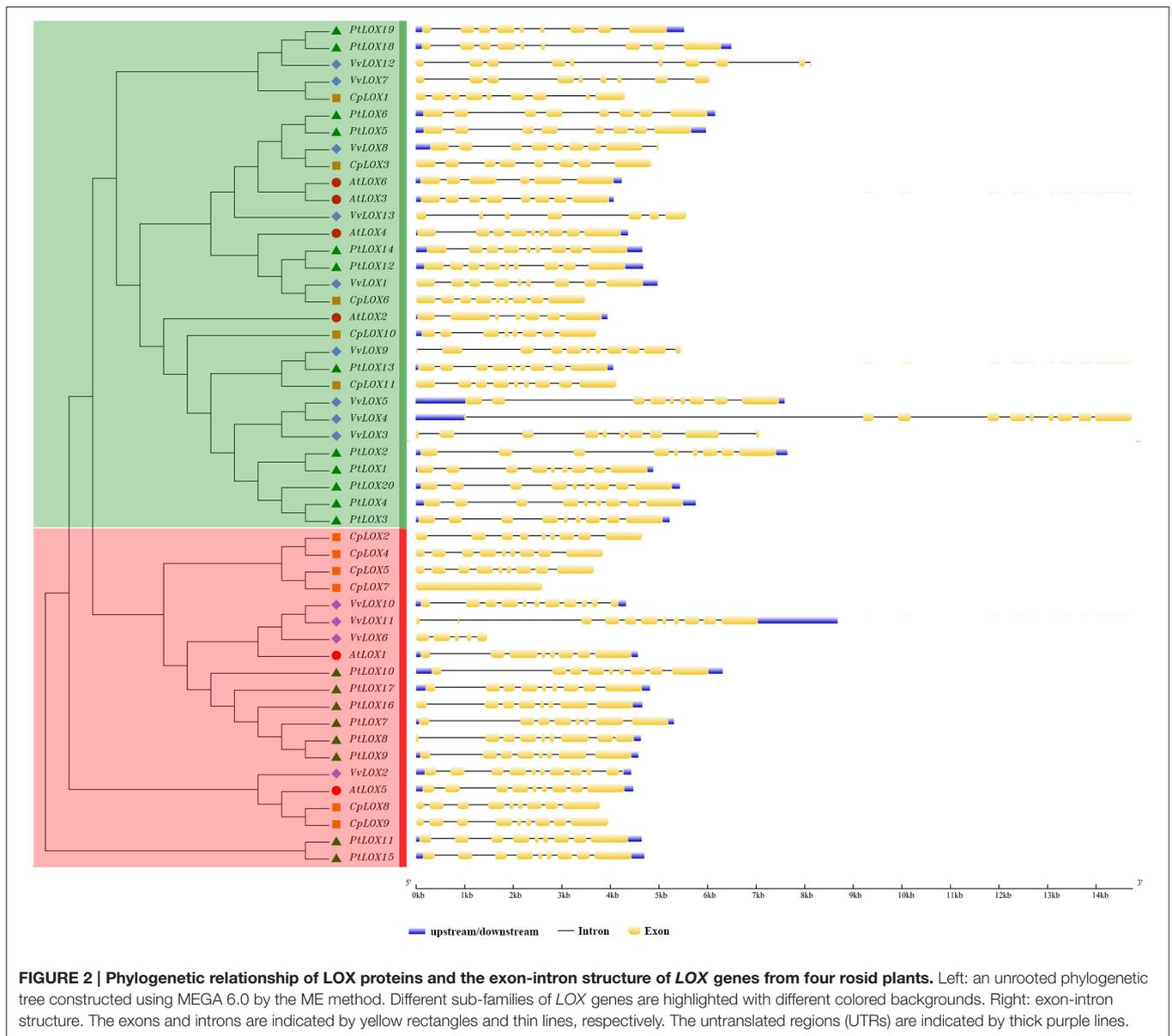
In order to probe the relationship between the genetic divergences within the *LOX* gene family and the corresponding expansion patterns, we further analyzed the gene duplication



events within each species. As previously mentioned, rosoid plants have experienced at least one polyploidy event. These events may have lasting implications for the evolution of LOX gene families. We used the MicroSyn software to investigate this possibility. If two members of the same gene family are homologous pairs, and three or more of the 50 upstream and downstream neighboring genes are also considered to be homologous pairs, we defined these two regions as those resulting from a duplication event. The number of LOX genes that arose from duplication events varied among the four rosoids. Our survey results showed that 10 collinear gene pairs occurred in the *Populus* genome and a total of four collinear gene pairs occurred in the *Vitis* genome; however, there was only one collinear gene pair in both *Arabidopsis* and *Carica* genomes (Figure 4).

In *Arabidopsis*, one gene pair (*AtLOX3/AtLOX6*) was found to have conserved neighboring regions and no syntenic relationships were detected within the other four *AtLOX* genes. In *Carica*, the microsynteny between the *CpLOX8* and *CpLOX9* genes is extensive and this gene pair is located next to each other on the same supercontig 58, believed to have derived from tandem duplication events. In *Vitis*, one pair, *VvLOX2/VvLOX8*, shares a substantial collinear region. In addition, one gene cluster *VvLOX3/VvLOX4/VvLOX5* and one gene pair, *VvLOX10/VvLOX11* are located near each other on the same chromosomes and these gene pairs might be evolved from tandem duplication. In *Populus*, three gene pairs, *PtLOX7/PtLOX16*, *PtLOX11/PtLOX15*, and *PtLOX12/PtLOX14* share extraordinary conserved synteny, with less conserved collinear genes surrounding *PtLOX5/PtLOX6*,

