



Evolution of the land plant exocyst complexes

Fatima Cvrčková^{1*}, Michal Grunt¹, Radek Bezdova¹, Michal Hála^{1,2}, Ivan Kulich¹, Anamika Rawat¹ and Viktor Žárský^{1,2}

¹ Department of Experimental Plant Biology, Faculty of Sciences, Charles University, Prague, Czech Republic

² Institute of Experimental Botany, Academy of Sciences of the Czech Republic, Prague, Czech Republic

Edited by:

Markus Geisler, University of Fribourg, Switzerland

Reviewed by:

Jürgen Kleine-Vehn, University of Natural Resources and Life Sciences Vienna, Austria

Frantisek Baluska, University of Bonn, Germany

*Correspondence:

Fatima Cvrčková, Department of Experimental Plant Biology, Faculty of Sciences, Charles University, Viničná 5, CZ 128 44 Praha 2, Czech Republic.
e-mail: fatima.cvrckova@natur.cuni.cz

Exocyst is an evolutionarily conserved vesicle tethering complex functioning especially in the last stage of exocytosis. Homologs of its eight canonical subunits – Sec3, Sec5, Sec6, Sec8, Sec10, Sec15, Exo70, and Exo84 – were found also in higher plants and confirmed to form complexes *in vivo*, and to participate in cell growth including polarized expansion of pollen tubes and root hairs. Here we present results of a phylogenetic study of land plant exocyst subunits encoded by a selection of completely sequenced genomes representing a variety of plant, mostly angiosperm, lineages. According to their evolution histories, plant exocyst subunits can be divided into several groups. The core subunits Sec6, Sec8, and Sec10, together with Sec3 and Sec5, underwent few, if any fixed duplications in the tracheophytes (though they did amplify in the moss *Physcomitrella patens*), while others form larger families, with the number of paralogs ranging typically from two to eight per genome (Sec15, Exo84) to several dozens per genome (Exo70). Most of the diversity, which can be in some cases traced down to the origins of land plants, can be attributed to the peripheral subunits Exo84 and, in particular, Exo70. As predicted previously, early land plants (including possibly also the Rhyniophytes) encoded three ancestral Exo70 paralogs which further diversified in the course of land plant evolution. Our results imply that plants do not have a single “Exocyst complex” – instead, they appear to possess a diversity of exocyst variants unparalleled among other organisms studied so far. This feature might perhaps be directly related to the demands of building and maintenance of the complicated and spatially diverse structures of the endomembranes and cell surfaces in multicellular land plants.

Keywords: exocyst, phylogeny, land plants, co-evolution, gene duplication

INTRODUCTION

Exocyst, or the Sec6/8 complex, is an evolutionarily conserved heterooligomeric protein complex, generally believed to function especially in the last stage of exocytosis – i.e., vesicle tethering, preceding fusion of trans-Golgi network-derived vesicles with the plasmalemma, although additional, also mostly vesicle trafficking-related, exocyst roles have been described (reviewed, e.g., in He and Guo, 2009; Zhang et al., 2010; Heider and Munson, 2012). The eight canonical exocyst subunits, Sec3, Sec5, Sec6, Sec8, Sec10, Sec15, Exo70, and Exo84, were originally identified in yeast (Ter-Bush et al., 1996; Guo et al., 1999). Subsequently, their homologs were found also in metazoans (Guo et al., 1997; Kee et al., 1997) and higher plants (Eliáš et al., 2003). Angiosperm exocyst subunits form complexes *in vivo* (Hála et al., 2008), and participate in exocytosis- or vesicle trafficking-dependent processes, such as cell growth including both tip growth and diffuse surface expansion (Cole et al., 2005; Wen et al., 2005; Synek et al., 2006; Hála et al., 2008), cell division (Fendrych et al., 2010), delivery of materials to the periplasm and cell wall (Wang et al., 2010), biogenesis of specialized cell wall structures such as the myxosperm seed coat (Kulich et al., 2010), pathogen response (Pečenkova et al., 2011), and mycorrhiza (Genre et al., 2012). The Exo70 subunit has been also previously implicated in the pollen-stigma interaction

in *Brassica* and *Arabidopsis* (Samuel et al., 2009), though its specific role remains controversial (Kitashiba et al., 2011) and the observed phenotypes may be rather due to a generalized secretion defect affecting stigma function (Synek et al., 2006).

Exocyst belongs, together with related COG, GARP, and DSL1 complexes, to the large, evolutionarily ancient family of eukaryotic quatrefoil vesicle tethering complexes (Whyte and Munro, 2002; Koumandou et al., 2007). Structural studies (recently reviewed by Hertzog and Chavrier, 2011) and theoretical sequence-based modeling revealed common structural elements involving rod-like helical bundles in all eight subunits, and a model of exocyst architecture based on aggregation of these bundles has been proposed (Munson and Novick, 2006; Croteau et al., 2009). Electron microscopy observations consistent with this model have been made also in the case of the putative plant exocyst (Seguí-Simmaro et al., 2004). Bundled Sec6, Sec8, Sec10 subunits probably form a core of the complex. At least in the yeast model, Sec6 also participates in its anchoring to the target membrane, and the remaining, more peripherally located subunits mediate interactions with membrane vesicles destined for delivery (as in the case of Sec15, interacting with the vesicle-borne Sec4 GTPase), with the target membrane and associated small GTPases of the Rho family (Sec3 and Exo70), and possibly with other structural or

regulatory proteins (Songer and Munson, 2009). The Exo70 subunit, which can bind to phosphoinositides, is crucial for targeting the complex to the destination membrane also in metazoans (He et al., 2007). Exo84 is also required for proper localization of the exocyst in yeast (Zhang et al., 2005). Surprisingly, the function of these subunits is not restricted to participation in exocytosis, as Exo70 and Exo84 subunits also participate in pre-mRNA splicing (Awashi et al., 2001; Dellago et al., 2011).

While exocyst subunits are encoded by a single gene in yeast or at most a few paralogs in metazoans, a puzzling number of plant isoforms has been identified in particular for the Exo70 subunit, which is encoded by 23 distinct loci in *Arabidopsis thaliana* (Eliáš et al., 2003; Synek et al., 2006). Some other subunits are also encoded by duplicated or triplicated (as in case of *A. thaliana* Exo84) loci. However, the only published phylogenetic studies of the plant exocyst so far are devoted solely to the Exo70 subunit (Eliáš et al., 2003; Synek et al., 2006) or restricted to a very limited species selection (Chong et al., 2010). With growing number of sequenced genomes, and increasing quality of genomic sequence annotations, a broader coverage of plant lineages can now be achieved. Here we present the results of a phylogenetic analysis of the canonical exocyst subunits encoded by 10 land plant genomes representing dicot and monocot angiosperms, a lycophyte (*Selaginella moellendorffii*) and a moss (*Physcomitrella patens*), and propose an evolutionary scenario consistent with our results.

