Table S1. List of sequences used in this study.

Species	Identifier	Database
Angiosperm NRPE1 homologs		
Amborella trichopoda	evm_27.model.AmTr_v1.0_scaffold00131.41	Amborella Genome
Arabidopsis thaliana	AT2G40030	TAIR
Aquilegia coerulea	Acoerulea.22047564 Acoerulea.22025488	Phytozome 10
Citrus clementina	Cclementina.20791338	Phytozome 10
Citrus sinensis	Csinensis.18120100	Phytozome 10
Cucumis sativus	Csativus.16978203	Phytozome 10
Eucalyptus grandis	Egrandis.23599136	Phytozome 10
Fragaria vesca	Fvesca.27249047	Phytozome 10
Glycine max	Gmax.30498456 Gmax.30505760	Phytozome 10
Gossypium raimondii	Graimondii.26760711	Phytozome 10
Linum usitatissimum	Lusitatissimum.23162059	Phytozome 10
Medicago truncatula	Mtruncatula.31075688	Phytozome 10
Mimulus guttatus	Mguttatus.28945928	Phytozome 10
Oryza sativa	LOC_Os01g73430 LOC_Os02g05880	Phytozome 10
Panicum virgatum	Pvirgatum.30305942 Pvirgatum.30252426	Phytozome 10
Phaseolus vulgaris	Pvulgaris.27152868	Phytozome 10
Prunus persica	Ppersica.17669039	Phytozome 10
Populus trichocarpa	Ptrichocarpa.27003419 Ptrichocarpa.26998728	Phytozome 10
Salix purpurea	Spurpurea.31444571	Phytozome 10
Setaria italica	Sitalica.19697436 Sitalica.19688672	Phytozome 10
Solanum lycopersicum	Slycopersicum.27300428	Phytozome 10
Sorghum bicolor	Sbicolor.28384871	Phytozome 10
Theobroma cacao	Tcacao.27443915	Phytozome 10
Vitis vinifera	Vvinifera.17824747	Phytozome 10
Oryza and related *NRPE1a*		
Brachypodium distachyon	Bradi4g45065	Phytozome
Leersia perrieri	LPERR02G03610	EnsemblPlants
Oryza barthii	OBART02G04100	EnsemblPlants

Oryza brachyantha	OB02G13460	EnsemblPlants
Oryza glaberrima	ORGLA02G0038200	EnsemblPlants
Oryza glumaepatula	OGLUM02G03900	EnsemblPlants
Oryza longistaminata	OLONG_014977	EnsemblPlants
Oryza meridionalis	OMERI0204920	EnsemblPlants
Oryza nivara	ONIV02G04030	EnsemblPlants
Oryza punctata	OPUNC02G03300	EnsemblPlants
Oryza rufipogon	ORUFI02G04170	EnsemblPlants
Oryza sativa ssp japonica	LOC_Os02g05880	Phytozome
Oryza sativa ssp indica	BGIOSGA007122	EnsemblPlants
Triticum aestivum	genomic region near Traes_6AS_B5ED3BA5B	Phytozome

Brassicaceae and related NRPE1

Aethionema arabicum	n.a.[1]	CoGe
Arabidopsis arenosa	n.a.[2]	SRA PRJNA196706[3]
Arabidopsis halleri	Araha.54531s0001	Phytozome 10
Arabidopsis lyrata	483042	Phytozome 10
Bochera strica	Bostr.23794s0061	Phytozome 10
Brassica oleraceae	Bo3g034650	EnsemblPlants
Brassica rapa	Bra000162	EnsemblPlants
Camelina sativa	n.a.[1]	CoGe
Capsella grandiflora	Cagra.2495s0018	CoGe
Capsella rubella	Carubv10022497m	Phytozome 10
Eutrema salsugineum	Thhalv10016128m	Phytozome 10
Leavenworthia alabamica	n.a.[1]	CoGe
Neslia paniculata	n.a.[1]	CoGe
Schrenkiella parvula	Tp4g22340	CoGe
Sisymbrium irio	n.a.[1]	CoGe
Tarenaya hassleriana	Th2v22033	CoGe

Angiosperm SPT5L homologs

Amborella trichopoda	evm_27.model.AmTr_v1.0_scaffold00002.356	Amborella Genome
Arabidopsis lyrata	487210	Phytozome 11
Arabidopsis thaliana	AT5G04290.1	TAIR
Aquilegia coerulea	Aquca_038_00139	Phytozome 11

Boechera stricta	Bostr.2128s0018.1.p	Phytozome 11
Brachypodium distachyon	Bradi2g20517.2.	Phytozome 11
Brassica rapa	Brara.J02756.1.p	Phytozome 11
Capsella grandiflora	Cagra.9217s0009 Org_	Phytozome 11
Capsella rubella	Carubv10000039m	Phytozome 11
Citrus clementina	Ciclev10030480m.g	Phytozome 11
Citrus sinensis	orange1.1g000247m	Phytozome 11
Carica papaya	evm.model.supercontig_188.8 evm.model.supercontig_53.150	Phytozome 11
Eucalyptus grandis	Eucgr.J01043.1	Phytozome 11
Eutrema salsugineum	Thhalv10015730m	Phytozome 11
Fragaria vesca	mrna23431.1-v1.0-hybrid	Phytozome 11
Glycine max	Glyma.10G276900.1 Glyma.20G112800.1	Phytozome 11
Gossypium raimondii	Gorai.011G234600.1	Phytozome 11
Kalanchoe marnieriana	Kalax.0009s0224.1.p	Phytozome 11
Linum usitatissimum	Lus10037946 Lus10038680	Phytozome 11
Manihot esculenta	Manes.07G094600.1.p	Phytozome 11
Malus domestica	MDP0000201216 MDP0000232228	Phytozome 11
Medicago truncatula	Medtr1g111560.1	Phytozome 11
Mimulus guttatus	Migut.F01946.1.p	Phytozome 11
Oryza sativa	LOC_Os05g43060.1	Phytozome 11
Panicum hallii	Pahal.C02415.1	Phytozome 11
Panicum virgatum	Pavir.Ca01318.1.p Pavir.Cb01635.1.p	Phytozome 11
Phaseolus vulgaris	Phvul.007G024600.1	Phytozome 11
Populus trichocarpa	Potri.008G032800.1 Potri.010G228800.1	Phytozome 11
Ricinus. communis	29898.t000002	Phytozome 11
Salix purpurea	SapurV1A.0411s0080.1.p	Phytozome 11
Theobroma cacao	Thecc1EG043328t1	Phytozome 11
Zea. mays	GRMZM2G370715_P01 GRMZM2G375222_P01	Phytozome 11

[1] Orthologs identified by similarity in unannotated genome. See methods for details.
[2] Short reads mapped to *Ar. lyrata* genomes for ortholog and SNP identification.
[3] from Hollister, *et al*. (2012). Genetic adaptation associated with genome-doubling in autotetraploid Arabidopsis arenosa. *PLoS Genetics*, *8*(12), e1003093.