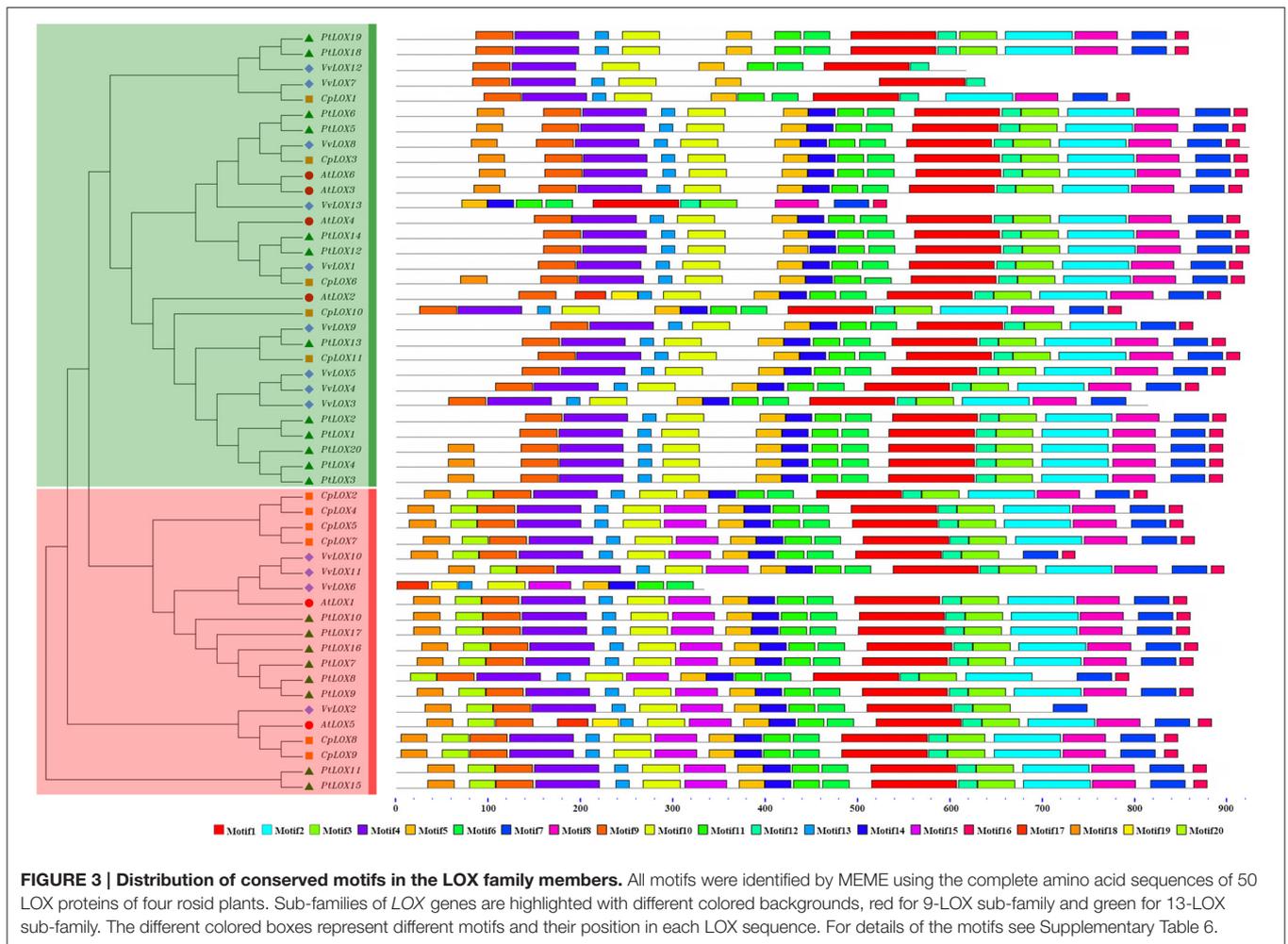
PtLOX8/PtLOX16, and *PtLOX15/PtLOX16*. Besides, one gene cluster, *PtLOX7/PtLOX8/PtLOX9/PtLOX10*, and one gene pair, *PtLOX18/PtLOX19*, appear to have evolved from tandem duplication events. Using this approach, we made a preliminary judgment on the duplication events within each species. To better examine the gene duplication events of LOX genes, we retrieved the syntenic chromosomal blocks (SCBs) associated with the expansion of LOX genes through S/WGDs from the plant genome duplication database. The results are consistent with the findings of most studies done in the MicroSyn software. The biggest difference occurs in *Vitis* and *Populus*. From the PGDD database, we found that the other two gene pairs *VvLOX2/VvLOX10* and *VvLOX3/VvLOX9* are associated with S/WGDs in *Vitis*. In *Populus*, two counterparts (*PtLOX1/PtLOXN1* and *PtLOX13/PtLOXN1*) of LOXs on SCBs are found to have sub-functionalized into other gene family members (indicated by the code name preceded by the letter “N”; Guo et al., 2014). Beyond that, the two gene pairs *PtLOX8/PtLOX16* and *PtLOX15/PtLOX16* achieved through MicroSyn software were not found in the PGDD database. Generally to consider the results from both ways, in current findings the number of LOX genes that arose from S/WGD are 10, five and two in *Populus*, *Vitis*, and *Arabidopsis*, respectively. Because the duplicated gene located on a SCB is simultaneous with another one, the median Ks value of duplicated genes in SCBs can be used to infer the dates of the large-scale duplication events. In this analysis, the duplicated gene pairs as well as the homologous genes in neighbor regions are used to date duplication events. The mean Ks values for each duplication



pair in the *LOX* genes are shown in **Figure 5** and **Table 2**. In *Populus*, the median K_s value of the γ triplication event is 1.54, and that related to the P-WGD is 0.27 (Tang et al., 2008b). We detected eight conserved gene pairs, which most likely resulted from SCB events. The median K_s of five in eight shows one range: 0.27–0.5. The median K_s values of the rest of the gene pairs is 2.1 and this pair is considered to associate with the most ancient γ -triplication event. In *Arabidopsis*, the median K_s values that have a relationship with β - and γ -WGDs are almost indistinguishable, and the K_s value is 2.00 (Tang et al., 2008b). To our knowledge the overall median K_s value for α -duplication in *Arabidopsis* is nearly 0.86. Therefore, the only one duplicated gene pair in *Arabidopsis* should be related with the α -duplication event. We also examined the expansion of *LOX* genes within the genomes of *Vitis*. According to previous reports, the overall median K_s value of SCBs in *Vitis* associated

with the γ triplication is 1.22. In our research results, the synonymous silent substitutions per site are calculated over these three possible gene pairs. The K_s values for *VvLOX2/VvLOX10*, *VvLOX2/VvLOX8*, and *VvLOX3/VvLOX9* are 1.2, 0.83, and 1.3, respectively. Based on the predicted K_s value, *VvLOX2/VvLOX10* and *VvLOX3/VvLOX9* appear to evolve from the γ triplication, while *VvLOX2/VvLOX8* evolve from a duplication event that occurred more recently.

Based on the gene-collinearity analysis within each species, we established an idealized gene tree of the duplication groups of *LOX* genes in four rosoid plants. As shown in **Figure 6**, in the *Arabidopsis* *LOX* genes duplicated network, one ancestor in the ancient genome duplication should have produced at least 12 *AtLOX* genes, but actually there is only one gene pair that is considered from α -duplication in our study. So we think that a possible ancient gene loss event occurred. In



Populus, after two rounds of duplications, one ancestor in ancient genome duplication should have produced at least six LOX genes. However, two of these lines lacked the copies, which would have been obtained from p genome duplication. Moreover, *PtLOX7*, *PtLOX8*, *PtLOX11*, *PtLOX15* and *PtLOX16* originate from the same ancestral gene. *PtLOX1* and *PtLOXN1* evolve from a prior duplication event, while *PtLOXN1* and *PtLOX13* result from a duplication event that occurred more recently. In addition, two LOX gene pairs could be matched to the γ triplication in *Vitis*. In contrast, there is no LOX-containing segments in *Carica* being matched in any duplicated pairs. Such a huge difference existed in the expansion manners of LOX gene within the four rosoid plants, so where did the remaining LOX genes in these species originate from?