MATERIALS AND METHODS

IDENTIFICATION OF EXOCYST SUBUNIT SEQUENCES

The collection of exocyst subunit sequences has been assembled by exhaustive mining of multiple data sources. For each subunit, a “seed” collection was generated as a non-redundant union of sequences originating from *A. thaliana*, *Arabidopsis lyrata*, *Populus trichocarpa*, *Vitis vinifera*, *Oryza sativa* var. *japonica*, *O. sativa* var. *indica* (omitted in case of Exo70 to keep the project at a manageable scale), *Sorghum bicolor*, *Brachypodium distachyon* and *P. patens*, and identified on the basis of their annotation among (i) components of the exocyst complex as recorded in the COG section of the STRING protein interaction database¹ (Skłarczyk et al., 2011) and (ii) reference sequences from GenBank (Benson et al., 2012). BLAST (McGinnis and Madden, 2004) searches of species-specific portions of the non-redundant section of GenBank and several species-specific resources (see below) have been employed to identify additional sequences from the above listed species, as well as from *S. moellendorffii* and selected members of the genus *Solanum* (see Results).

The additional databases mined included Uniprot (The Uniprot Consortium, 2012), Phytozome² (Goodstein et al., 2012), and JGI³ for multiple species, Solgenomics⁴ (Bombarely et al., 2011) and PGSC⁵ (Potato Genome Sequencing Consortium, 2011) for *Solanaceae*, The Arabidopsis Information Resource⁶ (Lamesch

et al., 2012) for *Arabidopsis*, and COSMOSS⁷ (Lang et al., 2005) for *Physcomitrella*. Final round of searches was performed between February and May 2012.

Redundancies within the collection were removed on the basis of pairwise BLAST alignments. In case of multiple protein predictions originating from the same locus, protein sequences closest to the most frequent splicing variety were chosen. In some cases, predicted protein sequences were revised based on re-evaluation of the available gene models, taking into account multiple methods of splicing prediction, ESTs, and homologous sequences as described previously (Grunt et al., 2008). The complete collection of sequences including the revised ones is available in the Supplement.

Additional BLAST searches of non-redundant GenBank sequences were performed to identify homologs of outlier sequences as described in Results.

PROTEIN SEQUENCE ALIGNMENTS

For initial estimation of sequence similarity and detection of possible problems with gene structure prediction (i.e., missing or extraneous exons), the interactive MACAW tool (Schuler et al., 1991; Lawrence et al., 1993), or the automated tools ClustalX (Thompson et al., 1997) and KALIGN (Lassmann and Sonnhammer, 2006) have been employed to generate preliminary versions of multiple protein sequence alignments. Final alignments for all subunits except Exo70 have been constructed manually with the aid of BioEdit (Hall, 1999), taking into account the preliminary alignments.

In case of the more numerous and more diverse Exo70 sequences, a similar manual approach has been employed first with a complete collection of *A. thaliana*, *A. lyrata*, and *P. trichocarpa* sequences, resulting in a “skeleton” alignment into which additional sequences in batches of up to 10 have been merged using the “realign selected sequences” feature of ClustalX; the alignments were manually adjusted after each batch using BioEdit with similarity shading for guidance, where considered appropriate.

Because of the admittedly subjective method of alignment construction, we are including the final alignments that have been used for phylogeny reconstruction in the Supplement. We have also performed parallel phylogeny estimations (as described below) with a manually constructed alignment and a KALIGN-constructed one for the Exo84 subunit, producing trees of essentially identical topology (i.e., sharing all significant branches, though differing somewhat in branch length and bootstrap support).

To identify conserved motifs in the divergent Exo70 N-termini, N-terminal sequence portions upstream of the conserved part used in phylogenetic analysis (see below) have been aligned *de novo* using ClustalX. Conserved sequence motifs have been identified visually after removal of obviously non-aligned sequences, manually adjusted in BioEdit and colored using the Dayhoff matrix (as implemented in BioEdit) for presentation.

PHYLOGENETIC ANALYSES

For phylogram construction, alignments except Exo70 were stripped of all columns containing gaps. For Exo70, which is more

¹<http://string-db.org/>

²<http://www.phytozome.net/>

³<http://genome.jgi.doe.gov/>

⁴<http://solgenomics.net/>

⁵<http://potatogenomics.plantbiology.msu.edu/>

⁶<http://www.arabidopsis.org>

⁷<http://www.cosmoss.org/>

divergent than the remaining subunits (especially in its N-terminal part) and where several sequences were C- or N-truncated, only the unreliably aligned N-terminal portion and regions containing gaps in multiple sequences were removed prior to phylogenetic tree calculation.

Trees were computed by the maximum likelihood (ML) method using PHYML v3.0 aLRT (Guindon and Gascuel, 2003; Anisimova and Gascuel, 2006) at Phylogeny.fr (Dereeper et al., 2008) with default settings, using the aLRT test to estimate internal branch reliability. Independently, phylograms were constructed also by the neighbor-joining (NJ) method using ClustalX with 1000 bootstrap samples. Trees were visualized with the aid of the MEGA5 software (Tamura et al., 2011) and manually colored using CorelDraw for presentation.

K_a/K_s ESTIMATIONS

Nucleotide sequences corresponding to selected Exo70 subunits (see Results) have been retrieved from GenBank, and portions corresponding to reliably aligned protein parts have been realigned manually using BioEdit in the “toggle nucleotide to protein” mode to re-create the protein alignment used to calculate the phylogenetic trees. Resulting nucleotide sequence alignments have been analyzed using Selecton (Stern et al., 2007) to obtain codon-specific values of non-synonymous to synonymous mutation rates, providing information on residue-specific selection in the history of the examined sequences.

RESULTS

AN INVENTORY OF EXOCYST SUBUNITS IN 10 PLANT SPECIES

We performed exhaustive searches of sequenced genomes of eight angiosperm and two non-seed plant species with the aim to identify all genes encoding putative exocyst subunits. Among the angiosperms, we included the eudicots *A. thaliana*, *A. lyrata*, poplar (*P. trichocarpa*), and grapevine (*V. vinifera*) as representatives of the rosids. To gain insight also into the asterid lineage, we attempted to find the exocyst subunits in the publicly available tomato (*Solanum lycopersicon*) genome and cDNA sequences, which, however, did not yet cover the complete genome at the time

of analysis. In particular, we found no sequences corresponding to Sec5 and Sec8. We therefore located the missing subunits in data from two potato species (*S. phureja* and *S. tuberosum*, respectively); we shall refer to these asterids collectively as *Solanum* sp. From the monocot class, four grass species (rice – *O. sativa*, represented by both *japonica* and *indica* varieties, sorghum – *S. bicolor*, and the model grass *B. distachyon*) have been included. Finally, we also analyzed genome data from one “lower” vascular plant – the lycophyte *S. moellendorffii*, and from the model moss *P. patens*. In total, we have collected 392 distinct protein sequences corresponding to presumed exocyst subunits (Table 1).

In agreement with the expected essential character of the exocyst in plants and with previous reports, all genomes encoded at least one copy of each subunit, and most of the subunits were encoded by one or a few loci, except Exo70, which always formed an extensive family of paralogs. Among the remaining subunits, we could upon closer inspection distinguish genuine single-copy or low copy subunits that were never present in more than two versions in the vascular plants (this was the case for Sec3, Sec5, Sec6, Sec8, and Sec10), and intermediate size gene families with more than two and less than eight paralogs in at least one of the species (Sec15 and Exo84). We shall further discuss these three groups separately.

