

Additional file 3: Number and divergence of expressed PsbO isoforms in the analyzed land plant species. Number of distinct *psbO* cDNA sequences found in EST databases and maximal number of different amino acid residues in mature proteins derived from these sequences.

	Family	Species	Number of PsbO isoforms	Maximal number of different amino acid residues between any two isoforms ^a
Bryophyta	Funariaceae	<i>Physcomitrella patens</i>	4	24
Lycopodiophyta	Selaginellaceae	<i>Selaginella moellendorffii</i>	1 ^b	
Cycadophyta	Cycadaceae	<i>Cycas rumphii</i>	1	
Coniferophyta	Cupressaceae	<i>Cryptomeria japonica</i>	1	
		<i>Sequoia sempervirens</i>	1	
	Pinaceae	<i>Picea abies</i>	1	
		<i>Picea sitchensis</i>	1	
		<i>Pinus sylvestris</i>	1	
		<i>Pinus taeda</i>	1	
Taxaceae	<i>Taxus baccata</i>	1		
monocots	Poaceae	<i>Hordeum vulgare</i>	1	
		<i>Oryza sativa</i>	1	
		<i>Triticum aestivum</i>	1	
		<i>Zea mays</i>	2	6
	Zingiberaceae	<i>Curcuma longa</i>	2	13
		<i>Zingiber officinale</i>	2	12
dicots	Asteraceae	<i>Artemisia annua</i>	2	13
	Brassicaceae	<i>Arabidopsis thaliana</i>	2	11
		<i>Brassica napus</i>	5	7
		<i>Brassica oleracea</i>	3	9
		<i>Brassica rapa</i>	4	7
		<i>Thellungiella halophila</i>	3	10 (41) ^c
	Euphorbiaceae	<i>Manihot esculenta</i>	2	14
	Fabaceae	<i>Glycine max</i>	4	13
		<i>Lotus japonicus</i>	2	15
		<i>Medicago truncatula</i>	2	12
		<i>Phaseolus vulgaris</i>	2	7
	Linaceae	<i>Linum usitatissimum</i>	1 ^d	
	Malvaceae	<i>Gossypium raimondii</i>	1	
		<i>Theobroma cacao</i>	1	
	Myrtaceae	<i>Eucalyptus grandis</i>	1	
	Phrymaceae	<i>Mimulus guttatus</i>	1	
	Rosaceae	<i>Fragaria vesca</i>	1	
		<i>Malus domestica</i>	2	11
	Rutaceae	<i>Citrus reticulata</i>	1	
		<i>Citrus sinensis</i>	1	
		<i>Citrus trifoliata</i>	1	
	Salicaceae	<i>Populus x canadensis</i>	2	23
		<i>Populus deltoides</i>	2	23
		<i>Populus trichocarpa x Populus deltoides</i>	2	22
		<i>Populus trichocarpa</i>	2	20
	Solanaceae	<i>Capsicum annuum</i>	2	18
		<i>Nicotiana benthamiana</i>	2	12
<i>Nicotiana tabacum</i>		4	14	
<i>Solanum habrochaites</i>		2	15	
<i>Solanum lycopersicum</i>		2	13	
<i>Solanum melongena</i>		2	21	
		<i>Solanum tuberosum</i>	2	12
Vitaceae	<i>Vitis vinifera</i>	2	17	

^a in mature form (without a transit peptide)

^b Two *psbO* sequences of *S. moellendorffii* can be found in EST databases. One of them (PlantGDB: PUT-165a-Selaginella_moellendorffii-12671) has far different mutation rate and seems to be a slightly expressed pseudogene, thus was omitted from analyses.

^c Two isoforms of *T. halophila* that are most similar to isoforms of *A. thaliana* have only 10 different amino acid residues, another, the most diverged isoform, is orthologous to a truncated pseudogene of *A. thaliana*.

^d *L. usitatissimum* expresses two *psbO* genes, but the corresponding mature proteins have the same amino acid sequence.