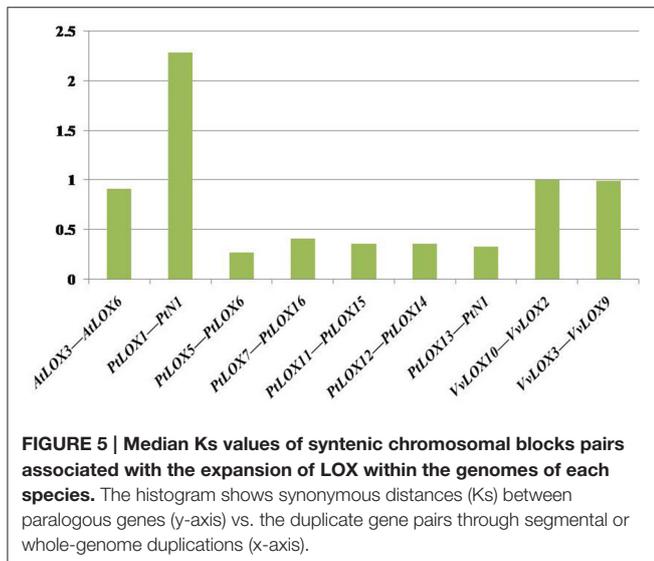
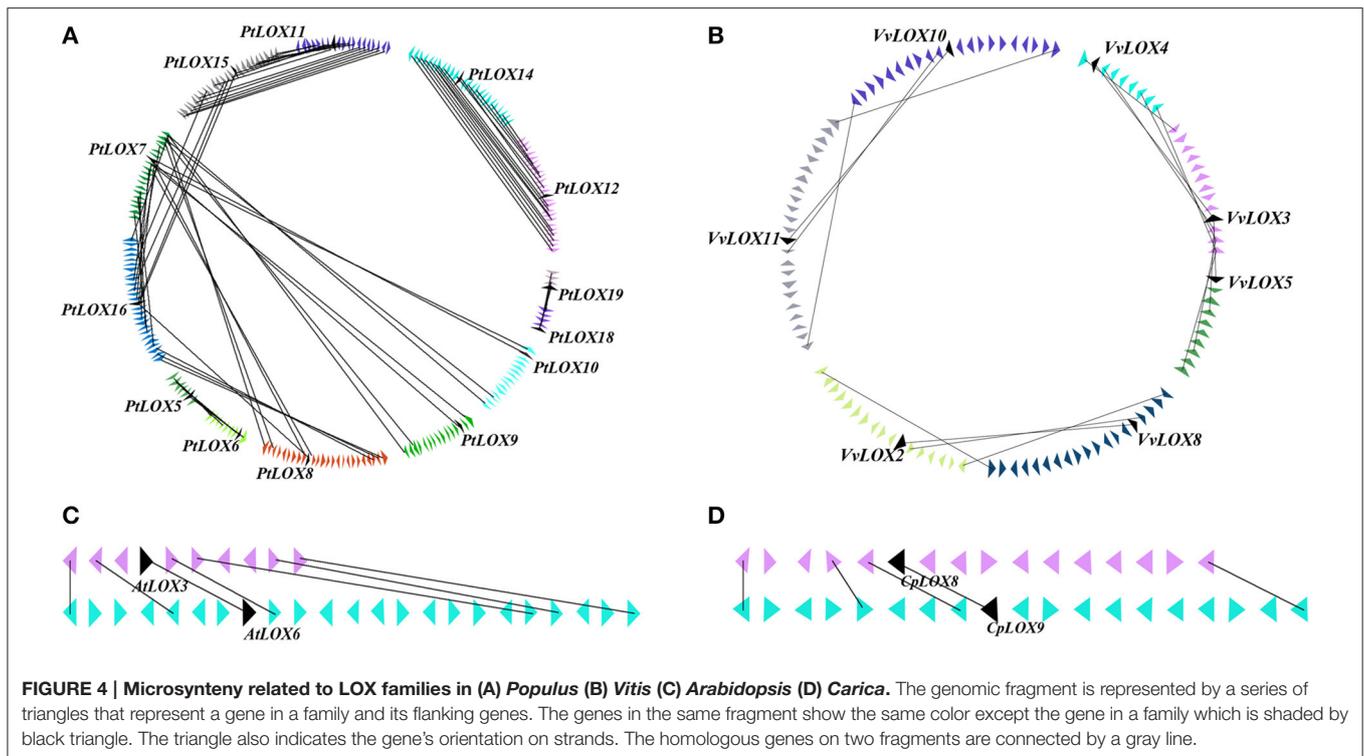
Evolutionary History of LOX Gene Families in Four Rosid Plants

We used the LOX gene family members as anchor genes to further examine the orthologous relationships and evolutionary history of LOX genes among four rosoid plants. After this interspecies microsynteny analysis, the relationships between syntenic orthologs of LOX genes in four rosoid plants are displayed

in **Figure 7**, indicating that the strongly conserved microsynteny among these regions across four species is observed significantly.

We obtained the collinear correlations of LOX genes in the four plant genomes by using MicroSyn. In total, 33 conserved syntenic segments were found (Supplementary Figure 2), and these syntenic segments are divided into six groups. Four of the groups contains all the four species with the LOX gene. *AtLOX1* in *Arabidopsis*, *PtLOX7/8/16/17* in *Populus*, *VvLOX11* in *Vitis*, and *CpLOX7* in *Carica* have conserved collinearity, and are identified as group “A.” *AtLOX3/6* in *Arabidopsis*, *PtLOX5/6* in *Populus*, *VvLOX8* in *Vitis*, and *CpLOX3* in *Carica* are classified into the group “B.” *AtLOX4* in *Arabidopsis*, *PtLOX12/14* in *Populus*, *VvLOX1* in *Vitis* and *CpLOX6* in *Carica* are grouped as group “C.” *AtLOX5* in *Arabidopsis*, *PtLOX11/15* in *Populus*, *VvLOX2* in *Vitis* and *CpLOX8/9* in *Carica* are grouped as group “D.” Group “E” consists of LOX genes from three species, which are *PtLOX9* in *Populus*, *VvLOX12* in *Vitis*, and *CpLOX1* in *Carica*. The LOX genes from two species comprise the group “F” as they are *PtLOX1* in *Populus* and *VvLOX3/4/5* in *Vitis*. The results are consistent with the findings of the phylogenetic analysis.

Subsequently, the synteny quality was calculated in four rosoid plants. The quality was calculated as twice the number of