LOW COPY SUBUNITS: SEC3, SEC5, SEC6, SEC8, AND SEC10

The first group of subunits includes Sec3 (in *A. thaliana* encoded by two genes in tandem – AtSEC3A/At1g47550/Arath1_Sec3 and AtSEC3B/At1g47560/Arath2_Sec3), Sec5 (again two *A. thaliana* genes – AtSEC5A/At1g76850/Arath1_Sec5 and AtSEC5B/At1g21170/Arath2_Sec5), Sec6, Sec8, and Sec10 (all encoded by single genes in *A. thaliana* – AtSEC6/At1g71820/Arath_Sec6, AtSEC8/At3g10380/Arath_Sec8 and AtSEC10/At5g12370/Arath_Sec10 – but see comments on possible Sec10 duplication below). Though these subunits are single-copy in some species, each of them is duplicated in at least one angiosperm genome, and all but Sec6 are triplicated in *P. patens*, showing that there is no strict functional requirement on keeping only a single protein version in cells. In fact, multiple splicing variants have been proposed

Table 1 | Numbers of exocyst subunit paralogs encoded by the studied plant genomes.

	Sec3	Sec5	Sec6	Sec8	Sec10	Sec15	Exo70	Exo84
<i>A. thaliana</i>	2	2	1	1	1	2	23	3
<i>A. lyrata</i>	2	2	1	1	1	2	23	3
<i>P. trichocarpa</i>	2	2	2	2	2	5	29	8
<i>Solanum</i> sp.	2 ²	1 ³	1 ²	1 ⁴	1 ²	2 ²	22 ²	4 ²
<i>V. vinifera</i>	1	1	2	1	1	2	15	3
<i>O. sativa</i> ¹	2(2)	1(1)	1(1)	1(1)	1(1)	4(4)	47	3(3)
<i>S. bicolor</i>	2	1	1	1	1	3	31	3
<i>B. distachyon</i>	2	1	1	1	1	3	27	3
<i>S. moellendorffii</i>	2	1	2	2	2	1	8	2
<i>P. patens</i>	3	3	1	3	3	2	13	7

The complete list of the 392 analyzed genes or proteins including database accession numbers, as well as protein sequences and sequence alignments used in phylogeny calculations, is provided as Supplementary Material.

¹*japonica* variety, with numbers for *indica* in brackets; ²*S. lycopersicon*; ³*S. phureja*; ⁴*S. tuberosum*.

for most *Arabidopsis* subunits in the recent genome annotation (Lamesch et al., 2012).

Tandem duplications affecting angiosperm exocyst genes are apparently not restricted to *Arabidopsis* Sec3. The *A. thaliana* genomic assembly might be problematic in the area around Sec10, since inspection of available GenBank sequences suggests a possible tandem duplication of the Sec10 locus differing by a couple of silent mutations and variant non-coding ends. The duplicated gene appears to be transcribed (see GenBank cDNAs AF479280.1 and AK318699.1 which are in good mutual agreement but differ from the reference genome sequence, though they encode an identical protein). Also in tomato, we found a single possibly functional Sec10 locus and three closely related pseudogenes with multiple stop codons, two of them in tandem (the pseudogenes are not included in the phylogeny; see Supplementary Material for accession numbers).

As a rule, protein sequences of the low copy subunits consist of a single well-defined domain, are well conserved along the whole length (exceptions will be discussed below) and their phylogenetic trees (Figure 1) exhibit striking overall mutual similarity. Within the angiosperms, all gene duplications except monocot Sec3 appear to be relatively recent, resulting in within-species paralogs that share at least 80% of identical amino acids in the most distant pair of the *A. thaliana* Sec5 paralogs. Duplicated paralogs cannot be matched among genomes more distant than the two rice varieties, or the two *Arabidopsis* species. The only exception from this pattern of apparently late gene duplications is the Sec3 subunit that has obviously split into two paralogous lineages early in the evolution of monocots or at least grasses.

Rice and *Arabidopsis* versions of any of the low copy subunits share between 59% (Arath1_Sec5 vs. OrysaJ_Sec5) and 81% (Arath_Sec6 vs. Orysa_Sec6) of identical amino acids. The Sec6, Sec10, and Sec8 subunits, believed to form the central core of the complex (Munson and Novick, 2006; Croteau et al., 2009), are the best conserved ones. Notably, one of the ancient Sec3 branches (the clade “monocot 1” in Figure 1) has considerably diverged from the cluster of dicot sequences and the remaining monocot clade, suggesting a possible release of selection pressure followed by neo- or subfunctionalization. Compared to the degree of conservation found in the angiosperms, the *Physcomitrella* paralogs exhibit major within-genome differences, with the most distant paralogs Phypa1_Sec10 and Phypa3_Sec10 sharing only 51% of identical amino acids.

Two sequences deviate from the standard overall conserved domain structure of the relevant low copy subunits and can be perhaps viewed as “structural outliers” of their corresponding gene families. In the case of the *A. lyrata* Sec3 paralog Araly2_Sec3, the N-terminal part of the conserved domain is replaced by a domain related to a family of RING box/E3 ligases, encoded by a single-exon and flanked at least from one side by a sequence related to Copia-like retroelements, suggesting a very recent retrotransposition-mediated gene fusion. However, this domain combination appears to be unique in the whole of GenBank, and there are no ESTs documenting that this gene is expressed *in planta*; therefore, its functionality and biological significance remains problematic.