Table S2: McDonald-Kreitman test of adaptive evolution.

	Polymorphic		Fixed		
	Synon.	**Nonsyn.**	**Synon.**	**Nonsyn.**	**p-value**
Ar. thaliana – Ar. lyrata					
Catalytic	58	59	71	51	n.s.
Ago platform	26	77	18	44	n.s.
Ar. thaliana – Ar. arenosa					
Catalytic	58	59	75	51	n.s.
Ago platform	26	77	24	48	n.s.
Ar. thaliana – Bo. stricta					
Catalytic	58	61	93	63	n.s.
Ago platform	26	80	39	59	n.s.
Ar. thaliana – Cap. rubella					
Catalytic	58	61	128	69	0.0481
Ago platform	26	77	42	64	n.s.
Ar. thaliana – Ne. paniculata					
Catalytic	59	59	104	65	n.s.
Ago platform	26	76	32	56	n.s.

n.s., not significant (FDR-adjusted $p < 0.05$)

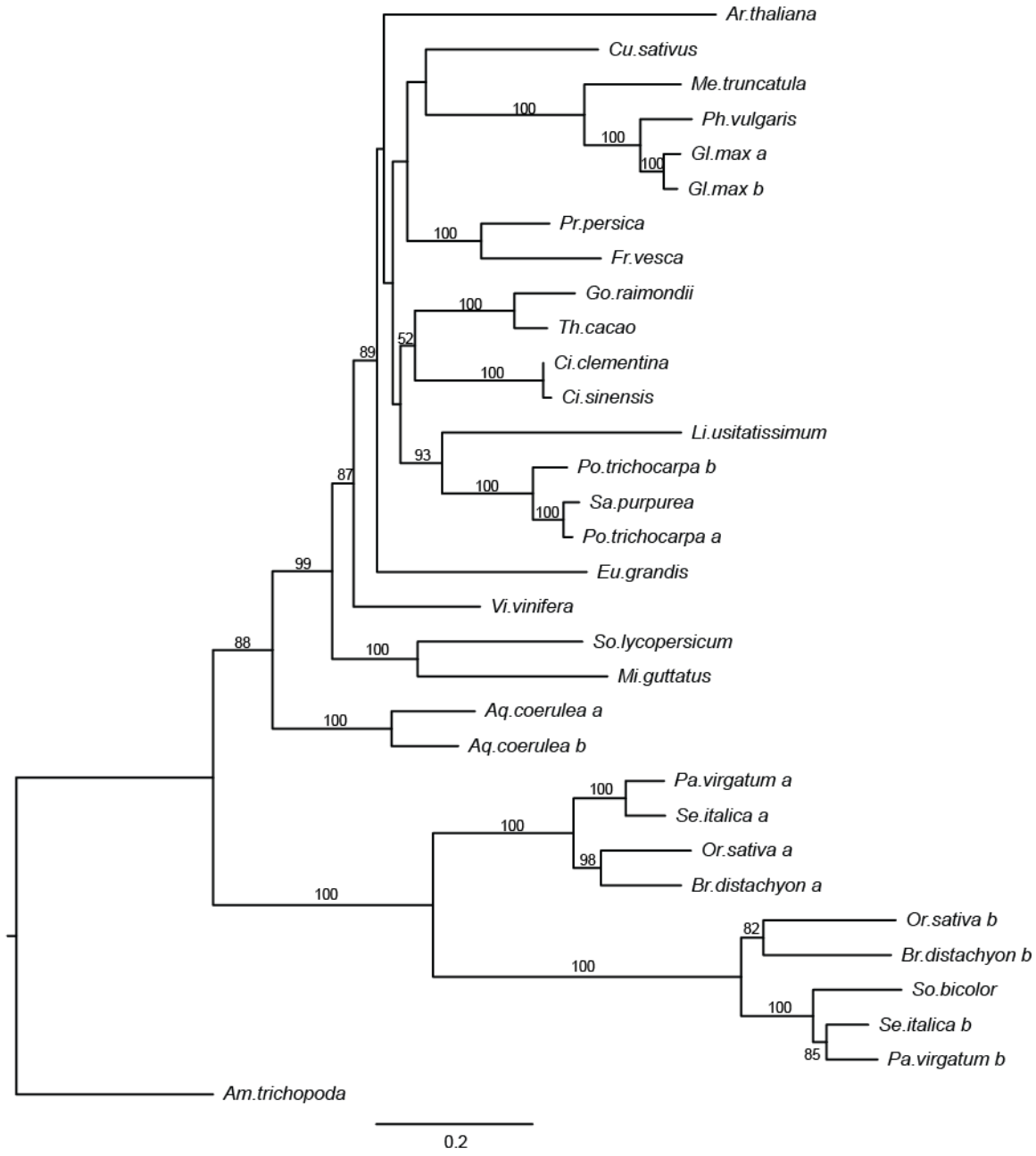

Figure S1. Phylogenetic relationship between NRPE1 orthologs analyzed in Figure 1. A maximum-likelihood tree generated from nucleotide sequence of catalytic domains A-H of 32 angiosperm NRPE1 orthologs. Positions in the alignment with a gap in >20% of taxa were removed. Branches list bootstrap support from 100 replicates when ≥50. The gene tree conforms to accepted organismal phylogeny.