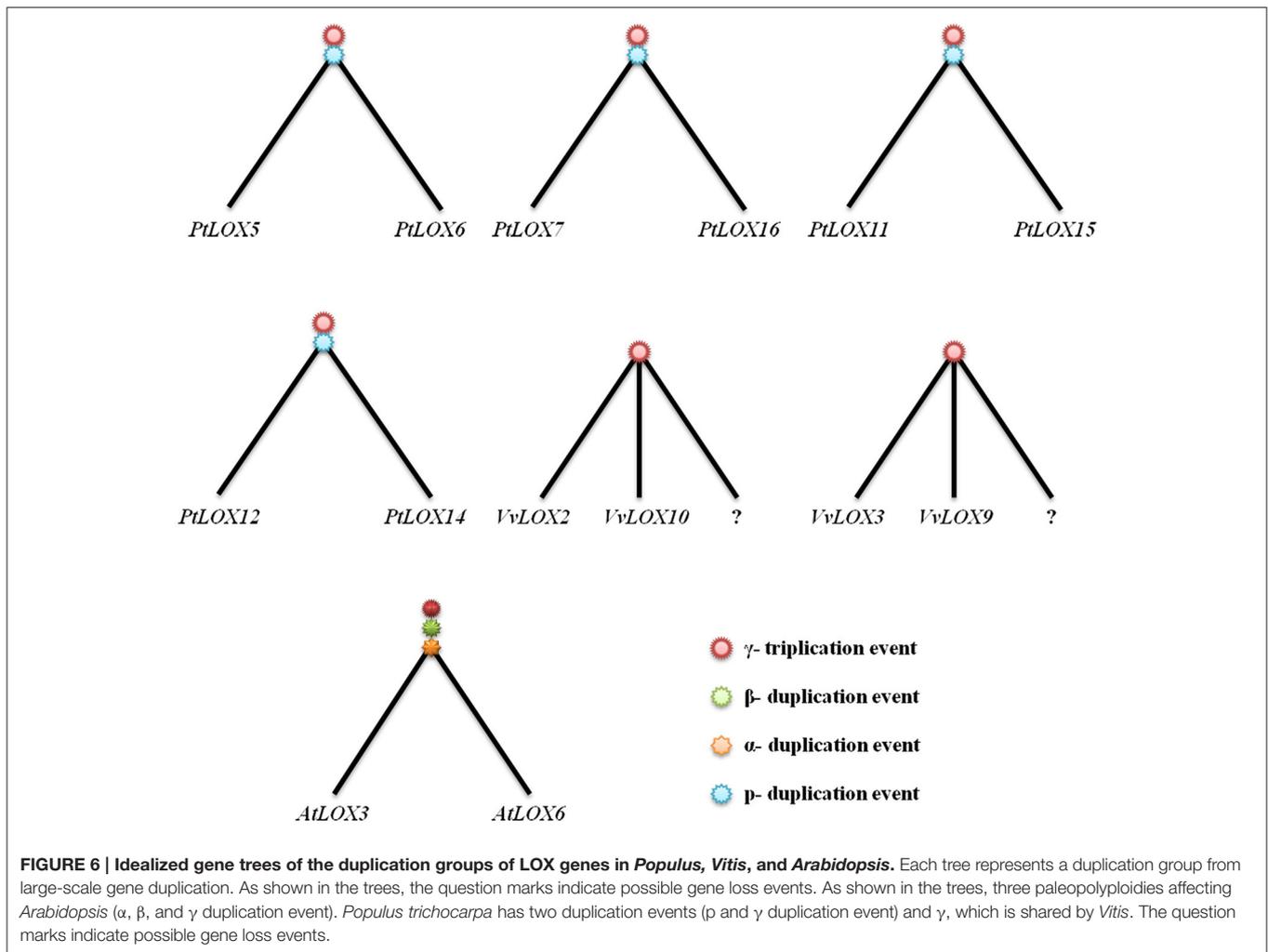
matches divided by the total number of genes in both segments (Cannon et al., 2006). These four species have a synteny quality of 67.77% for orthologous regions. The minimum value of synteny quality observed between *Arabidopsis* and *Vitis* was 48.97%, and the maximum value was 97.70%. The average synteny quality in the *Carica/Populus* syntenic regions reached over 89.24%, followed by *Carica/Vitis*, for which the average synteny quality was 76%. The average synteny quality in the *Arabidopsis/Populus* and *Arabidopsis/Carica* syntenic regions was 53.34 and 45.35%,

TABLE 2 | Median Ks values of SCB pairs associated with the expansion of LOXs within the genomes of each species.

Species	Locus_1 gene code	Locus_2 gene code	Ka	Ks	Block median Ks
<i>Arabidopsis thaliana</i>	AtLOX3	AtLOX6	0.10	0.98	0.91
<i>Populus trichocarpa</i>	PtLOX1	PtN1	0.21	2.14	2.28
	PtLOX5	PtLOX6	0.07	0.32	0.27
	PtLOX7	PtLOX16	0.05	0.26	0.41
	PtLOX11	PtLOX15	0.04	0.21	0.36
	PtLOX12	PtLOX14	0.06	0.25	0.36
	PtLOX13	PtN1	0.13	0.48	0.33
	<i>Vitis vinifera</i>	VvLOX10	VvLOX2	0.29	1.21
	VvLOX3	VvLOX9	0.30	1.30	0.99

respectively. Details of this comparative analysis are shown in **Table 3**.

It is thought that LOX families evolved from a process of different duplication events. However, within those LOX homologs, what kind of role is the genome-wide duplication playing? Since previous studies, the expansion of LOX genes within the genome of each species have been researched by using the PGDD database. Similarly, from the database, we also examined the SCBs associated with the expansion of LOXs between species. To better understand the gene-collinearity between species, a panoramic picture about the differential retention and evolution of the ancestral LOXs related to paleopolyploidy in the four rosoid plants was built (**Figure 8**, **Table 2**). The study, building on previous research, has identified



differences in the duplicates of these ancestral genes through S/WGDs in each species. Five, one, eight and four LOX genes have been linked to paleopolyploidy in *Vitis*, *Carica*, *Populus*, and *Arabidopsis*, respectively.

In theory, it should always be possible to find out if the microsynteny were maintained among the members of four rosoid plants. But this was not what we found. As shown in **Figure 8**, in the process of analyzing the gene-collinearity in individual species, we found out that one gene pair *AtLOX3/AtLOX6* in *Arabidopsis* originated from α -WGD. When we expanded to analyze the gene-collinearity between species, we found that *AtLOX3* and *AtLOX6* were all orthologous SCBs of the chromosomal block containing *PtLOX5* and *PtLOX6* in *Populus*. In addition, in *Carica*, one gene, *CpN1* on SCBs, was also found to have collinearity with *AtLOX3/AtLOX6* and *PtLOX5/PtLOX6*. This may be because LOX gene sub-functionalized into other gene family members over evolutionary time. A similar dynamic can be seen in the gene-collinearity analysis between *PtLOX12* and *PtLOX14*. The SCB analysis revealed that this gene pair is related to *AtLOX4* and *CpN2*. In addition, based on the results of the present study, *VvLOX2*, *VvLOX10*, *PtLOX11*, *PtLOX15*, and

AtLOX5 genes might have originated from the same ancestral gene. The orthologous *CpLOX6* genes that might have arose through S/WGDs were sub-functionalized into *VvN1* and *PtN2* in *Vitis* and *Populus*, respectively. Beyond that, the orthologous pair *CpN3* and *VvLOX12* were found to have originated from the same ancestral gene. Besides, the orthologous relationship of *VvLOX3/9* and *PtLOX1/13/N1* are no longer traceable to the LOX genes in the other three species. Furthermore, our opinion is that LOX genes which are unique to their species might represent the oldest relics of ancient LOXs differentially retained in each species.

