New Phytologist Supporting Information

Article title: Divergence of duplicated genes by repeated partitioning of splice forms and subcellular localization.

Authors: Yichun Qiu, Yii Van Tay, Yuan Ruan, Keith L. Adams

Article acceptance date: 21 August 2019

The following Supporting Information is available for this article:

Fig. S1. Scheme for primer design.

Fig. S2. The translated sequences of the last exon of alternatively spliced and specialized tAPX showing the conserved hydrophobic thylakoid-anchorage chain.

Fig. S3. Case studies of sub-localized duplicated *cpAPX* genes.

Fig. S4. The evolution of *cpAPX* genes in Solanaceae.

Table S1. CpAPX genes examined in this study (submitted separately).

Table S2. Plant materials in this study (submitted separately).

Table S3. Primers designed for this study (submitted separately).

Table S4. The selection analyses of cpAPX genes (submitted separately).

Fig. S1. Scheme for primer design. Alternatively-spliced single copy *cpAPX*, specialized *sAPX* (Type I and II) and specialized *tAPX* are aligned with only the last three exons shown by boxes. Blue and green squares are splicing acceptors corresponding to sAPX and tAPX transcripts, respectively. Red bars are stop codons. Striped bars represent the hydrophobic tail coding sequences. The forward (F) and reverse (R) primers (arrows) are designed outside of the orange bars to distinguish the coding potential for tAPX and sAPX.

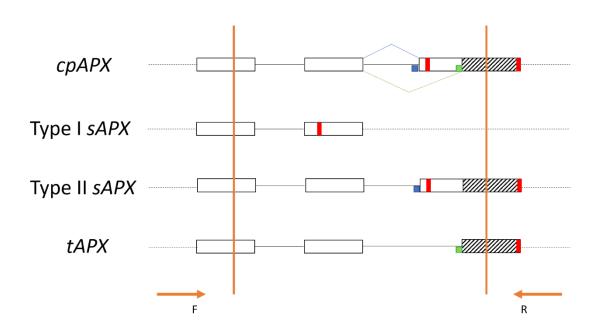


Fig. S2. The translated sequences of the last exon of alternatively spliced (blue) and specialized (red) tAPX showing the conserved hydrophobic thylakoid-anchorage chain. The green triangle points to the spinach tAPX which was purified from the thylakoidal membrane by Miyake and Asada (1992).

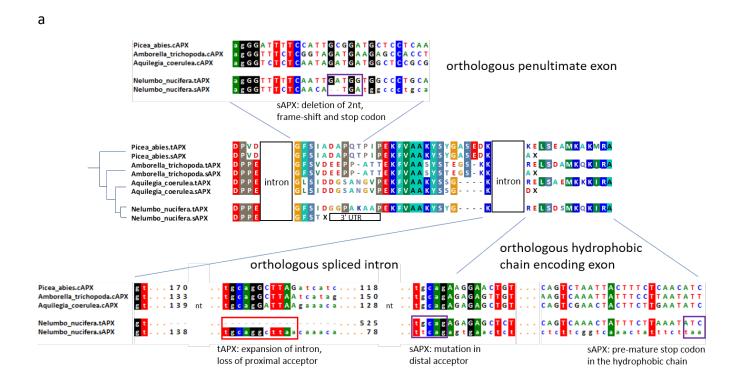
AS

Phalaenopsis equestris.tAP) Dendrobium_catenatum.tAP) Aquilegia_coerulea.tAPX Macleaya_cordata.tAPX Beta_vulgaris.tAPX Spinacia_oleracea.tAPX Lactuca_sativa.tAPX Daucus_carota.tAPX Coffea_canephora.tAPX Mimulus_guttatus.tAPX Mentha_longifolia.tAPX Mentha_piperita.tAPX Utricularia_gibba.tAPX Genlisea aurea.tAPX Ipomoea_trifida.tAPX Petunia_axillaris.tAPX Nicotiana_sylvestris.tAPX Solanum_tuberosum.tAPX Capsicum annuum.tAPX Kalanchoe_marnieriana.tAPX Cephalotus_follicularis.tAPX Jatropha_curcas.tAPX Linum_usitatissimum.tAPX1 Linum_usitatissimum.tAPX2 Medicago_truncatula.tAPX Phaseolus vulgaris.tAPX Glycine_max.tAPX1 Glycine_max.tAPX2 Citrullus_lanatus.tAPX Humulus_lupulus.tAPX Ziziphus_jujuba.tAPX Fragaria_vesca.tAPX Rubus occidentalis.tAPX Prunus_persica.tAPX Eucalyptus_grandis.tAPX Punica_granatum.tAPX Citrus_sinensis.tAPX Theobroma cacao.tAPX Corchorus_capsularis.tAPX Nymphaea colorata.tAPX Spirodela_polyrhiza.tAPX Lemna_minor.tAPX Zostera_marina.tAPX Xerophyta_viscosa.tAPX Eichhornia_paniculata.tAPX Elaeis_guineensis.tAPX Phoenix_dactylifera.tAPX Musa_acuminate.tAPX Ensete_ventricosum.tAPX Ananas comosus.tAPX Oryza_sativa.tAPX Brachypodium_distachyon.tA Sorghum_bicolor.tAPX Zea_mays.tAPX Nelumbo_nucifera.tAPX Fraxinus_excelsior.tAPX Sesamum indicum.tAPX Solanum_lycopersicum.tAPX Manihot esculenta.tAPX Hevea_brasiliensis.tAPX Populus_trichocarpa.tAPX Salix_purpurea.tAPX Anacardium_occidentale.tAP Gossypium_raimondii.tAPX Gossypium_arboreum.tAPX Tarenaya hassleriana.tAPX Aethionema_arabicum.tAPX Brassica_rapa.tAPX Schrenkiella_parvula.tAPX Capsella_rubella.tAPX Arabidopsis_thaliana.tAPX