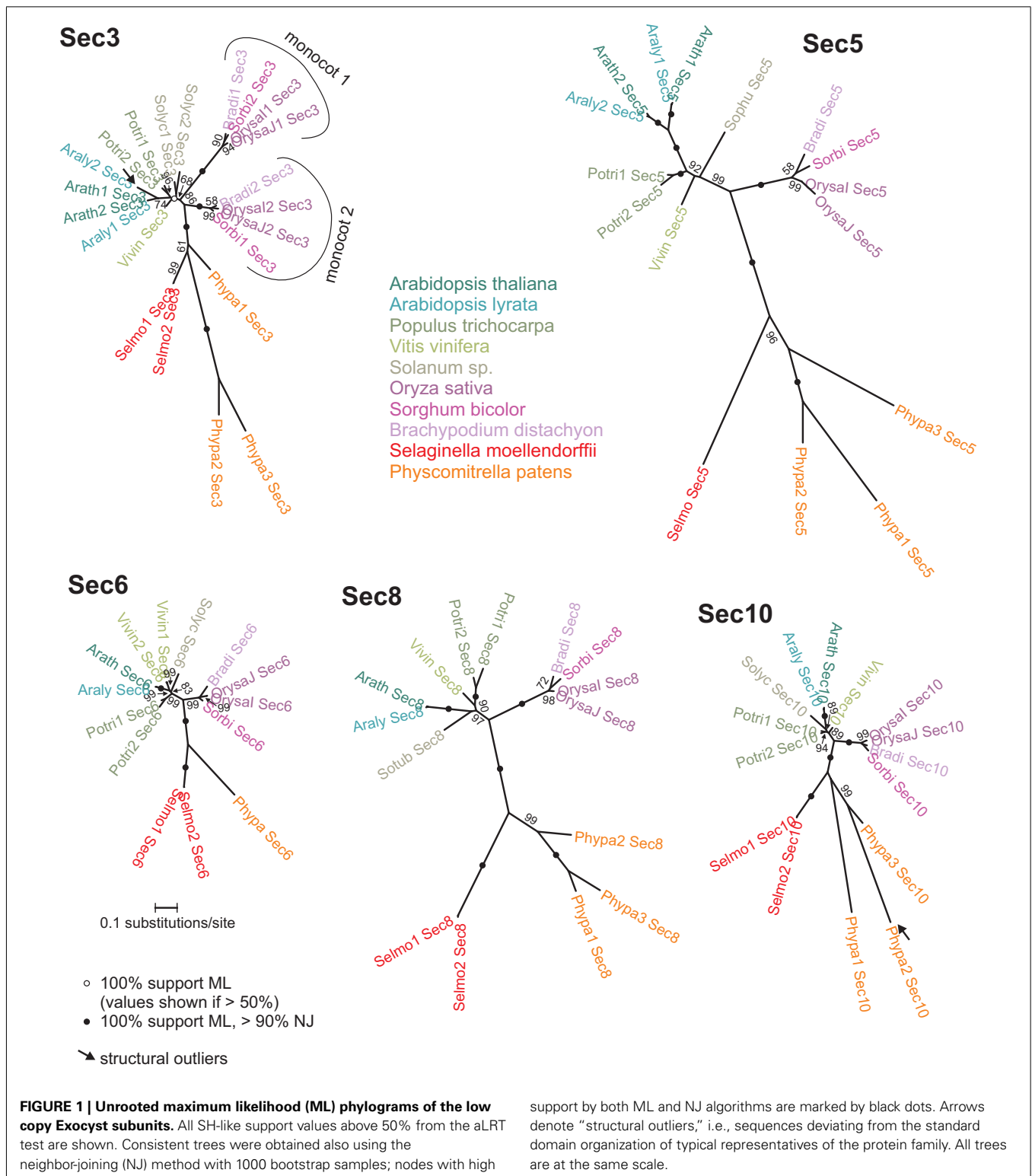
The second structural outlier is the *P. patens* Sec10 paralog Phypa2_Sec10, noticed in our previous study (Grunt et al., 2008) because of its unique combination of a N-terminally located Sec10 domain with the formin-specific FH2 domain at the C-end of the protein (see sequence Phypa5 in Grunt et al., 2008). An alternative splicing prediction separates these two domains into two distinct proteins (a short version corresponding to standard Sec10 is included in our phylogeny). The combination of Sec10 and FH2 domains is again unique in GenBank. However, there is a partial cDNA of the formin end (GenBank BY987890.1) indicating gene expression in the moss, albeit it is unclear which splice variants are biologically relevant.

INTERMEDIATE SIZE FAMILIES: SEC15 AND EXO84

The second group of subunits consists of two gene families, Sec15 (with two *A. thaliana* paralogs – AtSEC15A/At3g56640/Arath1_Sec15 and AtSEC15B/At4g02350/Arath2_Sec15) and Exo84 (with three paralogs in *A. thaliana* – AtEXO84a/At5g49830/Arath2_Exo84, AtEXO84b, At1g10385/Arath3_Exo84, and AtEXO84c/At1g10180/Arath1_Exo84). In other studied genomes, Sec15 is encoded by two to five subunits (except *S. moellendorffii*, where only a single protein was found) and Exo84 by three to eight (again except *S. moellendorffii* with only two genes). In both cases the highest number was found in *P. trichocarpa*, and the final poplar subunit count may be even higher, since there is cDNA evidence of additional transcripts encoding proteins identical to the Sec15 paralogs included in our analysis but differing in their non-translated ends, reminiscent of the situation in *A. thaliana* Sec10 (see Supplementary Material).

Phylogenetic trees of both families indicate that at least a part of the observed diversity is ancient, and can be traced back at least to the origins of angiosperms (Figure 2). Both gene families can be split into two branches in seed plants, with multiple additional within-branch amplifications. In Sec15, most of these later amplification events (generating two clusters of poplar genes in both branches and a pair of rice genes in the B branch) appear to be fairly recent, reminiscent of those detected for low copy subunits. However, a duplication of the A subunit apparently occurred early in the monocot lineage (no later than at the emergence of grasses), resulting in two monocot- or grass-specific subfamilies, A1 and A2.

In Exo84, the situation is somewhat more complex. Clearly defined A and B branches (named according to the corresponding *A. thaliana* subunits) were found only in the dicots, while related monocot sequences form a rather compact cluster, probably closer to the A branch than to B. The dicot A and B branches and the monocot cluster will be further referred to as the A/B clade (Figure 2). Monocot A/B sequences also bear traces of early gene duplication preceding the radiation of grasses, but clearly distinct from the event that produced the A and B branches and followed by little actual sequence divergence (the rice A/B sequences OrysaJ1_Exo84 and OrysaJ2_Exo84 share 78% of identical amino acids). Besides of the A/B clade, a second ancient branch (the C clade) is shared by all examined angiosperms and contains, as a rule, products of single-copy genes with exception of an apparently recent cluster of three poplar sequences. Two mutually related dicot outliers (the CX sequences) are apparently related to the C clade, and proteins similar to them have been predicted also in



Ricinus communis (GenBank XM_00252625.1) and *Glycine max* (GenBank XM_003541318.1). A genomic DNA sequence fragment from *V. vinifera* (GenBank AM443616.2) contains patches of a possible ORF similar to the CX sequences in an area annotated as non-coding. While the grapevine genome annotation may require

updating, it is also possible that these patches are vestiges of a lost gene that may have had a wider distribution.