DISCUSSION

In terms of evolution, *A. thaliana*, *P. trichocarpa*, *C. papaya*, and *V. vinifera* are believed to originate from a common paleohexaploid ancestor demonstrated by numerous studies (Ohno, 2013). Single and more recent multiple WGD events have been found in the genomes of *Populus* and *Arabidopsis*. Paleopolyploidy events provided opportunities for gene

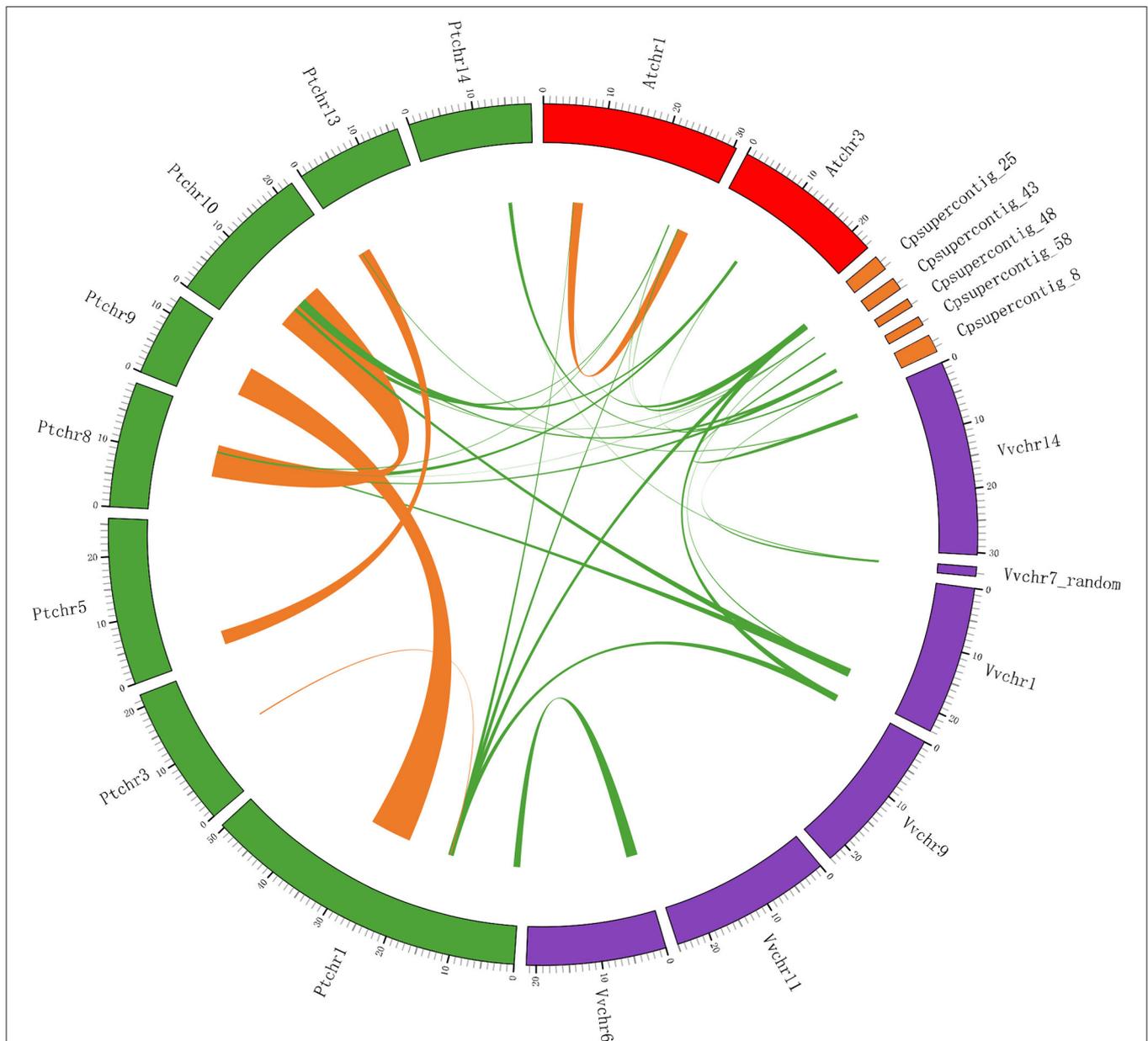


FIGURE 7 | Extensive microsynteny of LOX regions across *Populus*, *Vitis*, *Arabidopsis* and *Carica* chromosomes. The *Populus*, *Vitis*, *Arabidopsis*, and *Carica* chromosomes, shown in different colors, are labeled Pt, Vv, At, and Cp, respectively. Numbers along each chromosome box indicate sequence lengths in megabases. The whole chromosomes of these four rosoid plants, harboring LOX regions, are shown in a circle. Green lines represent the syntenic relationships between LOX regions within species. Orange lines represent the syntenic relationships between LOX regions between species.

duplication, and those duplicated genes have been shown to act as an important role in evolutionary innovation (Hittinger and Carroll, 2007). Functional diversification with duplicated genes results in more complex organisms. Typically, ancient genome duplications have always been thought to be a powerful source of functional innovation and genome complexity, and is also followed by substantial gene loss. Lipoxygenase and its products are involved in the regulation of a variety of processes. In this paper, we used the model rosoid plant *Arabidopsis*, as well as

Carica, *Populus*, and *Vitis* to study the evolutionary history of this gene family.

In this study, we identified 6, 20, 13, and 11 LOXs in *Arabidopsis*, *Populus*, *Vitis*, and *Carica*, respectively. Except *Arabidopsis*, previous surveys indicated that a very high proportion of most LOX members in other three species are paralogs. In order to improve our understanding on what affects the evolutionary constraints, we measured the Ka/Ks ratios of paralogous pairs of the four species. Amidst all of the pairwise

TABLE 3 | The synteny quality of regions orthologous across four modern rosids.

	<i>Arabidopsis</i>	<i>Carica</i>	<i>Populus</i>	<i>Vitis</i>
ARABIDOPSIS				
<i>Carica</i>	45.35%			
<i>Populus</i>	53.34%	89.24%		
<i>Vitis</i>	45%	76%	97.70%	

comparison data, only one gene pair, *VvLOX6/11* exhibits a Ka/Ks ratio larger than 1, suggesting that accelerated devolution with positive selection occurred in this gene pair. Other than that, Ka/Ks ratios of all the other paralog pairs are lower than 1, indicating that the *LOX* genes at the protein level are very slow-changing and the majority of sites are often controlled by strong purifying selection.