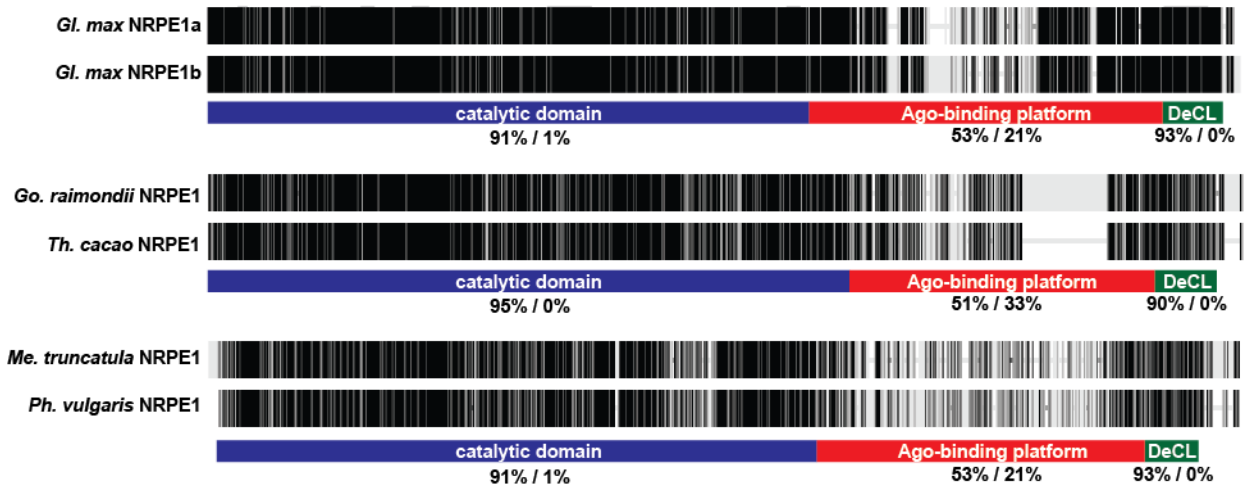

Figure S2. Pairwise comparison of NRPE1 orthologs. Amino acid sequences were aligned with MUSCLE. Identical positions are black and mismatches and gaps are gray. Values below the domain schematics are % positives / % gaps as measured with BLASTP.

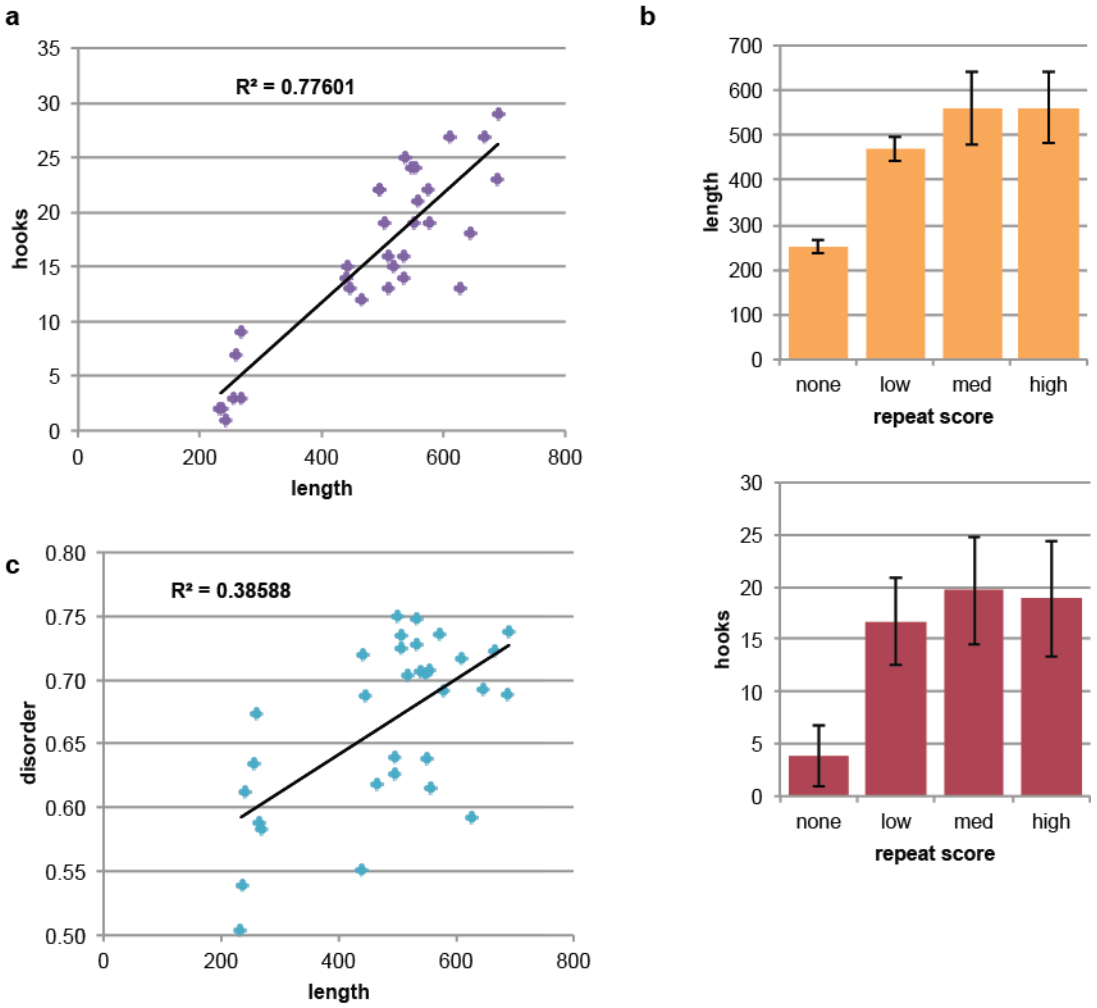

Figure S3. Conserved features of Ago-binding platforms are correlated. (a) Longer Ago-binding platforms have more Ago hooks. (b) Ago-binding platforms with higher repeat scores are longer (top) and have more Ago hooks (bottom). Error bars represent standard deviation of values in each group. Repeat score as listed in Table 1. (c) Longer Ago-binding platforms also have greater average disorder scores. Trend lines and R^2 values from a linear regression.

