





convert to AND

negate

Gene | lox9

negate

Plant Reactome Pa Jasmonic acid bio

URL: https://maize-pangenome.gramene.org

Lu, Zhenyuan¹; Chougule, Kapeel¹; Olson, Andrew¹; Tello-Ruiz, Marcela K¹; Wei, Sharon¹; George, Nancy²; Papatheodorou, Irene²; Ware, Doreen¹³ ¹Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724; ²EMBL-European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, CB10 1SD, UK; ³USDA ARS NEA Plant Soil & Nutrition Laboratory Research Unit, Ithaca, NY, 14853, USA

Abstract

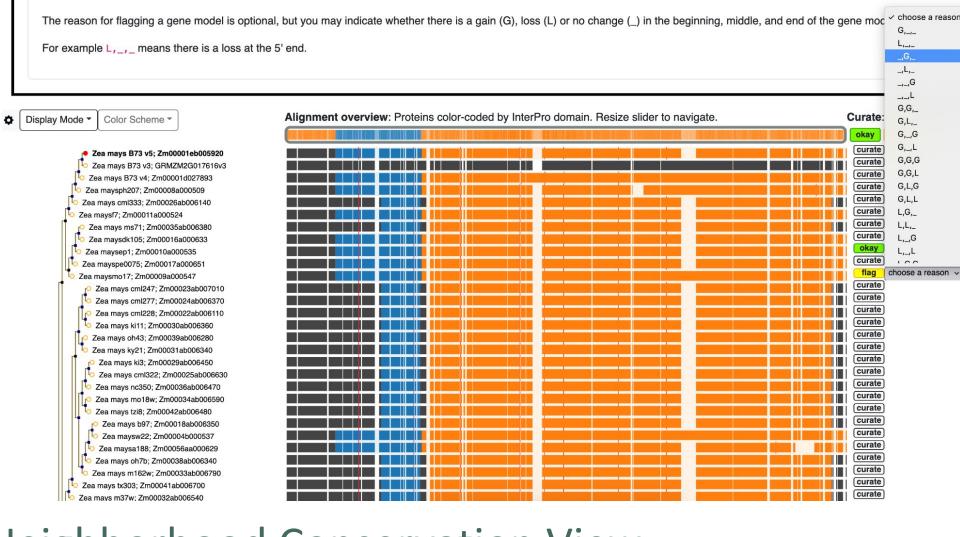
The Gramene project has developed pan-genome subsites, as a natural extension of the Gramene and Ensembl infrastructures. Each pan-genome site is dedicated to the study of individual crop groups (e.g., maize, rice, sorghum, and grape). The Gramene project (http://www.gramene.org) was a key player in sequencing the first maize reference B73, more recently the 25 NAM founders, continues to be engaged with the community. The maize pansite (https://maize-pangenome.gramene.org), established in 2021, hosts reference assemblies for 26 maize accessions. Each maize accession is hosted in a separate genome browser, providing access to gene-based views with entry points via text-based searches.. The B73 V5 genome serves as the reference assembly for anchoring expression, population and pathway views; transcript abundance across a gene model at different tissues and stages of development and views of paralog expression support decision making on candidate gene selection. Phylogenetic analyses are based on maize-specific gene trees built from the assigned canonical protein for each gene locus. The protein-based trees allow users to rapidly traverse between the different maize accessions, where alleles are considered orthologs in the trees, as well as other species; copy-number variations within maize can also be evaluated alongside lineage-specific gene expansions. Protein homology is viewable as amino acid alignments and as gene neighborhood conservation. We adapted these views to deploy community curation tools-from the homology tab-for users to flag potential structural annotation issues. More recent work has used these gene trees to build a pan-gene index used as the inputs to improve and extend the previous gene structural annotation workflow. Release 3 will include updated annotations and access to 8 maize genomes (4 European Flint, Mo17, W22, PH207 & A188), expression, and Maize ENCODE (NSF-IOS-1445025) related studies. Funded by USDA-ARS-8062-21000-041-00D.