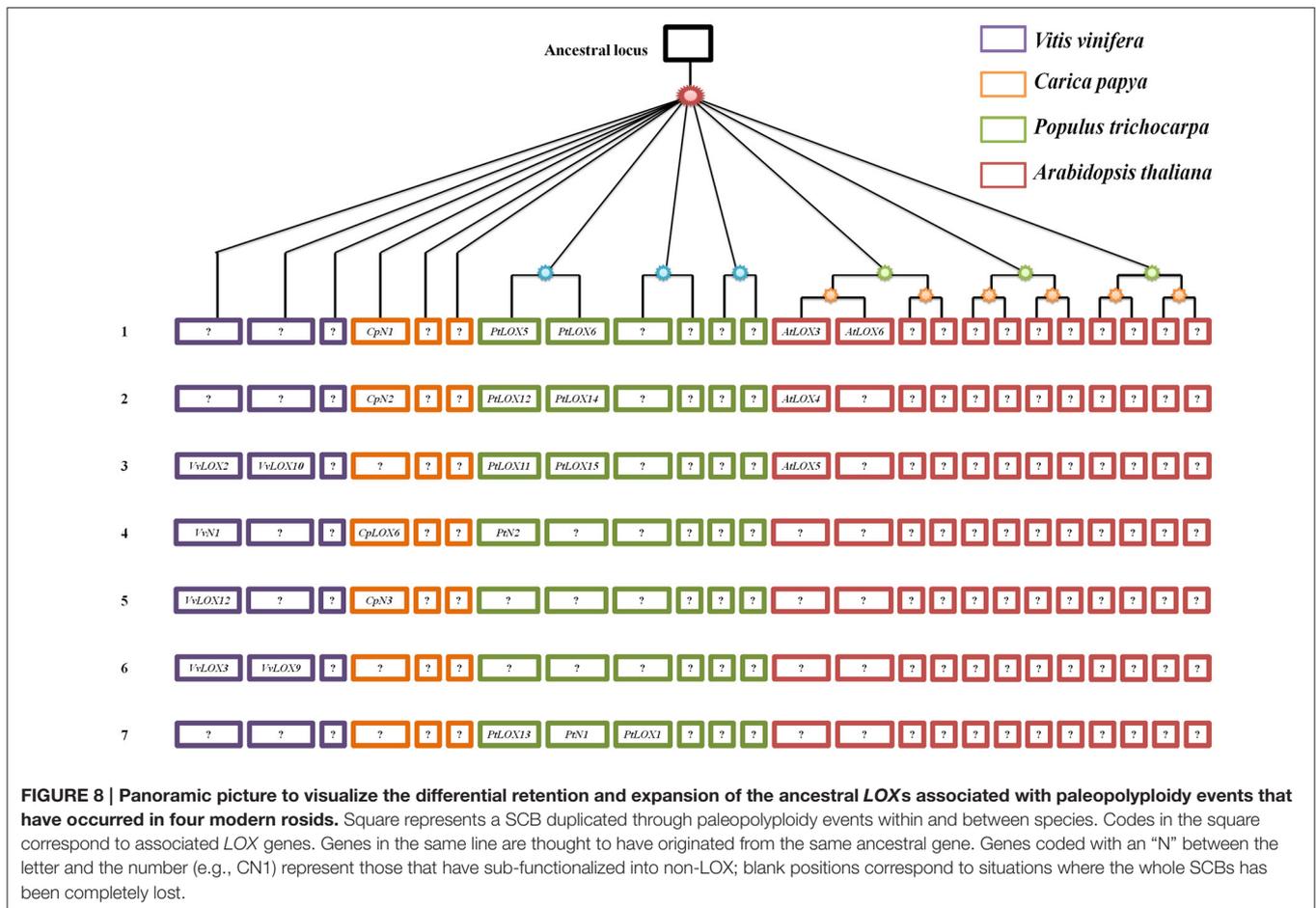
Phylogenetic trees are quite informative for obtaining the *LOX* gene relationships with each other. In this study, the *LOX* genes are divided into two groups, 13-*LOX* and 9-*LOX*, consistent with many previous studies. The calculated genetic distances among the two *LOX* subfamilies were computed and the results show that the *LOX* genes of 9-*LOX* sub-families appear to be more closely related to each other than those *LOX* genes in 13-*LOX* sub-families. The number of exons in *Carica* and *Populus* *LOX* genes is relatively stable, whereas the exon numbers has changed dramatically in *Vitis* and *Arabidopsis*. *Vitis* have the most number of exons with 11 and the least number of exons with five. Exon-intron structural diversification has been confirmed in the evolution of many gene families, and the reason why exon-intron gain or loss occurs is because of the genetic assortment of different chromosome fragments. The MEME server identifies that each sub-family shares a similar motif, and the results could have implications for functional similarities about these *LOX* proteins (Paterson et al., 2006). The differential motifs in each sub-family may endow the *LOX* proteins with new functions or to raise their performance. Our study shows that the results meet the similarities in gene structure and motif composition of most *LOX* proteins from phylogenetic analysis of the *LOX* gene family. *LOX* genes differentiate into various characteristics among the different sub-families for a variety of possible reasons, and the most probable cause is that the *LOX* members were functionally diversified (Blanc and Wolfe, 2004).

The current tools related to investigate the relationship among genes in modern plants include sequence similarity, microsynteny analysis, and retrieve the syntenic chromosomal blocks from PGDD. As might be expected each of these approaches has advantages and shortcoming in certain situations. For example, some ancient duplicates could not be retrieved preferences from the Blast method because sequence similarity may have severely eroded in long evolutionary process. As a result, we could no longer track the paralog-ship for such duplicates based on this method. For instance, our results showed that *VvLOX3* and *VvLOX9* are duplicates resulting from γ -WGD in *Vitis* (Figure 8), but this gene pairs in the paralog analysis based on sequence similarity that remained undetected (Supplementary Table 5). Thanks to the plant genome

duplication database, more duplicates which had arisen from segmental or whole-genome duplications could be discovered easily. But this approach had a number of drawbacks, such as we could not detect the homology pairs that proliferate via other duplication strategies. The large paralogous group unique to only one rosid plant were found abundant in our results (Supplementary Table 5). It is unscientific to infer evolutionary relationships for homologous genes among different lineages reducing only based on Microsynteny analysis. Microsynteny between two members of a gene family is calculated from their flanking genes. And we have already known that the quality of gene prediction in different genome sequencing programs is drastically different. If the flanking regions contain assembly errors, gaps or annotation errors, would cause the microsynteny that be artificial. In conclusion, to approach the evolutionary relationships for members of *LOX* gene family across four modern rosids, the above methods should be utilized compositely.

Based on the gene collinearity on syntenic chromosomal blocks within and between these four rosid plants, we set up a panoramic picture to trace the evolutionary history of *LOX* gene families (Figure 8). Our analysis shows that five (line 1–5 in Figure 8) of these ancestral *LOXs* are retained in more than two species. By contrast, two of the ancestral *LOX* genes were retained in only one of the four rosids. These observations suggests that, no matter which rosid plant is used as the model plant, the functions of a gene family inevitably get the limited amount. We could speculate genes uniquely retained in only one species may have a specific and indispensable function. As it turns out, this finding will help us dig deeper into the unique genes retained in the rosid plants and further research the function of these genes. A surprising finding of this study is that we have not found any line containing the *LOX* genes in all four modern rosids. Thus, we determined that all of the *LOXs* in each modern rosid are offspring coming from different ancestral genes.