Picea_abies.tAPX Pseudotsuga_menziesii.tAPX Ginkgo_biloba.tAPX Amborella_trichopoda.tAPX Dioscorea_alata.tAPX Asparagus_officinalis.tAPX

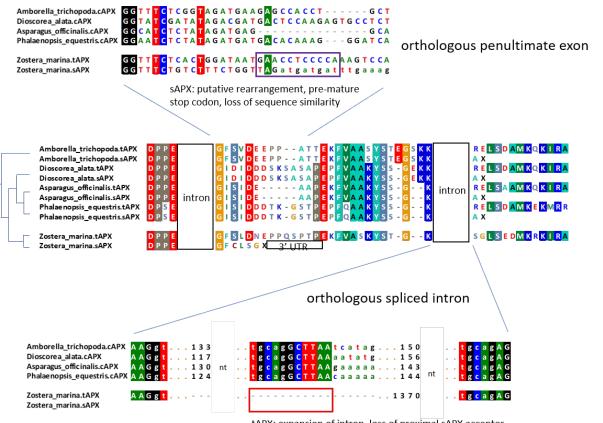
	ELSEAMKAKM	RAFYLALGGS	PNKPLOSNYF		A V L A S F F G - G N X	_
ĸ	ELSEAMKSKM	RAEYLAFGGS	PNKPLQSNYF	LNIIIFVAVL	AVLASFFG-GNX	-
	ELSEAMKKKI	RAEYLAFGGS	PNKPLQSNYF	LNIIFVAVL	A I L A A L F G N X	-
	ELSDAMKQKI ELSDAMKOKI	RAEYEAIGGS	PDKPLKSNYF	EN IMILVAAL	A F L T S L L G S X	2
	ELSAAMKQKI	RAEYEALGGG	PDKPLQSNYF	LNIIFVAVL	A L L S S L V V N X	_
x	ELSDAMKEKM	I r r e y e a v g g s	PDKPLKSNYF	LNIMIVVALL	A F L T S L I G N X	-
×	ELSEAMKQKI	RREYEAVGGS	PDKPLKSNYF		AFLTSLLGNX	1
	ELSDAMKOKI	RAEYEAVGGS	PDKPLQSNYF		A I L T Y F F G N X	2
	ELSDSMKQKI	RAEYEGLGGS	PDKPLSSNYF	LNIMIVIGVL	AVLSYLA GNX	-
	ELSDSMKEKI	RAEYEGFGGS	PNKPLPTNYF	LNIMIVIGVL	A V L S Y L A G N X	-
	ELSDSMKQKI ELSDNMKOKI	RAEYEAIGGS	PDKPLQSNYF PDKPLQSNYF		A I L T S L F G N X	-
	ELSDAMKEKI	RAEYOAVGGS	PDKPLOSNYF	LNIMIVIAVE	AILTSLLGNX	
	ELSEAMKQKI	RAEYEGLGGS	PDKPLPSNYF	LNIMIVIGVL	A I L T S L L G N X	-
	ELSDSMKQKI	RAEYEGFGGS	PDKPLQSNYF	LNIIIVIAVL	A F V T S F V M N X	-
	ELSDAMKQKI ELSDAMKOKI	RAEYEGLGGI	PDKPLQSNYF	EN IMIVIAVL	A F V T S L V V N X	-
	ELSDSMKQKI	RAEYEAVGGS	PDKPLSSNYF	LNIIIVVAVL	AYVTSLVMNX	-
	QLSQSMKEKI	RAEYEAVGGS	PDKPLQSNYF	LNIMIFIAVL	AYVTSLV MNX	-
	ELSDAMKQKI	RAEYEGLGGS	PDKPLPTNYF	LNIMIVIGVL	A L L T S L L G N X	-
	ELSDAMKQKI ELSDAMKOKI	RAEYEGEGGS	ADKPLPINYF		AIVTSLL-GNX	2
	ELSDAMKQKI	RAEYEGLGGT	PDKPLPTNYF	LNIIVIGVL	A I L T Y L L G N X	-
	E L S D AM K Q K I	QAEYEGLGGS	PDKPLPTNYF	LNIMILIGVL	A L L T S L L G N X	-
x	E L S D AM K Q K I	RAEYEGEGGS	PDKPLKTNYF		ALLTSLI-GNX	-
	ELSDAMKOKI	RAEYEAVGGS	PDKPLQSNYF		A F L T S L V G N X	2
	EFSDSMKQKM	I R A E Y E A I G G S	PSKALSSNYF	LNIIIVISVL	. A V L T Y L F G N X	-
	EFSDSMKQKM	RAEYEAIGGS	PNKALSSNYF	LNIIIVISVL	A V L T Y L F G N X	-
	ELSDAMRKKI ELSEAMKOKI	RAEYEAVGGS	PDKALKSNYF	ENTITVIAAL	A I L T Y L F G N X	2
	ELSDGMKQKI	RAEYEAIGGS	PDKPLQSNYF	LNIIIIAVL	A F L T S L L G N X	_
	ELSDSMKQKI	RAEYEAIGGS	PDKPLQSNYF	LNIMIVIAVL	A F L T S L L G N X	-
	ELSDPMKQKI	RAEYEGFGGS	PDKPLPTNYF		A I L T S L L G N X	-
	ELSENMKQKI	RAEYEAVGGS	PDKPLQSNYF		A F L T S L F G N X	2