Funding acknowledgement: United States Department of Agriculture (USDA)

New Search and Quick Views

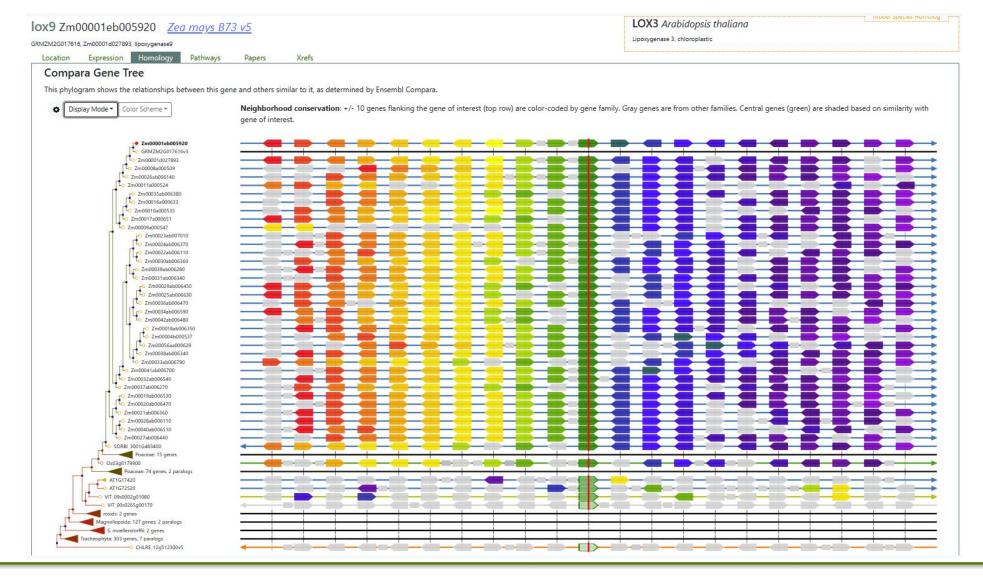
Plant Reactome Pathway (/ox 13-LOX and 13-HPL pathway 27 Divinyl ether biosynthesis II (13-LOX) (InterPro Homologous_superfamily (/ox LipOase_C_sf 17 PLAT/LH2_dom_sf 27 Peptidase_S1_PA 623 PH-like_dom_sf 799 Homing_endonucl 11 Ntn_hydrolases_N 574 PIN-like_dom_sf 199 InterPro Family XPG/Rad2_eukaryotes 11 Lysyl_oxidase 2 Tfb1/GTF2H1 13 LipOase 123 Alpha-hydoxy_acid_DH_FMN 66 Proteasome_sua/b 1425 InterPro Domain LAGLIDADG_2 27 LipOase_C 477 TFIIH_BTF_p62_N 7 PLAT/LH2_dom 605 InterPro Binding_site Lage 100 State State State	rch for <i>lipoxygen</i> (9) gene
13-LOX and 13-HPL pathway 27 Divinyl ether biosynthesis II (13-LOX) (2) InterPro Homologous_superfamily LipOase_C_sf 17 PLAT/LH2_dom_sf 27 Peptidase_S1_PA 523 PH-like_dom_sf 799 Homing_endonucl 11 Ntn_hydrolases_N 574 PIN-like_dom_sf 199 InterPro Family XPG/Rad2_eukaryotes 11 Lysyl_oxidase 2 Tfb1/GTF2H1 13 LipOase 123 Alpha-hydoxy_acid_DH_FMN 36 Proteasome_sua/b 1425 InterPro Domain LAGLIDADG_2 37 LipOase_C 471 TFIIH_BTF_p62_N 7 PLAT/LH2_dom 505 InterPro Binding_site InterPro Binding_site InterPro Binding_site InterPro Binding_site	(9) gene
13-LOX and 13-HPL pathway 22 Divinyl ether biosynthesis II (13-LOX) 4 InterPro Homologous_superfamily LipOase_C_sf 17 PLAT/LH2_dom_sf 22 Peptidase_S1_PA 523 PH-like_dom_sf 799 Homing_endonucl 11 Ntn_hydrolases_N 574 PIN-like_dom_sf 