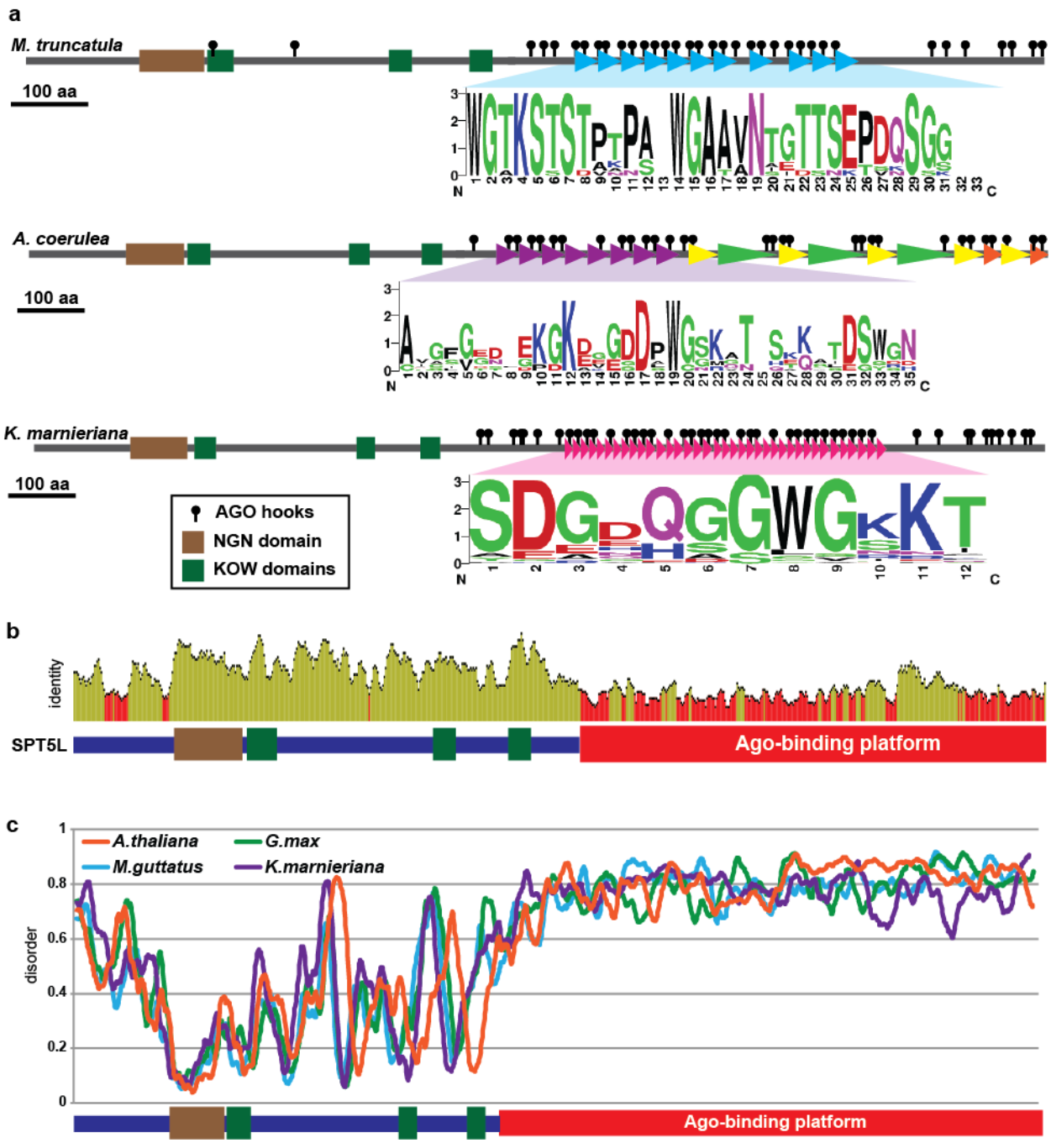

Figure S4. The Ago-binding platform of SPT5L is repetitive, non-conserved, and intrinsically disordered. (a) Representative SPT5L orthologs from across angiosperms demonstrate variation in the number, sequence, and conservation of tandem repeats in the Ago-binding platforms (colored triangles). (b) Identity between SPT5L orthologs from angiosperms is lowest in the C-terminal Ago-binding platform. After MUSCLE alignment columns with >20% gaps were stripped; a 10-amino acid sliding window of identity is displayed (above). The SPT5L domain structure is displayed below. (c) The Ago-binding platform is intrinsically disordered. IUPred disorder prediction for four SPT5L orthologs (*Arabidopsis thaliana*, *Glycine max*, *Mimulus guttatus*, *Kalanchoe marnieriana*); a 10-amino acid rolling average is plotted.

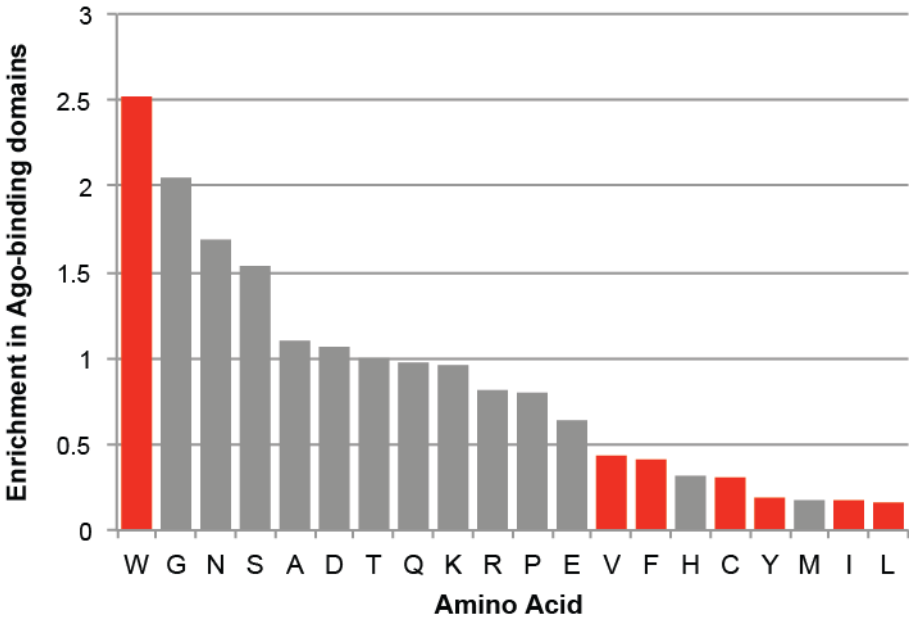

Figure S5. Depletion of order-promoting amino acids in Ago-binding platforms. Amino acids were arrayed by enrichment in Ago-binding platforms relative to other protein regions. Order-promoting amino acids are marked in red. Enrichment data from (Karlowski et al., 2010); order-promoting data from (Radivojac et al., 2007).