	ELSDSMKQKI	RAEYQAIGGS	PDKPLQTNYF	LNIIIVISVL	A I L T S L F G N X	-
	ELSDSMKQKI	RAEYQAIGGS	PDKPLQSNYF	LNIIIVIAVL	AILTSLF-GNX	1
	ELSENMKQKI FISDAMKOKI	RAEYQAVGGS	PDKPLQSNYF		A L L T S L V G N X	2
	ELSDAMKQKI	RAEYETIGGS	PDKPLKSNYF	LNIMIVIAAL	A V L T S L F G N X	-
	ELSEAMKQKI	RAEYEAVGGS	PDTPLRSNYF	LNIMIVVAVL	A L L T S L F G N X	-
	ELSEAMKQKI ELSDSMKOKI	RAEYEGLGGS	PD K P L P T N Y F	LNIMITIGVL		-
	ELSEAMKQKI	RAEYEAVGGS	PDKPLKSNYF	LNIMIVVAVL	A I L T S L L G S Y X	2
	ELSNSMKQKI	RAEYEAIGGS	PDKPLQSNYF	LNIMIFIAGL	A F L T S L L G N X	-
	ELSDSMKQKI	RAEYEAIGGS	PDKPLQSNYF	LNIILFIGGL	ALLSSLL ANX	-
	ELSETMKOKI	RAEYEGEGGS	PDRPLOSNYF		V V L T Y F L G N X	2
	ELSDSMKQKI	RAEYEAVGGS	PDKPLQSNYF	LNIMILIAGL	A F L T S L V G N X	-
	ELSEAMKQKI	RAEYEALGGS	PDKPLKSNYF	LNIMILIAAL	AFLTSLLGNX	-
	ELSEAMKQKI ELSDAMKOKI	RAEYEAVGGS	PDKPLKSNYF	EN IMILVAAL	A F L T S L L G N X	-
	ELSDAMKQKI	RAEYEALGGS	PDKPLKSNYF	FNIMILIAGE	A F L T S L V M N X	_
	ELSDAMKQKI	RAEYEAIGGS	PDKPLKSNYF	LNIMIVIAAL	A F L T S L L G N X	-
APX	ELSDSMKQKI	RAEYEGEGGS	PDKPLQSNYF		A F L T S L L G S X	-
AFA	ELSESMKOKI	RAEYEGEGGGS	PDKPMQSNYF		AFLTSLVGNX	2
	ELSDAMKQKI	RAEYEGFGGS	PDKPMQSNYF	LNIMILIAGL	A F L T S L V G N X	-
	ELSDSMKQKI	RAEYEAIGGS	PDKPLQSNYF	LNIMIVIAVL	A F L A S L V V N X	-
	ELSESMKQKI	RAEYEGEGGS	PDKPLQSNYF	EN IMIVIAIL	A L V T S L V M N X	2
1	ELSDAMKQKI	RAEYEGLGGT	PDKPLPTNYF	LNIIIVIGVL	AILTYLL GNX	-
	ELSEAMKQKI	RAEYEAVGGS	PDKPLKSNYF	LNIMIVIAVL	A L L T S L L G N X	-
	ELSEAMKQKI	RAEYEAIGGS	PDKPLQSNYF	LNIMIVIAVL	A F L T S L L G N X	-
	ELSETMKOKI	RAEYEAVGGS	PDKPLQSNYF			2
PX	ELSEAMKQKI	RAEYEAVGGS	PDKPLRSNYF	LNIMIVIAVL	A F V T S L V V N X	-
	ELSDSMKQKI	RAEYESFGGG	PDKALPTNYF	LNIIIIGVL	AILTSLL GNYX	-
	ELSDSMKQKI	RAEYESFGGG	PDKPLPTNYF		A I L T S L L G N N X	1
	ELSDSMKKKI	RAEYEAIGGS	PNKHLPTNYF	LNIIIAITVL	V L L T Y L S G N F S D P S S F X - V L L G T L L G N N N T S D Y S G F S V L L F T F L G N N N S D Y S G F S V L L F T F F G N N N S D Y S G F X	x
	ELSDSMKKKI	RAEYEAIGGS	PDKPLPTNYF	LNIIAISVL	VLLFTFLGNNNSSDYSGF	x
	ELSDSMKKKI	RAEYEAIGGS	PNKPLPTNYF	LNIIAITVL	VLLFTFFGNNNSDYSGFX	-
					V L L S T L F G G G N N N S D F S G F X V L L S T L F - G G N N N S D F S G F X	-
				- ALL REAL OF		