199 InterPro Family XPG/Rad2_eukaryotes 11 Lysyl_oxidase 2 Tfb1/GTF2H1 13 LipOase 123 Alpha-hydoxy_acid_DH_FMN 65 Proteasome_sua/b 1425 InterPro Domain LAGLIDADG_2 27 LipOase_C 471 TFIIH_BTF_p62_N 7 PLAT/LH2_dom 505 InterPro Binding_site InterPro Binding_site InterPro Binding_site InterPro Binding_site	107 Berre
LipOase_C_sf 17 PLAT/LH2_dom_sf 27 Peptidase_S1_PA 523 PH-like_dom_sf 799 Homing_endonucl 11 Ntn_hydrolases_N 574 PIN-like_dom_sf 199 InterPro Family XPG/Rad2_eukaryotes 11 Lysyl_oxidase 2 Tfb1/GTF2H1 13 LipOase 123 Alpha-hydoxy_acid_DH_FMN 36 Proteasome_sua/b 1425 InterPro Domain LAGLIDADG_2 37 LipOase_C 471 TFIIH_BTF_p62_N 7 PLAT/LH2_dom 505 InterPro Binding_site	
InterPro Family XPG/Rad2_eukaryotes 11 Lysyl_oxidase 2 Tfb1/GTF2H1 13 LipOase 123 Alpha-hydoxy_acid_DH_FMN 36 Proteasome_sua/b 1425 InterPro Domain LAGLIDADG_2 37 LipOase_C 471 TFIIH_BTF_p62_N 7 PLAT/LH2_dom 505 InterPro Binding_site	
XPG/Rad2_eukaryotes 11 Lysyl_oxidase 2 Tfb1/GTF2H1 13 LipOase 123 Alpha-hydoxy_acid_DH_FMN 36 Proteasome_sua/b 1425 InterPro Domain LAGLIDADG_2 37 LipOase_C 471 TFIIH_BTF_p62_N 7 PLAT/LH2_dom 505 InterPro Binding_site	
InterPro Domain LAGLIDADG_2 37 LipOase_C 471 TFIIH_BTF_p62_N 7 PLAT/LH2_dom 505 InterPro Binding_site	
LAGLIDADG_2 37 LipOase_C 471 TFIIH_BTF_p62_N 7 PLAT/LH2_dom 505 InterPro Binding_site	
InterPro Binding_site	
LipOase_Fe_BS 84	
InterPro Conserved_site	
LipOase_CS 82	
GO function	
protein-lysine 6-oxidase activity 2	Complex querie