In contrast to the angiosperms, the lycophyte and moss Sec15 sequences exhibit only minimum diversification, while Exo84 underwent duplication in *S. moellendorffii* and extensive

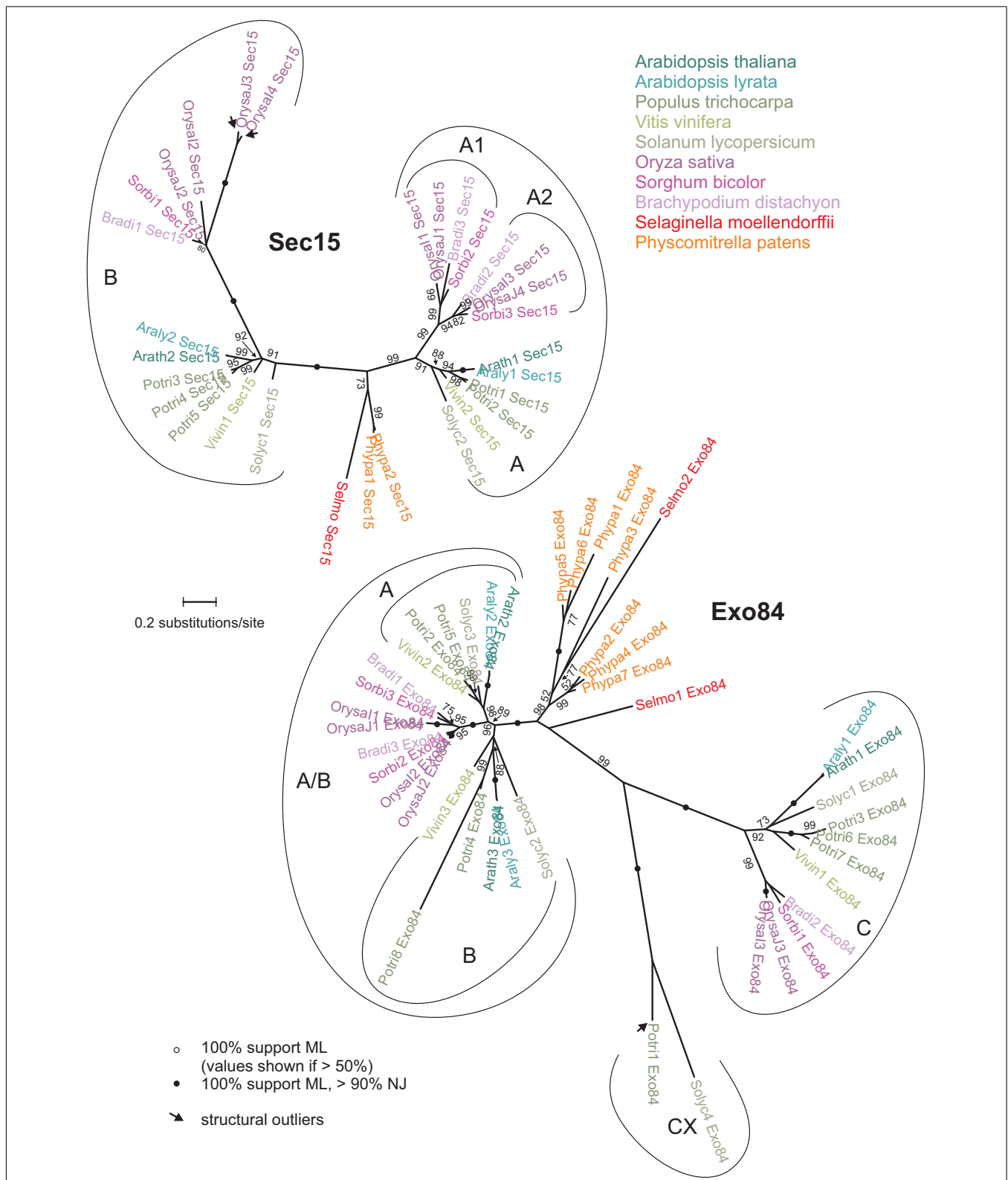


FIGURE 2 | Unrooted maximum likelihood (ML) phylogenies of Sec15 and Exo84. All SH-like support values above 50% are shown. Consistent trees were obtained also using the neighbor-joining (NJ) method; nodes

with high support by both ML and NJ algorithms are marked by black dots. Arrows denote “structural outliers” (see **Figure 1**). Both trees are at the same scale.

amplification, producing seven rather diversified paralogs, in *P. patens*. While the branching order is not reliable in the non-angiosperm sequences, the resulting tree does not exclude the possibility that the two major Exo84 clades might have appeared already at the base of the vascular plants.

Compared to the low copy subunits, Sec15 and Exo84 sequences exhibit greater diversity, with *Arabidopsis* AtSec15A and AtSec15B sharing 48%, AtExo84A and AtExo84B 59%, and AtExo84B and AtExo84C only 34% of identical amino acids. Nevertheless, the angiosperm branches of the phylogram appear to be rather compact, and the only conspicuously diversified poplar Exo84B paralog, Potri8_Exo84, may not be expressed, as we could not find any corresponding ESTs.

All sequences from each family can be aligned reliably along the whole length, with only three exceptions. The predicted rice Sec15 paralogs OrysaJ3_Sec15 and OrysaI4_Sec15 are missing a C-terminal part of the characteristic Sec15 domain and have instead an unrelated sequence. No homologs with such a gene organization have been found in GenBank, and there are no ESTs matching these genes. Together with the long distance from the rest of B clade Sec15 sequences, indicating relaxed selection, this suggests that these *O. sativa* Sec15 outliers may actually correspond to a pseudogene that has arisen not long before the separation of the *japonica* and *indica* varieties and that is now in the process of decay. The third structural outlier, Potri1_Exo84, one of the outlier CX sequences with a long C-terminal extension, also lacks cDNA or EST support, and it is thus not clear if it is expressed at all.

THE ENORMOUS DIVERSITY OF EXO70 PARALOGS

The large Exo70 family consists of 23 paralogs in *A. thaliana*, representing eight previously identified clades (Synek et al., 2006): AtExo70A1/At5g03540/ArathA1_Exo70, AtExo70A2/At5g52340/ArathA2_Exo70, and AtExo70A3/At5g52350/ArathA3_Exo70 in clade A, AtExo70B1/At5g58430/ArathB1_Exo70 and AtExo70B2/At1g07000/ArathB2_Exo70 in clade B, AtExo70C1/At5g13150/ArathC1_Exo70 and AtExo70C2/At5g13990/ArathC2_Exo70 in clade C, AtExo70D1/At1g72470/ArathD1_Exo70, AtExo70D2/At1g54090/ArathD2_Exo70, and AtExo70D3/At3g14090/ArathD3_Exo70 in clade D, AtExo70E1/At3g29400/ArathE1_Exo70 and

AtExo70E2/At5g61010/ArathE2_Exo70 in clade E, AtExo70F1/At5g50380/ArathF1_Exo70 in clade F, AtExo70G1/At4g31540/ArathG1_Exo70 and AtExo70G2/At1g51640/ArathG2_Exo70 in clade G, and eight paralogs – AtExo70H1/At3g55150/ArathH1_Exo70, AtExo70H2/At2g39380/ArathH2_Exo70, AtExo70H3/At3g09530/ArathH3_Exo70, AtExo70H4/At3g09520/ArathH4_Exo70, AtExo70H5/At2g28640/ArathH5_Exo70, AtExo70H6/At1g07725/ArathH6_Exo70, AtExo70H7/At5g59730/ArathH7_Exo70, and AtExo70H8/At2g28650/ArathH8_Exo70 – in clade H. In other studied plants, the number ranges from eight in *Selaginella* to 47 in rice (Table 2; see Supplementary Material for a full list of genes), albeit the final count might still change in the genomes whose annotation is still under development (especially *Solanum* sp.).