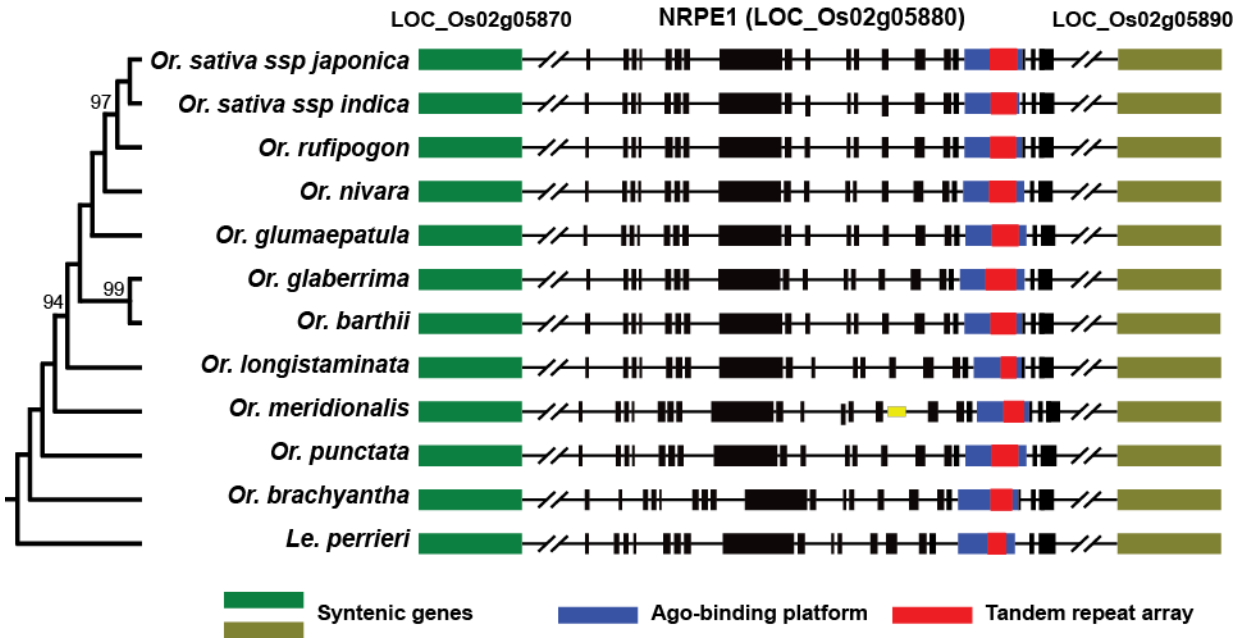

Figure S6. NRPE1a orthologs in the genus *Oryza*. Phylogenetic relationships between homologs are depicted on the left, based on a ML gene tree of the NRPE1a catalytic region. This gene tree recapitulates the expected species tree. Syntenic flanking genes are colored green and gold, with the *Or. sativa* gene identifier listed above. Boxes represent exons and lines represent introns. The Ago-binding platform is in blue, and the tandem repeats are in red. The small yellow box depicts missing sequence.