- Location Genomic context
- **Expression** Baseline gene expression from EBI
- **Homology** Customizable Ensembl Compara gene family trees with displays
- Alignment Overview (full-length gene)
- Multiple-Sequence Alignment (zoom into the amino acid level)
- Neighborhood Conservation View (zoom out +/- 10 flanking genes)
- **Pathways** Association to pathways in Gramene's Plant Reactome
- **Xrefs** Cross-references to other databases
- Aligned tree view enables structural gene curation



Mark genes as okay or flag genes that you think might have issues and choose a reason from the drop down menu. When finished, enter your email address and submi

Gene Neighborhood Conservation View





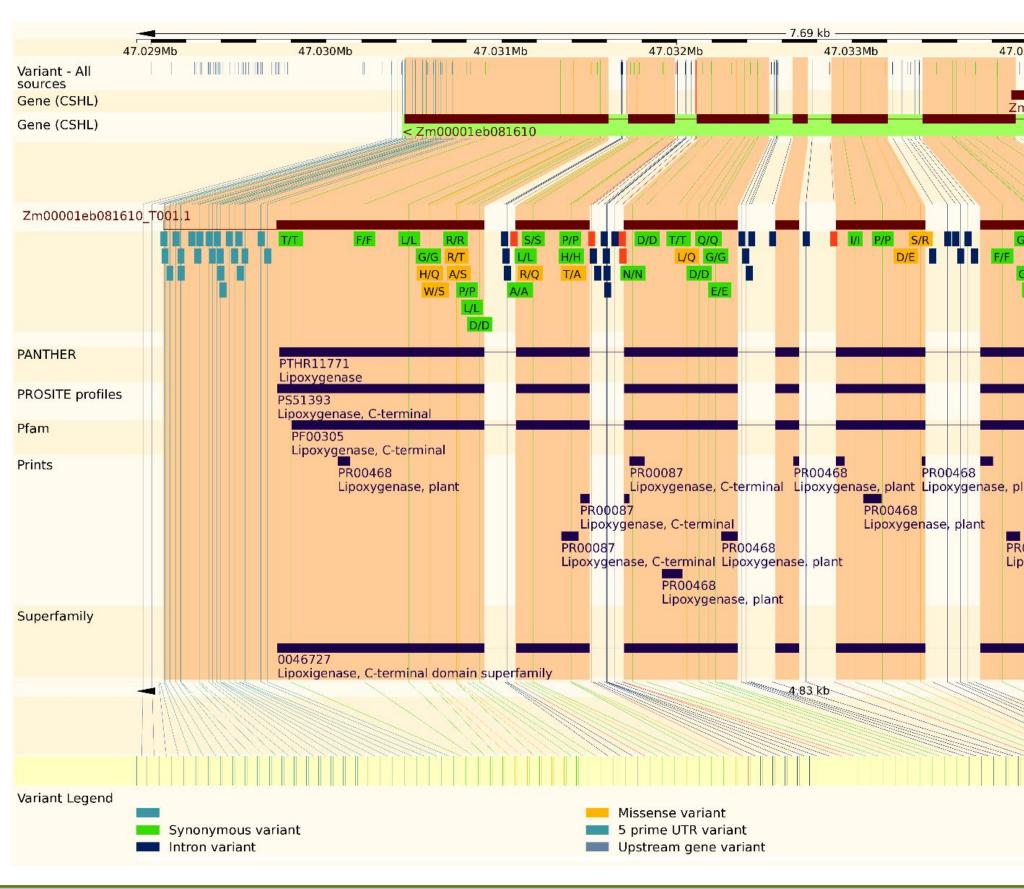
Gramene PanMaize: One-Stop Pan-Genome Browser for **Exploring the Rich Genetic Diversity in Maize**

A multi-species phylogenomic and pathway-informed community resource

- Release 3 includes 43 plant genomes with 26 Maize NAM lines and 8 new genomes (4 European Flint, Mo17, W22, PH207 & A188)
- Integrated gene function from the literature enabling searches by TO/PO terms > 36K Gene Tree families with ortholog & paralog calls
- Manually curated B73 v4 gene models mapped to v5 • Gene curation user interface
- Visualization of variant effect prediction over protein domains
- Gene expression (EBI-Atlas) & orthology-based pathway projections (Plant Reactome) for B73 v5
- Blast service
- Programmatic access
- Quick start guide