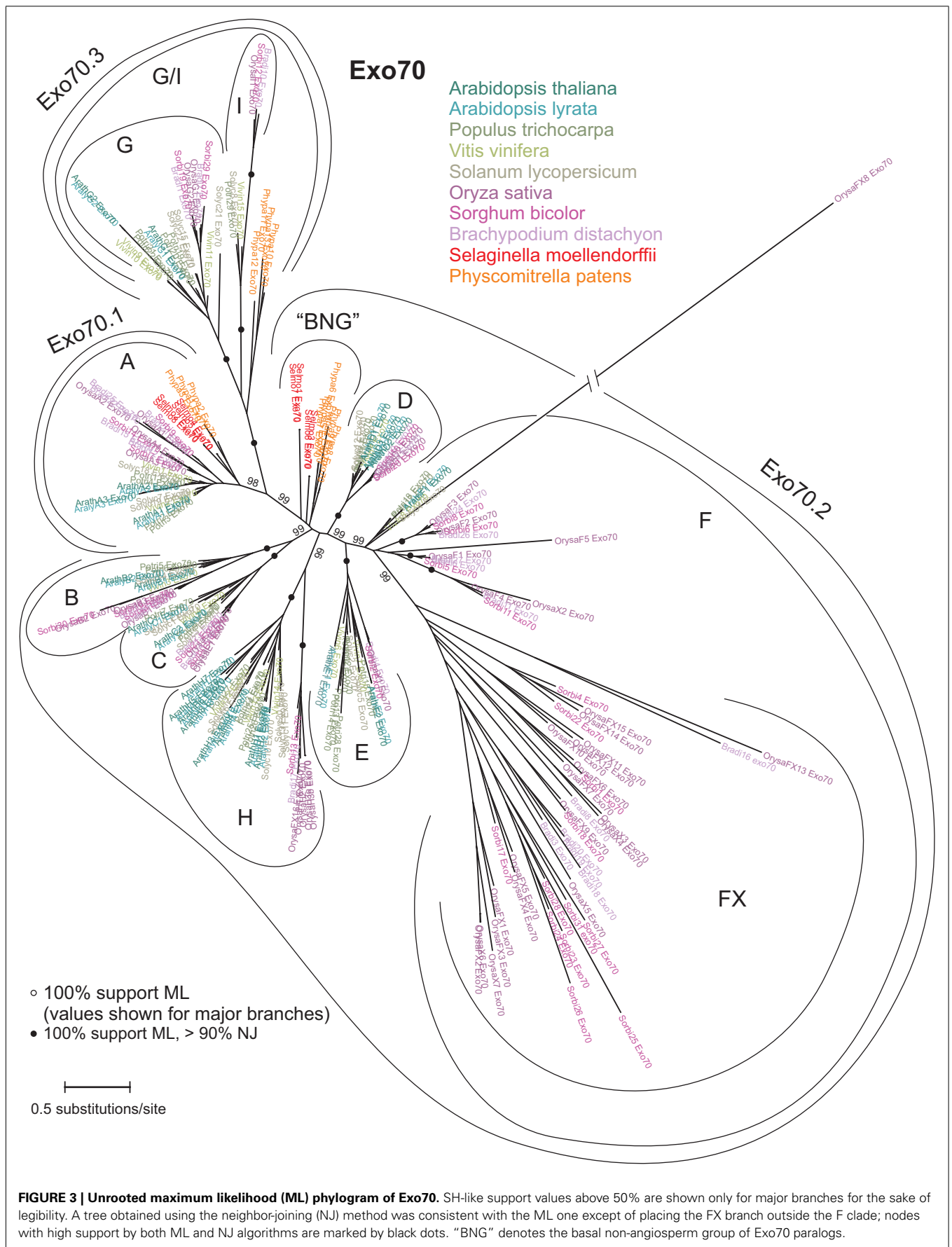
Unlike the other seven subunits, Exo70 paralogs are rather diverse and their N-terminal part of up to 300 amino acids could not be aligned reliably throughout all the 238 studied sequences. We have used only the well-aligned portion to construct a phylogram (Figure 3) that essentially corroborates the previous reports but brings some additional new insights. Our analysis confirms the existence of three major Exo70 lineages Exo70.1, Exo70.2, and Exo70.3 that contain both angiosperm and “lower plant” sequences, as well as the nine clades (A–I) with members of both monocot and dicot origin (Synek et al., 2006). The clade I, restricted only to some angiosperms (it is, e.g., missing in both *Arabidopsis* species), clusters within a branch that includes the compact angiosperm G clade and a group of moss sequences, but none from *Selaginella*, suggesting loss in the lycophyte lineage. We will refer to this wider branch, corresponding to the previously proposed Exo70.3 lineage, as the G/I clade.

Remarkable is the major expansion of a monocot- or grass-specific branch of the F family, the FX clade. Apparently, a single family of Exo70 subunits underwent major expansion in both monocots and dicots. Reverse transcription might have contributed to gene amplification in case of the abundant dicot H clade with a large proportion of single-exon genes (Synek et al., 2006; Chong et al., 2010), but not in case of the monocot FX with a large proportion of multi-exon genes (Chong et al., 2010). Somewhat surprisingly, a very distant paralog OrysaFX8_Exo84, which clusters within the F branch but outside the genuine FX clade,

Table 2 | Numbers of Exo70 paralogs encoded by the studied genomes (in total and in the individual clades).

	All	A	B	C	D	E	F (FX)	G/I	H	BNG ¹
<i>A. thaliana</i>	23	3	2	2	3	2	1(0)	2	8	–
<i>A. lyrata</i>	23	3	2	2	3	2	1(0)	2	8	–
<i>P. trichocarpa</i>	29	4	2	3	2	6	2(0)	5	5	–
<i>S. lycopersicon</i>	22	3	1	3	2	2	1(0)	4	6	–
<i>V. vinifera</i>	15	2	1	1	1	2	1(0)	4	3	–
<i>O. sativa</i>	47	4	3	3	2	1	26 (19)	3	5	–
<i>S. bicolor</i>	31	3	3	2	2	1	16 (12)	3	1	–
<i>B. distachyon</i>	27	5	2	2	2	1	11(6)	3	1	–
<i>S. moellendorffi</i>	8	4	–	–	–	–	–	–	–	4
<i>P. patens</i>	13	3	–	–	–	–	–	4	–	6

¹ Basal non-angiosperm group.



is supported by a full-length cDNA (GenBank AK109785.1) and there are even two closely related ESTs from *Lolium perenne* (GenBank GR511301.1) and *L. temulentum* (GenBank DT673816.1), indicating that this outlier is functional and possibly specific for some grasses.