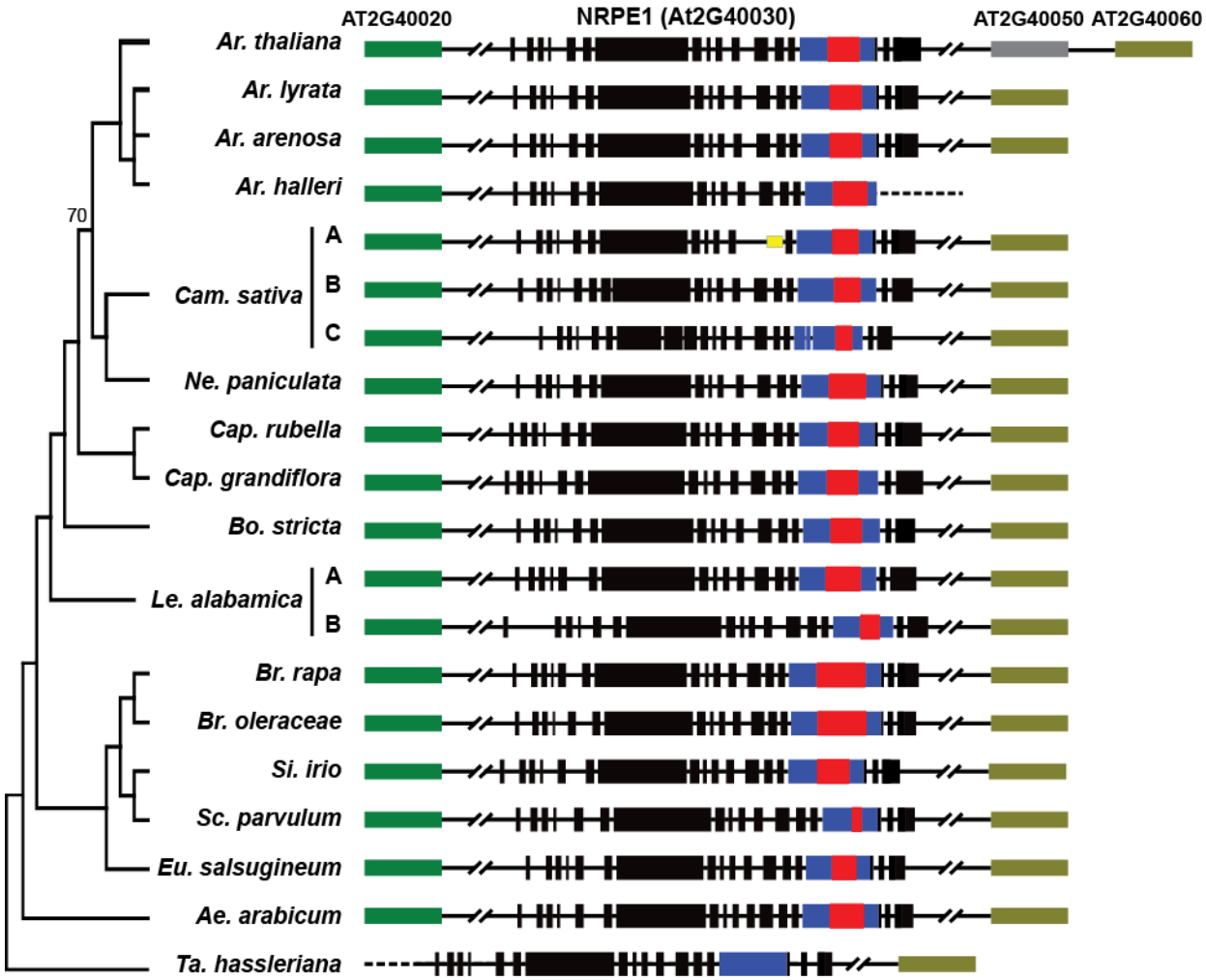

Figure S7. NRPE1 orthologs in the family Brassicaceae. Phylogenetic relationships between homologs are depicted on the left, based on a ML gene tree of the NRPE1 catalytic region. This gene tree recapitulates the expected species tree and includes recent polyploidization events in *Camelina sativa* and *Levenworthia alabamica*. All features are colored as in Figure S7. The *T. hassleriana* sequence is at the end of a contig and does not have a left flanking gene.

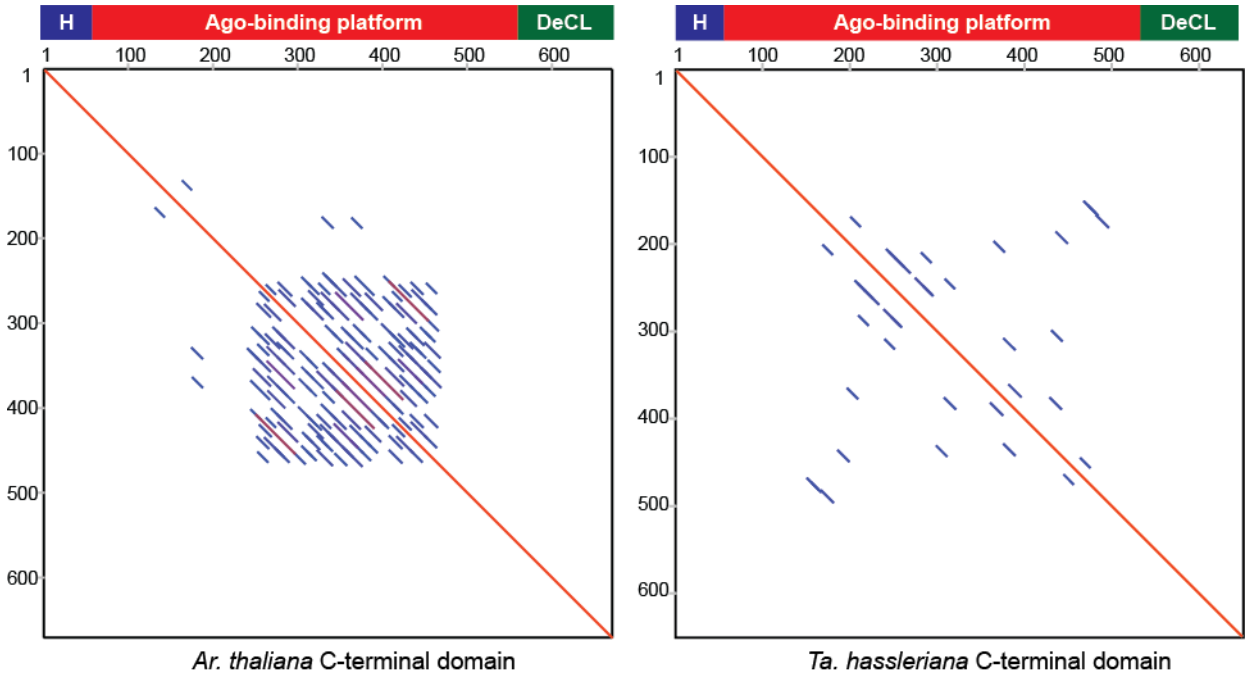

Figure S8. The *Ta. hassleriana* Ago-binding platform contains repetitive character. Self-alignment analysis (dot plots) demonstrates strong and weak repetitive character in *Ar. thaliana* (left) and *Ta. hassleriana* (right) carboxy terminal domains, respectively.