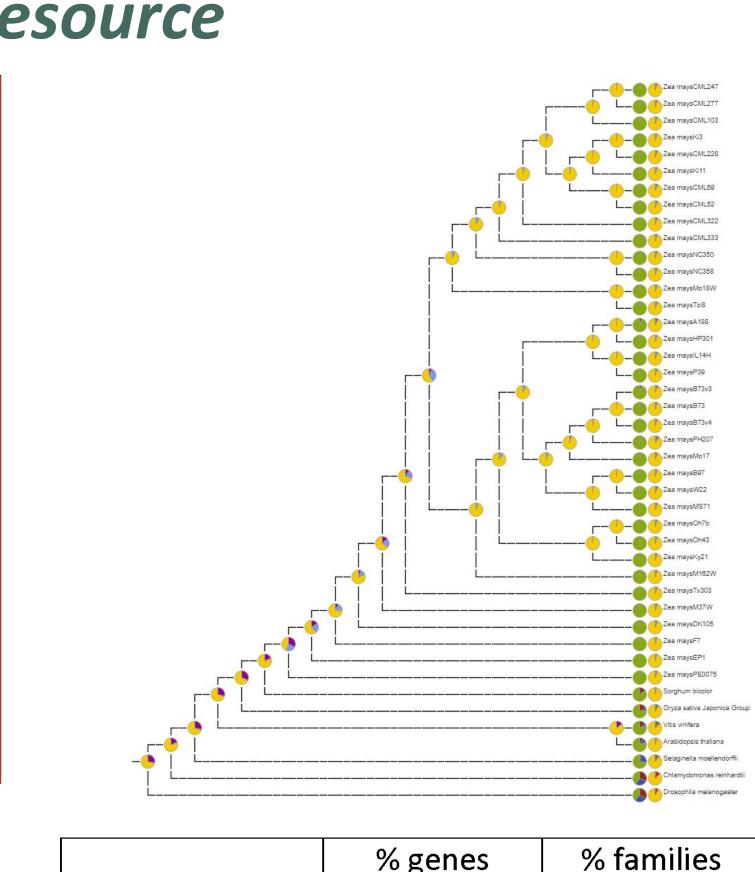
Structural variation for NAM genomes (B73 v5 browser)