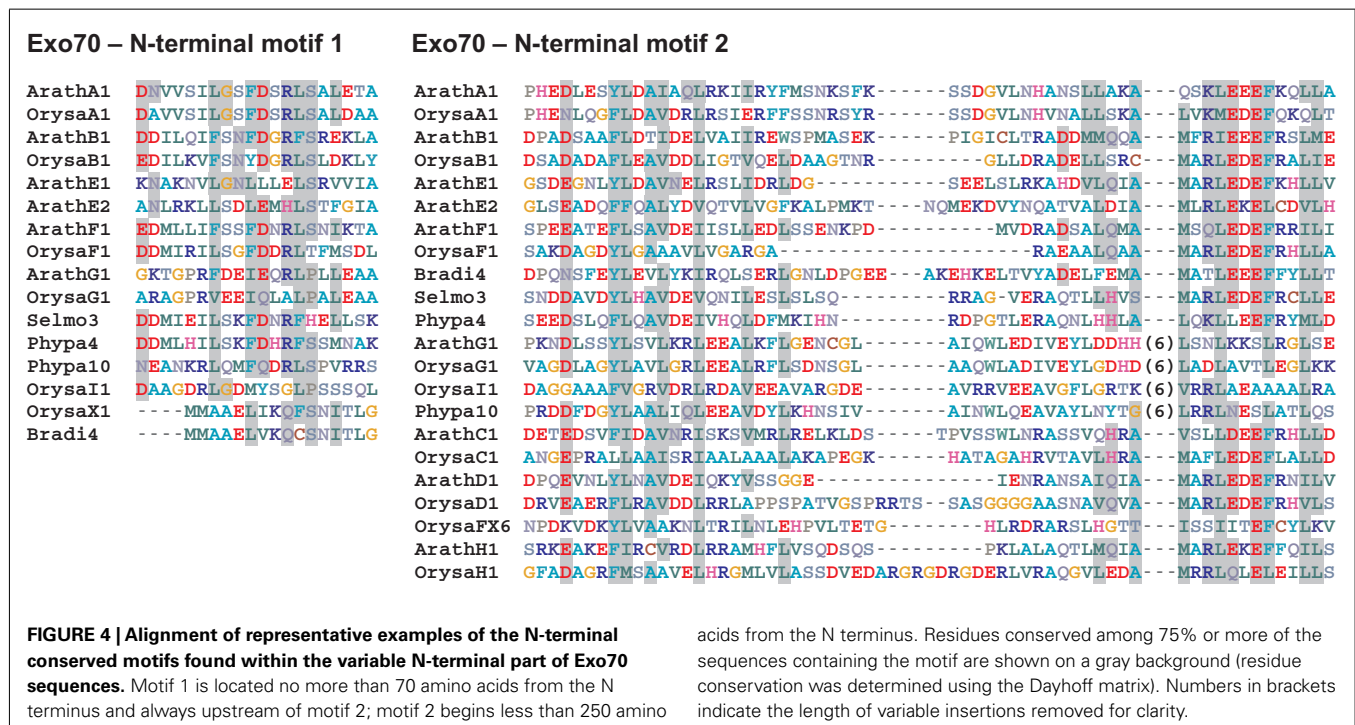
Duplicated genes tend to be rapidly eliminated by natural selection if they bring no advantage in terms of fitness. Question thus arises why there are so many Exo70 varieties maintained across large evolutionary distances. One possibility would be sub- or neofunctionalization of the conserved Exo70 domain itself. We have thus examined representative *A. thaliana* Exo70 sequences for traces of positive (diversifying) selection by estimation of the residue-specific ratio of non-synonymous to synonymous mutation rates (K_a/K_s). For this analysis, we chose two sequence collections – eight representatives of the main clades (AtExo70A1, AtExo70B1, AtExo70C1, AtExo70D1, AtExo70E1, AtExo70F1, AtExo70G1, and AtExo70H1) to identify markers of selection generating or enhancing between-clade differences, and eight representatives of the H clade (AtExo70H1, AtExo70H2, AtExo70H3, AtExo70H4, AtExo70H5, AtExo70H6, AtExo70H7, and AtExo70H8) to find traces of selection favoring within-clade differences. However, in both cases there was only evidence of purifying selection throughout the length of the sequence, but no positive selection, and we thus conclude that differences within the conserved part of the Exo70 subunits are not likely to play a decisive part in determining the function of the individual paralogs.

Functional diversification, however, may be due to the variable N-terminal sequences. We thus examined these regions in more detail and uncovered two sequence motifs conserved in many, but not all, Exo70 paralogs (Figure 4). Distribution of these motifs (see Supplementary Material) suggests that they are ancestral, and that they have been lost or eroded in some of the sequences. Motif

2 is present in most, if not all members of all stable Exo70 clades with exception of FX, and also in the members of the “basal non-angiosperm group,” i.e., in the sequences of lower plant origin with unclear mutual relationships that belong to the Exo70.2 supergroup. The more N-terminally located motif 1 was found in most members of the basal non-angiosperm group and of the A, E, and G/I clades. It is present also in members of the F branch except FX, and also except the distant F outlier *OrysaFX8_Exo70*. This suggests that Motif 1 is also ancestral but was lost in some angiosperm clades within the Exo70.2 supergroup. The presence of the conserved motifs indicates that the variable Exo70 N-termini have largely evolved through a process of mutations and selection rather than domain-shuffling, although this does not have to be the rule in all cases (especially the origin of the diverse and mutually largely unrelated N-termini of the FX proteins remains unclear).

DISCUSSION

The present study provides the first attempt to reconstruct evolution of the land plant, especially angiosperm, exocyst complex in the broader context of higher plant evolution. Our previous works (Eliáš et al., 2003; Synek et al., 2006) have focused only on the most abundant subunit, Exo70, and the only previous phylogenetic study addressing all the eight canonical exocyst subunits in plants (Chong et al., 2010) was based on only four species – *Arabidopsis*, rice and poplar as the representatives of angiosperms, and the moss *P. patens*, allowing only a limited possibility of generalization. We have included a broader and more representative collection of genomes, including a moss (*P. patens*), a lycophyte, i.e., a non-seed vascular plant (*S. moellendorffii*), and eight angiosperms (unfortunately, there is, to date, no sufficiently well-covered gymnosperm genome for an exhaustive search). The angiosperms are represented by five dicotyledonous and three grass species covering a



rather broad range of diversity (see the simplified scheme of plant evolution in **Figure 5**). Among dicots, the closest are the two *Arabidopsis* species (*A. thaliana* and *A. lyrata*) that have separated approximately five millions of years ago (Koch et al., 2000). Poplar (*P. trichocarpa*) is included as somewhat more distant representative of the rosid clade, grapevine (*V. vinifera*) as a basal rosid, and several members of the genus *Solanum* (where, unfortunately, no genome is annotated well enough to provide data for all subunits) are representing the asterids. The coverage of the monocot clade is narrower, as all the available genomes belong to grasses. Thus, although we propose some possible monocot-specific features of the exocyst family in this paper on the basis of data from three grass species (*O. sativa*, *S. bicolor*, and *B. distachyon*), we do not know at present if such features are present also in non-grass monocots.