To date, top-down analysis shows a high degree of collinearity between the four studied rosids. *Arabidopsis* (Lamesch et al., 2012), *Carica* (Ming et al., 2008), *Vitis* (Jaillon et al., 2007), and *Populus* (Tuskan et al., 2006) have been suggested to possess paleohexaploidy in a common ancestor (Jaillon et al., 2007). Previous results indicate that genome triplication (γ) occurred in a common ancestor of *Vitis*, *Arabidopsis*, *Carica*, and *Populus*. Meanwhile, the previous results show the two most recent paleopolyploidies affecting *Arabidopsis* that are often described as α and β duplication. *Populus trichocarpa* has had a unique duplication event in recent times, which is called salicoid lineage (p , following the usage in Tang et al., 2008a). Considering the paleopolyploidy events that occurred in each species, there should be 3 ancestral loci in *Carica* and *Vitis*, 6 ancestral loci in *Populus*, and 12 in *Arabidopsis*. Based on these results, the multiplicity ratio for an ancestral gene comparison in the genomes of four species should be 4:2:1:1. And in fact none of the ancestral *LOXs* included the extremes of each condition. However, collinear correlations of *LOX* genes in the four plant genomes have been obtained by using MicroSyn provides an interesting point. In our work, 33



conserved syntenic segments are divided into six groups, and in almost every group, LOX genes from *Populus* are present in at least an extra copy compared with *Vitis* and *Carica*, and the two copies are paralogs. The result accords with the well-documented fact and also provides powerful evidence that *Populus* has undergone an additional whole genome duplication, which is not shared with *Vitis* and *Carica*. However, our survey results show that there are not twice as many LOX genes in *Populus* vs. *Vitis*, suggesting them could have suffered differential gene loss events could have happened to these two species. The LOX gene family has shrunk in the herbaceous plants and retained a large number of LOX genes in the woody plants, leading to the hypothesized that some *Arabidopsis* LOX genes might have been lost during the evolutionary process due to functional redundancy. Previous studies showed that after the paleopolyploidy events, the exponential growth in gene numbers is often tempered by massive and progressive gene death in the subsequent diploidization process (Tang et al., 2008a). Another possibility is that LOXs may have expanded faster in the other three species than in *Arabidopsis*. This expansion to more abundant LOX genes in *Populus*, *Vitis*, and *Carica* genomes suggests a great need of LOX genes to participate in more complicated physiological and biological processes in these three woody species. These results probably suggest

the complex evolutionary history of the LOX family in rosid plants.

In this study, there are eight pairs of genes associated with S/WGDs, including six pairs from the PtLOX gene family, along with two pairs that have sub-functionalized into other gene family members. In contrast, there is only one collinear gene pair in both *Arabidopsis*, which is mainly caused by the rapid substitutions in *Arabidopsis*. The γ -triplication event has ever happened in the common ancestor of these four species. But all of them have different values median Ks about γ -paleologs from four modern rosids. In *Arabidopsis* the median Ks is close to 2.0, which was higher than that in *Populus* (1.54), *Carica* (1.76), and *Vitis* (1.22) (Tang et al., 2008a). Some studies have shown rapid substitutions at a rate proportional to the amount of synonymous sites (Guo et al., 2014). Over millions of years of evolution in *Arabidopsis* extensive chromosome have been actively rearranged. That might contribute to the high median Ks between γ -paleologs in *Arabidopsis* and this would destroy collinearity. The result accords with the fact that *Arabidopsis* which contains more paleopolyploidies has a smaller genome than that of *Populus*, though both originated from a common ancestor. Besides, our analyses showed that almost all *Vitis* contains many more paralog-ship LOX genes than *Carica*, although both of them were affected by the γ -WGD event. The

observations suggesting that the gene duplication impacts turn out to be small in CpLOX gene family. The dates of the large-scale duplication events have been obtained through calculating the median Ks value of duplicated genes in SCBs. In *Populus*, the number of LOX genes produced by the recent duplication event is much more than those produced from ancient duplication events. In *Arabidopsis*, the only unique gene pair is generated by α duplication event, which is also a recent duplication event. This illustrates that recent duplication host to those species which have undergone at least whole-genome duplication.

The current study provides an overview of LOX genes in four rosid plants, including their phylogenetic relationship, gene structure, conserved motifs, microsynteny and gene collinearity. Based on these findings, we tracked the evolutionary history of ancestral LOX genes among four modern rosids. The results suggest that all of the LOX genes in each species could have resulted from different ancestral genes. This study presented here may provide clues for exploring the unique genes retained in the rosid plants and aid in the research of the biological functions about these special genes.

REFERENCES

- Adams, K. L., and Wendel, J. F. (2005). Polyploidy and genome evolution in plants. *Curr. Opin. Plant Biol.* 8, 135–141. doi: 10.1016/j.pbi.2005.01.001
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Bailey, T. L., Williams, N., Misleh, C., and Li, W. W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* 34, W369–W373. doi: 10.1093/nar/gkl198
- Bannenberg, G., Martínez, M., Hamberg, M., and Castresana, C. (2009). Diversity of the enzymatic activity in the lipoxygenase gene family of *Arabidopsis thaliana*. *Lipids* 44, 85–95. doi: 10.1007/s11745-008-3245-7
- Barker, M. S., Vogel, H., and Schranz, M. E. (2009). Paleopolyploidy in the Brassicales: analyses of the Cleome transcriptome elucidate the history of genome duplications in *Arabidopsis* and other Brassicales. *Genome Biol. Evol.* 1, 391–399. doi: 10.1093/gbe/evp040
- Baysal, T., and Demirdöven, A. (2007). Lipoxygenase in fruits and vegetables: a review. *Enzyme Microb. Technol.* 40, 491–496. doi: 10.1016/j.enzmictec.2006.11.025
- Blanc, G., and Wolfe, K. H. (2004). Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell* 16, 1667–1678. doi: 10.1105/tpc.021345
- Bowers, J. E., Chapman, B. A., Rong, J., and Paterson, A. H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 422, 433–438. doi: 10.1038/nature01521
- Boyington, J. C., Gaffney, B. J., and Amzel, L. M. (1993). The three-dimensional structure of an arachidonic acid 15-lipoxygenase. *Science* 260, 1482–1486. doi: 10.1126/science.8502991
- Brash, A. R. (1999). Lipoxygenases: occurrence, functions, catalysis, and acquisition of substrate. *J. Biol. Chem.* 274, 23679–23682. doi: 10.1074/jbc.274.34.23679
- Brash, A. R. (2015). “Lipoxygenases: a Chronological Perspective on the Synthesis of S and R Fatty Acid Hydroperoxides,” in *Bioactive Lipid Mediators*, ed T. Y. M. Murakami (Japan: Springer), 69–84.
- Cai, B., Yang, X., Tuskan, G. A., and Cheng, Z.-M. (2011). MicroSyn: a user friendly tool for detection of microsynteny in a gene family. *BMC Bioinformatics* 12:1. doi: 10.1186/1471-2105-12-79
- Cannon, S. B., Sterck, L., Rombauts, S., Sato, S., Cheung, F., Gouzy, J., et al. (2006). Legume genome evolution viewed through the *Medicago truncatula* and

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: ZC, DC, and YX. Performed the experiments: ZC, DZ. Analyzed the data: ZC, WC. Wrote the paper: ZC, WC, and HY. Participated in the design of this study and revised manuscript: ZC, DC, WC.