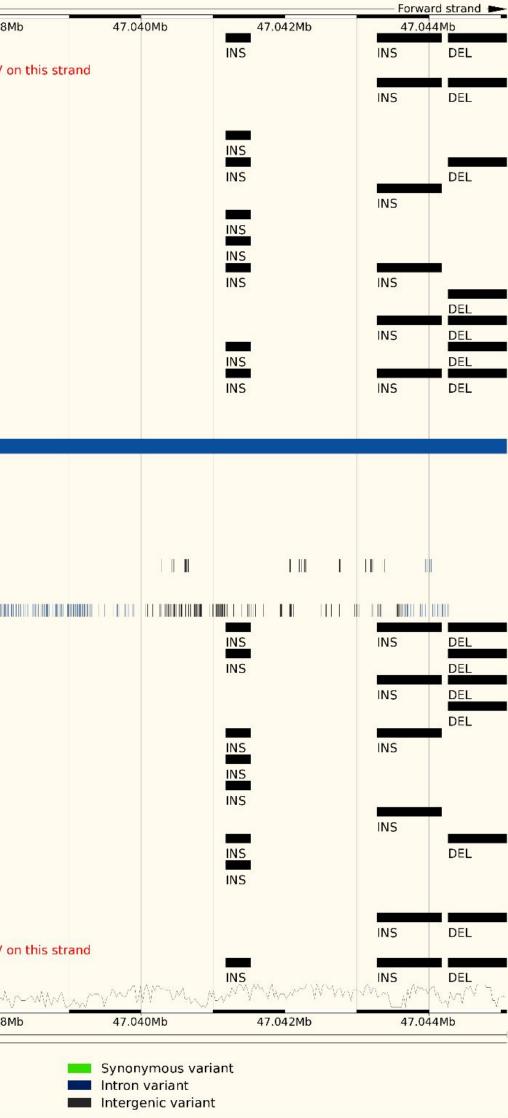
						18.27 kb					rd strand 🗭
	NAM Zm-Oh7B SV	47.028Mb	47.030Mb	47.032Mb	47.034Mb	47.036Mb	47.038Mb	47.040Mb	47.042Mb	47.04	
	NAM_Zm-Oh43_SV					No features from NAI	M_Zm-Oh43_SV on this st	rand	INS	INS	DEL
	NAM_Zm-NC358_SV					DEL				INS	DEL
	NAM_Zm-NC350_SV					DEL					
	NAM_Zm-Ms71_SV								INS		
	NAM_Zm-Mo18W_SV								INS		DEL
	NAM_Zm-M37W_SV					DEL			_	INS	
	NAM_Zm-M162W_SV								INS		
	NAM_Zm-Ky21_SV					DEL			INS		
	NAM_Zm-Ki11_SV								INS	INS	
	NAM_Zm-II14H_SV										DEL
	NAM_Zm-CML52_SV					DEL				INS	DEL
	NAM_Zm-CML228_SV								INS		DEL
	NAM_Zm-CML103_SV					F			INS	INS	DEL
	Gene (CSHL)				Zm00001e protein coo	b081620_T001 >					
	Contigs				protein coc	scaf_1 >					
	Gene (CSHL)		<	Zm00001eb081610_T001							
				rotein coding		_					
				< Zm00001eb081610_T00 protein coding)2						
	Variant - Panzea_	Variant - Panzea_2.7GBS				11.1	0	1.0.00	1.111 1		II.
he fly!!	2.7GBS				1.1			1 11 11	1 11 1		П
	Variant - HapMap2	Variant - HapMap2									
¢	NAM_Zm-CML103_SV								INS	INS	DEL
Ū	NAM_Zm-CML228_SV								INS	200	DEL
_	NAM_Zm-CML52_SV					DEL				INS	DEL
	NAM_Zm-II14H_SV										DEL
	NAM_Zm-Ki11_SV								INS	INS	
-	NAM_Zm-Ky21_SV					DEL			INS		
	NAM_Zm-M162W_SV								INS		
	NAM_Zm-M37W_SV					DEL				INS	
	NAM_Zm-Mo18W_SV								INS		DEL
	NAM_Zm-Ms71_SV								INS		
-	NAM_Zm-NC350_SV					DEL					
	NAM_Zm-NC358_SV					DEL				INS	DEL
	NAM_Zm-Oh43_SV					No features from NAI	M_Zm-Oh43_SV on this st	rand			
	NAM_Zm-Oh7B_SV %GC	NO MARKAN	10 - 10 m s	MWYM MALINA	MM/VWW.nov ~	MAT A	M MARA	- ANTAN		INS	
		47.028Mb	<u>/////////////////////////////////////</u>	47.032Mb	47.034Mb	47.036Mb	47.038Mb	47.040Mb	47.042Mb	47.04	4Mb
	M	Reverse strand	UNDCO.17	T1.032ND	47.034140	18.27 kb	UP1050,7+	47.04000	47.042MD	47.04	עויח
	Variant Legend	Missense variant			Splice region variant			Synonymous variant			
			5 prime UTR variant		 3 prime UTR variant Downstream gene variant 			Intron variantIntergenic varia	nt		
	Gene Legend	Upstream gene variant Protein Coding									
	Gene Legenu										
		protein coding									
		There are currently 42 trac Maize PanGenome Zea ma		87 1 (7m-B73-BEEEBENCE	-NAM-5 0) Chromosom	e 2. 47 026 800 - 47 0/	45.081				





Genetic variation (Hapmap2 & Panzea) & structural variation for the NAM genomes vs B73 v5