Fig. S3. Case studies of sub-localized duplicated *cpAPX* genes. Amino acid and nucleotide sequences of sub-localized paralogous are aligned with several single-copy orthologs. Red boxes and purple boxes indicate the nucleotide sequence changes contributing to the tAPX specialization and sAPX, respectively, and a short description is indicated beside the boxes. (a) *Nelumbo nucifera*; (b) *Zostera marina*; (c) Musaceae.





b



tAPX: expansion of intron, loss of proximal sAPX acceptor

С

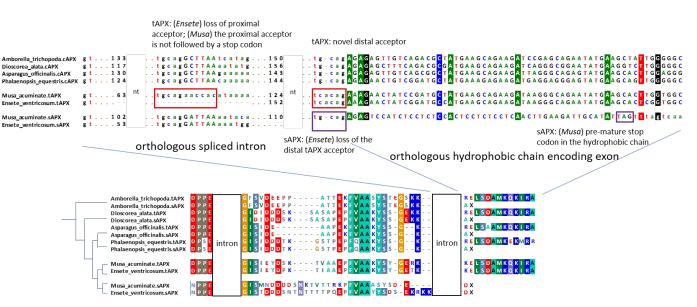
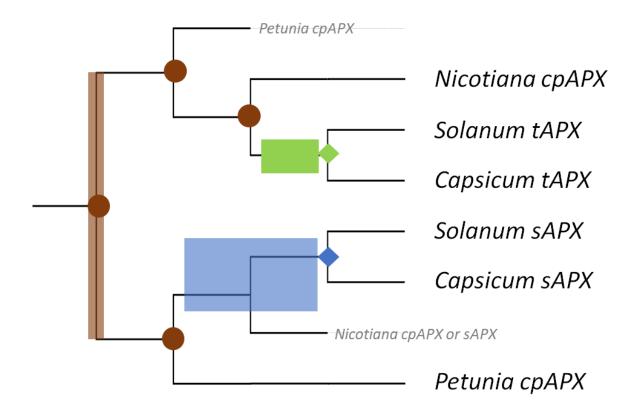


Fig. S4. The evolution of *cpAPX* genes in Solanaceae. The genes in black are in the extant genomes. The genes in grey were presumably present in the ancestral genome and were lost. Brown bar indicates gene duplication. Brown circles indicate ancestral alternatively spliced genes. Green and blue boxes indicate possible timing when the inferred gene function specialization to *tAPX* and *sAPX* took place. Green and blue diamonds indicate the ancestral *tAPX* and *sAPX* gene before the divergence of *Solanum* and *Capsicum* from a common ancestor.



References:

Miyake C, Asada K. 1992. Thylakoid-bound ascorbate peroxidase in spinach chloroplasts and photoreduction of its primary product monodehydroascorbate radicals in thylakoids. Plant Cell Physiology **33**(5):541–553.