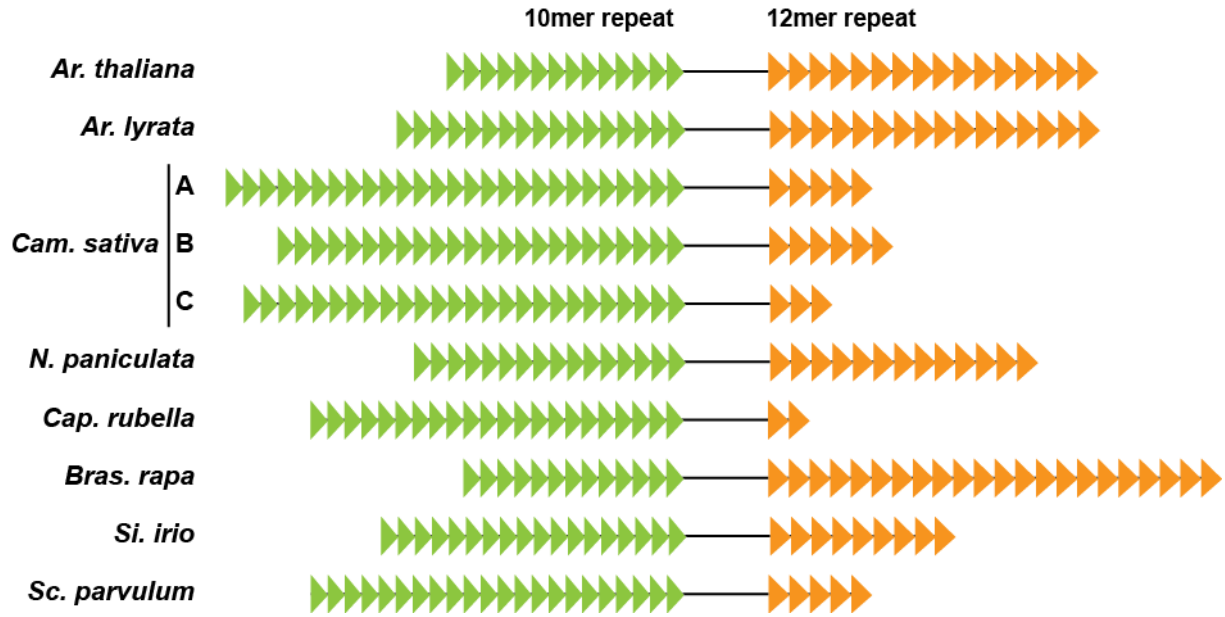

Figure S9. Variation in tandem repeats in the SPT5L Ago-binding platform. Diagram of tandem repeats from various Brassicaceae species, including three homeologs in *Camelina sativa* (A-C). All species have two distinct tandem repeat arrays found on separate exons and the sequence between the arrays is similar.

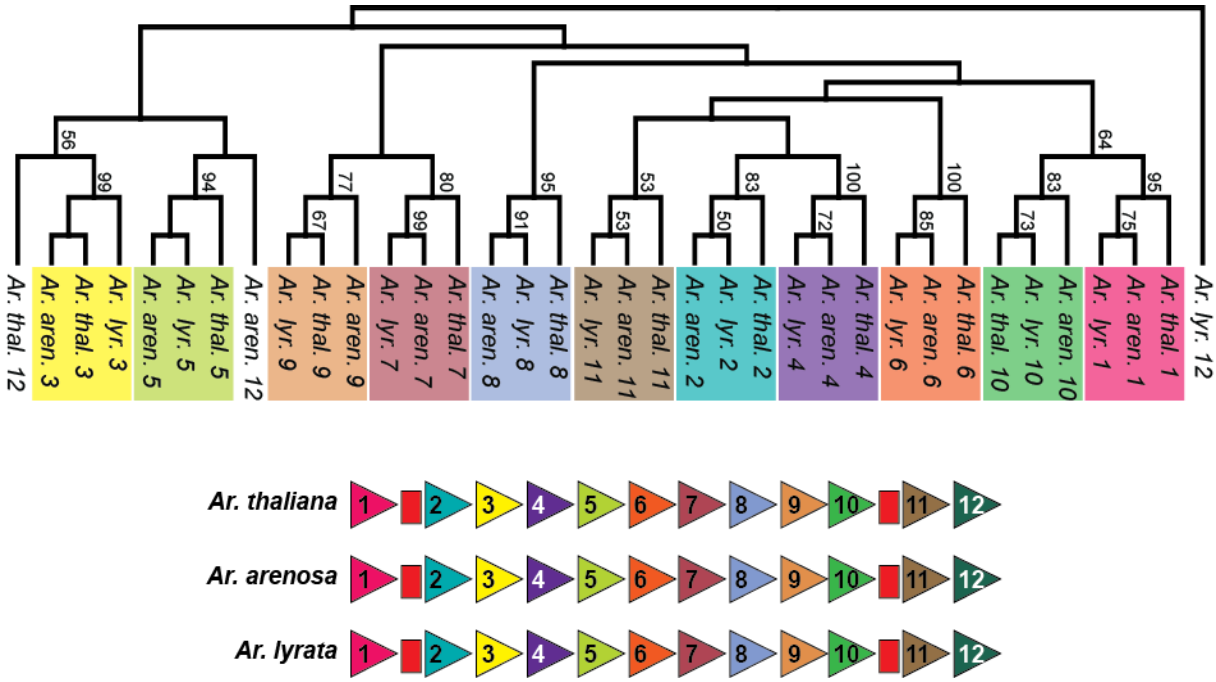

Figure S10. Phylogenetic analysis of repeats in the *Arabidopsis* genus demonstrates no changes in the NRPE1 repeat platform. (top) Cladogram based on ML tree of nucleotide sequence from individual repeats in *Ar. thaliana*, *Ar. arenosa*, and *Ar. lyrata* NRPE1 orthologs. Bootstrap values from 100 replicates are reported when ≥50. (bottom) Diagram of tandem repeats in these species colored by phylogenetic relationships.