FUNDING

Sub-project I under National Science and Technology Support Program (2015BAD07B070104). Anhui provincial Natural Science Foundation (1608085QC65). We thank the members of the Laboratory of Modern Biotechnology for their assistance in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fgene.2016.00176>

- Lotus japonicus* genomes. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14959–14964. doi: 10.1073/pnas.0603228103
- Chen, Z., Chen, X., Yan, H., Li, W., Li, Y., Cai, R., et al. (2015). The Lipoxygenase gene family in Poplar: identification, classification, and expression in response to MeJA treatment. *PLoS ONE* 10:e0125526. doi: 10.1371/journal.pone.0125526
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Feussner, I., and Kühn, H. (2000). 15 Application of Lipoxygenases and related enzymes for the preparation of oxygenated lipids. *Enzymes Lipid Modification* 40:309. doi: 10.1002/3527606033.ch15
- Feussner, I., and Wasternack, C. (2002). The lipoxygenase pathway. *Annu. Rev. Plant Biol.* 53, 275–297. doi: 10.1146/annurev.arplant.53.100301.135248
- Finn, R. D., Mistry, J., Schuster-Böckler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., et al. (2006). Pfam: clans, web tools and services. *Nucleic Acids Res.* 34, D247–D251. doi: 10.1093/nar/gkj149
- Guo, A., Zhu, Q., Chen, X., and Luo, J. (2007). [GSDS: a gene structure display server]. *Yi Chuan* 29, 1023–1026. doi: 10.1360/yc-007-1023
- Guo, L., Chen, Y., Ye, N., Dai, X., Yang, W., and Yin, T. (2014). Differential retention and expansion of the ancestral genes associated with the paleopolyploidies in modern rosid plants, as revealed by analysis of the extensins super-gene family. *BMC Genomics* 15:612. doi: 10.1186/1471-2164-15-612
- Hildebrand, D. F. (1989). Lipoxygenases. *Physiol. Plant.* 76, 249–253. doi: 10.1111/j.1399-3054.1989.tb05641.x
- Hittinger, C. T., and Carroll, S. B. (2007). Gene duplication and the adaptive evolution of a classic genetic switch. *Nature* 449, 677–681. doi: 10.1038/nature06151
- Hufton, A. L., and Panopoulou, G. (2009). Polyploidy and genome restructuring: a variety of outcomes. *Curr. Opin. Genet. Dev.* 19, 600–606. doi: 10.1016/j.gde.2009.10.005
- Jaillon, O., Aury, J.-M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., et al. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449, 463–467. doi: 10.1038/nature06148
- Jaillon, O., Aury, J.-M., and Wincker, P. (2009). “Changing by doubling,” the impact of Whole Genome Duplications in the evolution of eukaryotes. *C. R. Biol.* 332, 241–253. doi: 10.1016/j.crv.2008.07.007
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109

- Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., et al. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* 40, D1202–D1210. doi: 10.1093/nar/gkr1090
- Larkin, M. A., Blackshields, G., Brown, N., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal, W., and Clustal X Version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404
- Lee, T.-H., Tang, H., Wang, X., and Paterson, A. H. (2013). PGDD: a database of gene and genome duplication in plants. *Nucleic Acids Res.* 41, D1152–D1158. doi: 10.1093/nar/gks1104
- Leenhardt, F., Lyan, B., Rock, E., Boussard, A., Potus, J., Chanliaud, E., et al. (2006). Genetic variability of carotenoid concentration, and lipoxygenase and peroxidase activities among cultivated wheat species and bread wheat varieties. *Eur. J. Agronomy* 25, 170–176. doi: 10.1016/j.eja.2006.04.010
- Ming, R., Hou, S., Feng, Y., Yu, Q., Dionne-Laporte, A., Saw, J. H., et al. (2008). The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature* 452, 991–996. doi: 10.1038/nature06856
- Mulder, N. J., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., D224–D228. doi: 10.1093/nar/gkl841
- Ohno, S. (2013). *Evolution by Gene Duplication*. Berlin; Heidelberg: Springer Science & Business Media.
- Paterson, A. H., Chapman, B. A., Kissinger, J. C., Bowers, J. E., Feltus, F. A., and Estill, J. C. (2006). Many gene and domain families have convergent fates following independent whole-genome duplication events in Arabidopsis, Oryza, Saccharomyces and Tetraodon. *Trends Genet.* 22, 597–602. doi: 10.1016/j.tig.2006.09.003
- Podolyan, A., White, J., Jordan, B., and Winefield, C. (2010). Identification of the lipoxygenase gene family from *Vitis vinifera* and biochemical characterisation of two 13-lipoxygenases expressed in grape berries of Sauvignon Blanc. *Funct. Plant Biol.* 37, 767–784. doi: 10.1071/FP09271
- Rokas, A., and Holland, P. W. (2000). Rare genomic changes as a tool for phylogenetics. *Trends Ecol. Evol.* 15, 454–459. doi: 10.1016/S0169-5347(00)01967-4
- Sémon, M., and Wolfe, K. H. (2007). Consequences of genome duplication. *Curr. Opin. Genet. Dev.* 17, 505–512. doi: 10.1016/j.gde.2007.09.007
- Shibata, D., Steczko, J., Dixon, J., Hermodson, M., Yazdanparast, R., and Axelrod, B. (1987). Primary structure of soybean lipoxygenase-1. *J. Biol. Chem.* 262, 10080–10085.
- Steczko, J., Donoho, G. P., Clemens, J. C., Dixon, J. E., and Axelrod, B. (1992). Conserved histidine residues in soybean lipoxygenase: functional consequences of their replacement. *Biochemistry* 31, 4053–4057. doi: 10.1021/bi00131a022
- Tamura, K., Stecher, G., Peterson, D., Filipiński, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Tang, H., Bowers, J. E., Wang, X., Ming, R., Alam, M., and Paterson, A. H. (2008a). Synteny and collinearity in plant genomes. *Science* 320, 486–488. doi: 10.1126/science.1153917
- Tang, H., Wang, X., Bowers, J. E., Ming, R., Alam, M., and Paterson, A. H. (2008b). Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps. *Genome Res.* 18, 1944–1954. doi: 10.1101/gr.080978.108
- Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604. doi: 10.1126/science.1128691
- Umate, P. (2014). Genome-wide analysis of lipoxygenase gene family in Arabidopsis and rice. *Plant Signal. Behav.* 6, 335–338. doi: 10.4161/psb.6.3.13546
- Van Loon, L. C., Rep, M., and Pieterse, C. (2006). Significance of inducible defense-related proteins in infected plants. *Annu. Rev. Phytopathol.* 44, 135–162. doi: 10.1146/annurev.phyto.44.070505.143425
- Vogt, J., Schiller, D., Ulrich, D., Schwab, W., and Dunemann, F. (2013). Identification of lipoxygenase (LOX) genes putatively involved in fruit flavour formation in apple (*Malus domestica*). *Tree Genet. Genomes* 9, 1493–1511. doi: 10.1007/s11295-013-0653-5
- Wang, Y., Tang, H., Debary, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40, e49–e49. doi: 10.1093/nar/gkr1293
- Wang, D., Zhang, Y., Zhang, Z., Zhu, J., and Yu, J. (2010). KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genom. Proteom. Bioinform.* 8, 77–80. doi: 10.1016/S1672-0229(10)60008-3
- Yang, Z., and Bielawski, J. P. (2000). Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* 15, 496–503. doi: 10.1016/S0169-5347(00)01994-7

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Chen, Chen, Chu, Zhu, Yan and Xiang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.