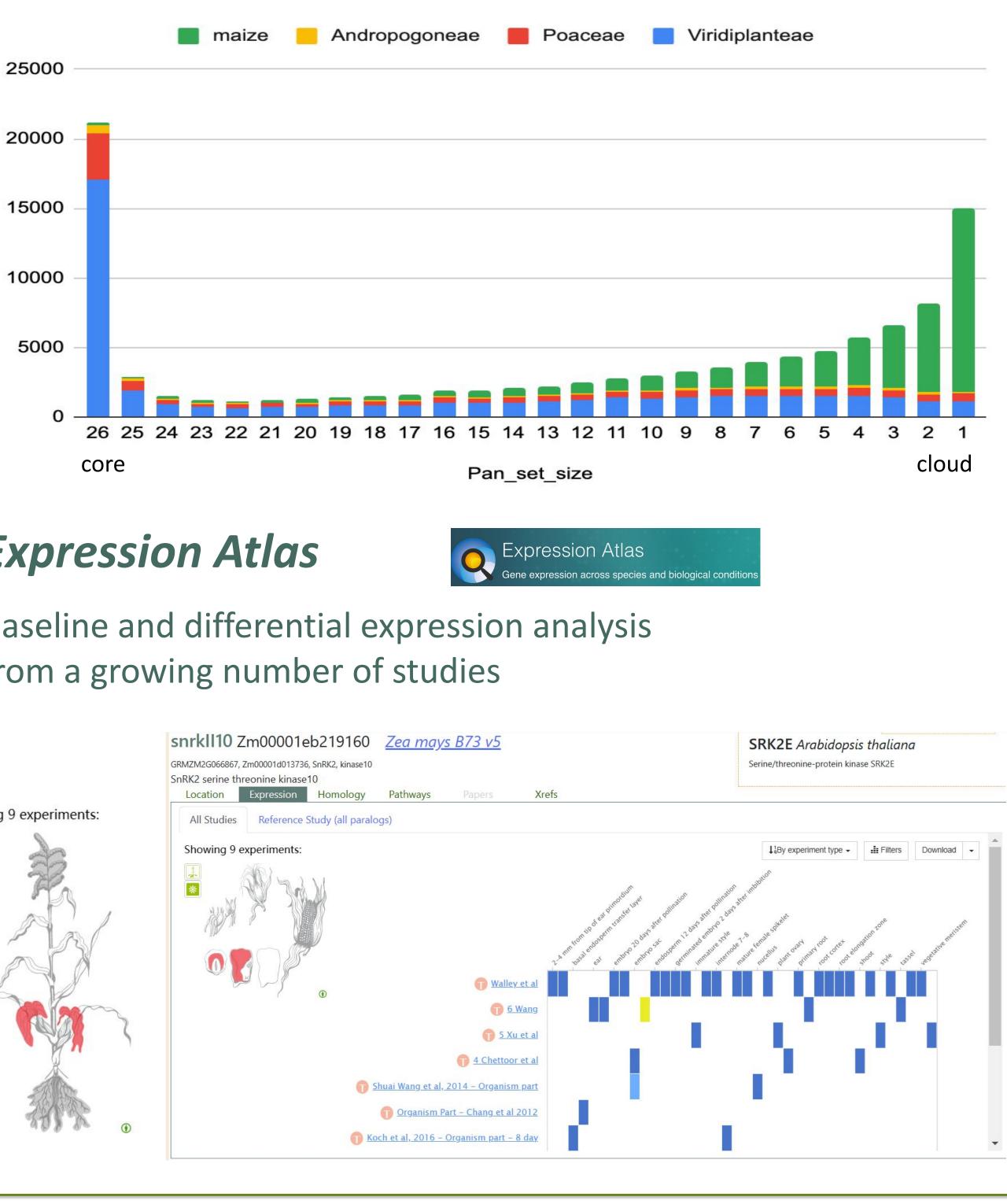




Origin of Maize Genes

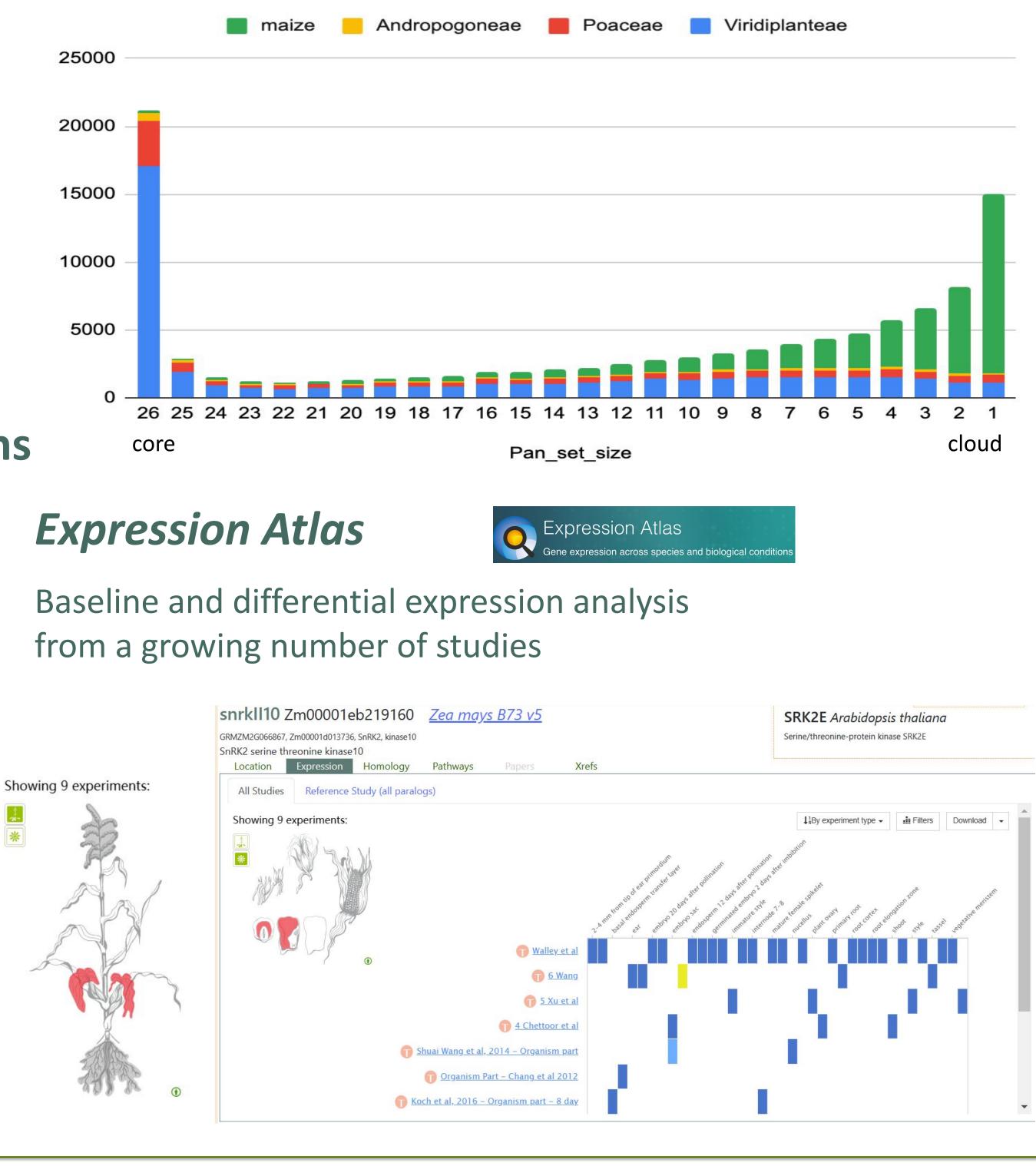


Pan gene distribution by age from Maize NAM genome annotations



G/G S/S Lipoxygenase, plan Lipoxygenase, plant PR00468 Lipoxygenase, plant PLAT/LH2 domain superfamil Splice region varian 3 prime UTR variant







	% genes	% families
nic Origin	n=39,324	n=9236
а	38.4	35.2
ntae	6.6	7.2
hyta	21.2	16.3
ohyta	2.8	2.9
ophyta	10.6	13.0
nids	0.8	1.3
	8.1	14.6
eae	0.9	1.7
goneae	0.6	1.5
5	5.3	6.2
	4.7	0.0