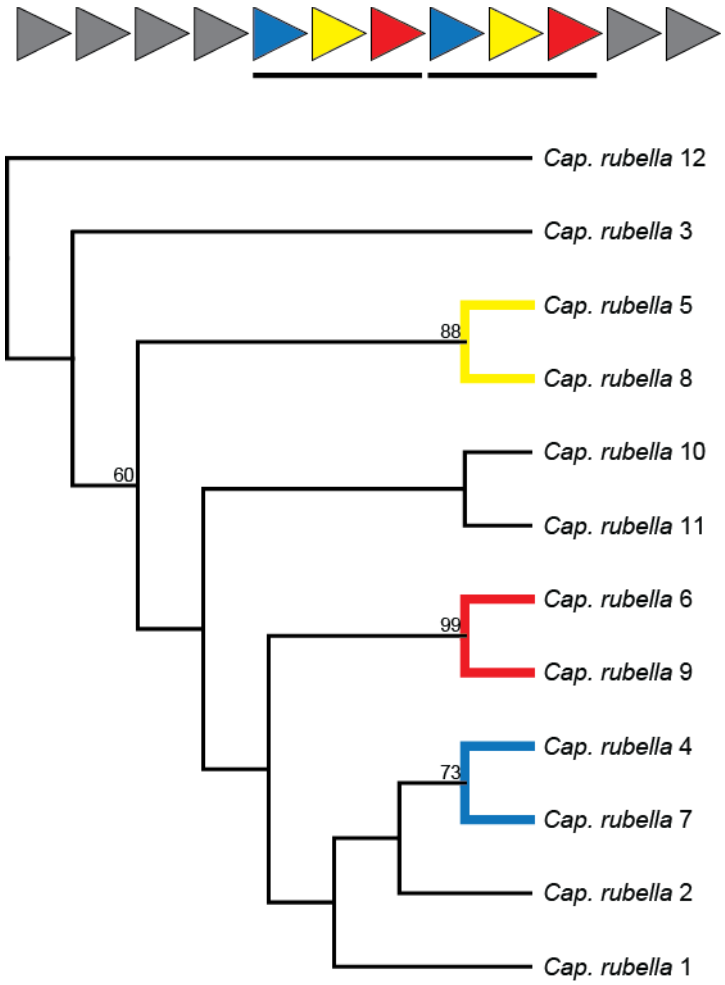

Figure S11. Phylogenetic analysis of *Capsella rubella* repeats supports a three-unit duplication. (top) Diagram of *Cap. rubella* tandem repeats colored by phylogenetic relationships. (bottom) Cladogram based on ML tree of individual repeats using nucleotide sequence. Bootstrap values from 100 replicates are reported when ≥50.

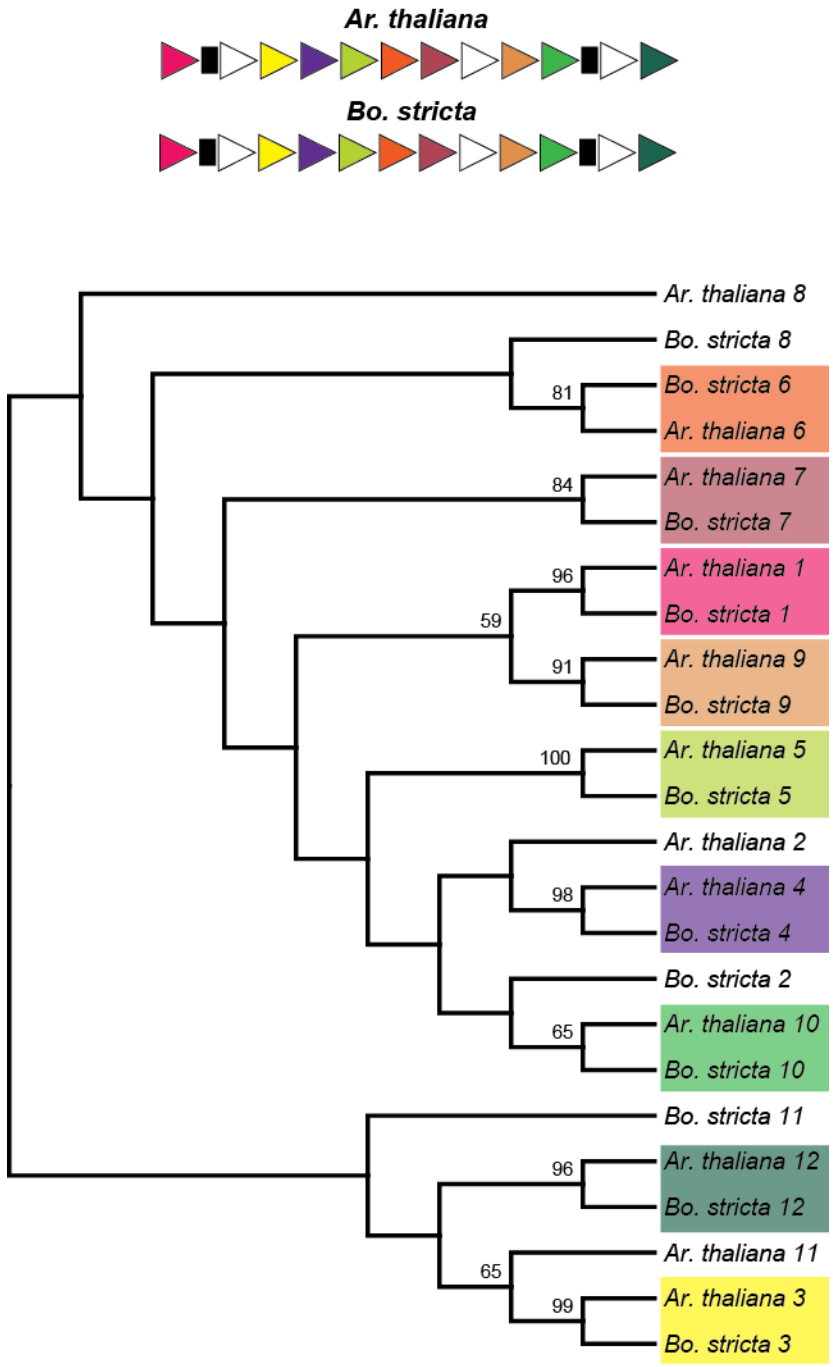

Figure S12. Phylogenetic analysis of *Boechera stricta* repeats demonstrates the ancestral repeat state. (top) Diagram of *Ar. thaliana* and *Bo. stricta* tandem repeat array showing identical placement of degenerate repeats. (bottom) Cladogram based on ML tree of individual *Ar. thaliana* and *Bo. stricta* repeat units shows colinearity of repeats. Bootstrap values from 100 replicates are reported when ≥50.

SUPPORTING INFORMATION NOTES

Note S1. CDS sequences from angiosperm NRPE1 sequences listed in Table S1.

Note S2. CDS sequences from angiosperm SPT5L sequences listed in Table S1.

Note S3. CDS sequences from *Orzya* NRPE1 sequences listed in Table S1.

Note S4. CDS sequences from Brassicaceae NRPE1 sequences listed in Table S1.

All SI notes are .fasta files and can be opened with any text editor or DNA analysis software.

REFERENCES

Karlowski WM, Zielezinski A, Carrère J, Pontier D, Lagrange T, Cooke R. **2010**. Genome-wide computational identification of WG/GW Argonaute-binding proteins in Arabidopsis. *Nucleic acids research* **38**: 4231–4245.

Radivojac P, Iakoucheva LM, Oldfield CJ, Obradovic Z, Uversky VN, Dunker AK. **2007**. Intrinsic disorder and functional proteomics. *Biophysical Journal* **92**: 1439–1456.