In total, we have analyzed nearly 400 exocyst subunit sequences. Our list, however, may not be complete especially in case of the *Solanum* sp. sequences, where genomic annotation is still under development, and some loci may have been missed. It is not surprising that our inventory yielded novel genes especially in the Exo70 family in addition to those reported previously for *P. trichocarpa* and *O. sativa* (Chong et al., 2010). On the other hand, in the absence of gene expression data and experimental observations, distinction between functional genes and pseudogenes may be somewhat blurry, especially in case of the extensive Exo70 family, containing numerous single-exon members that apparently underwent reverse transcription at some point in the course of their evolution (Synek et al., 2006). Thus, the determined numbers of subunits might still somewhat change in the future, even in well-characterized models (see the possible undocumented duplication of the *Arabidopsis* Sec10 locus). Also allelic diversity in heterozygous diploids may have resulted in identification of extra-neous loci in particular in the case of *S. moellendorffii*, where most

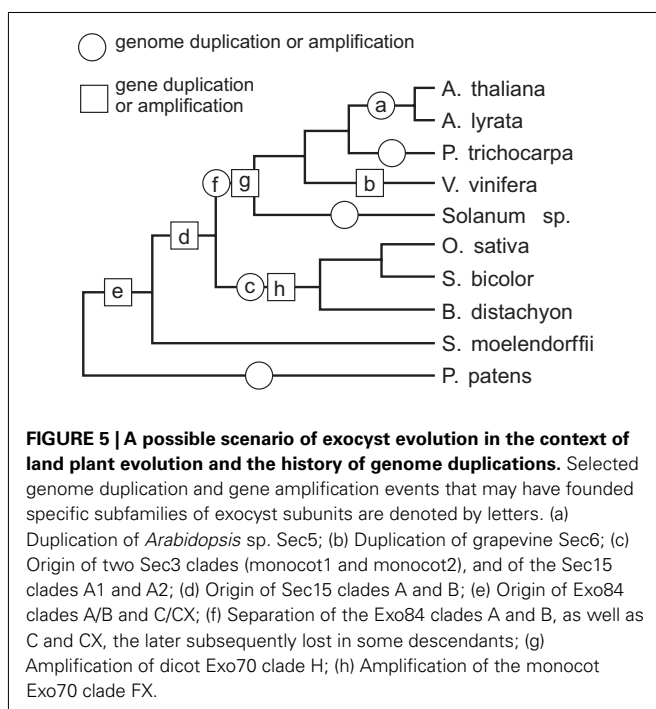
genes appear to have two closely related paralogs, albeit this species is believed to be one of the few plants without a recent history of whole-genome duplications (Jiao et al., 2011).

We have found that a subgroup of exocyst subunits, corresponding to the previously proposed core of the complex (Munson and Novick, 2006; Croteau et al., 2009), underwent little or no amplification in the vascular plants, though even these subunits have amplified to a some extent in non-seed plants. These low copy subunits, in particular Sec6, Sec8, and Sec10, but to a somewhat lesser extent also Sec5 and Sec3, exhibit evolutionary trees that are not only topologically similar but also obviously correlated in terms of branch length, which is consistent with co-evolution driven by the requirement of maintaining mutual compatibility of closely interacting complex subunits (Juan et al., 2008; Lovell and Robertson, 2010). The remaining subunits Sec15, Exo84, and in particular Exo70, exhibit greater diversity consistent with their function on the periphery of the complex, providing an interface to a variety of interactors that may be specific to particular lineages or even to particular paralogs.

Whole-genome duplications have played an important part in the evolution of land plants (Van de Peer et al., 2009; Jiao et al., 2011). They also provided an obvious source of “raw materials” for evolution of divergent paralog families of the exocyst subunits. We were able to pinpoint several of the proposed ancient genome duplication or polyploidization events in a widely accepted scenario of land plant evolution (Soltis et al., 2008; Van de Peer et al., 2009; Jiao et al., 2011; Woodhouse et al., 2011) as possible sources of distinct exocyst subunit clades especially in the Sec15 and Exo84 families (**Figure 5**). However, it has to be stressed that not every gene duplication coincident with a genome duplication must be a result of that duplication. For instance, the tandem duplication of *A. thaliana* Sec3 appears to be a local event, while equally distant *A. thaliana* Sec5 paralogs are obviously a product of a whole-genome duplication (see data from Woodhouse et al., 2011).

The greatest part of the putative exocyst diversity is due to the extremely amplified Exo70 subunit that apparently existed in at least three paralogs already in the common ancestor of land plants including Rhyniophytes, and diversified into seven clades prior to the separation of the monocot and dicot lineages (Synek et al., 2006). Reverse transcription may have contributed to early amplification of some clades, which contain mostly single-exon genes, among them also the dicot clade H that has expanded into an extensive family of paralogs. No such expansion, however, took place in the monocots, which have only a few H-type Exo70s. Instead, a branch of the multi-exon F family has amplified and diversified substantially, producing the monocot-specific FX clade.

While there is considerable sequence divergence among the Exo70 paralogs, we found no evidence of positive selection operating across their conserved part. An obvious source of functional diversity, however, would be the variable sequences at both ends of the Exo70 subunits. A possible participation of C-terminal motifs in differential binding to membrane phosphoinositides has been already proposed (Žárský et al., 2009). Here we have uncovered two obviously ancestral N-terminal motifs that document that the N-terminal segments, though highly diversified, have evolved from a common ancestor at least in most of the sequences, without contribution of major domain-shuffling events. Nevertheless,



they have possibly built up enough diversity to mediate interactions with a variety of cellular components, ensuring thus the apparently required functional diversification.

Assuming that the alternative paralogs of exocyst subunits are co-expressed, and that they can freely combine into complexes (which is by no means guaranteed), literally hundreds of distinct exocysts may exist within plant cells. Were the subunit combinations unrestricted (i.e., each Sec15 paralog working with each Exo84 and each Exo70), *Arabidopsis* would be capable of producing 552 distinct exocyst variants, and rice stunning 1128 variants. Alternative splicing may provide an additional source of exocyst diversity. Even in metazoans, an array of Exo70 splicing variants was uncovered, dependent on cell type and age of the tissue (Delgado et al., 2011). It is therefore possible that also in animals hidden multiplicity of exocysts may exist depending on the splice-isoforms of Exo70 (and possibly also other subunits). On the other hand, the actual numbers of plant exocyst varieties are undoubtedly much lower than the numbers of possible subunit combinations, since not all paralogs are co-expressed, and some may be expressed only under special circumstances or not at all. Nevertheless, we cannot avoid asking what is the biological relevance (or selective advantage) of such a profusion of exocyst varieties.

One possible reason may be the need to maintain and manage a variety of qualitatively distinct membranes – not only of intracellular compartments, but also within the cell cortex whose lateral mobility is restricted by the cell wall. Distinct exocyst variants in the same cell, defined especially by different “landmarking” Exo70 subunits, may participate in delimiting specific plasmalemma domains (“activated cortical domains”) engaging in distinctly regulated membrane turnover. Together with the underlying cytoplasm (in particular the connected recycling endosomes defined by distinct Rab11 paralogs), the activated cortical domains form larger functional units (recycling domains) that may play a central part in the control of different cortical or endomembrane domains of the many-sided plant cells (see detailed discussion in Žárský et al., 2009; Žárský and Potocký, 2010). Another possibility is functional separation of the diverse complexes in time and/or in

tissue or organ space through controlled gene expression of subunit variants optimized for a particular set of circumstances (e.g., specific cell differentiation stages, tissues, or environmental conditions). Participation of Exo84B in the establishment of mycorrhiza (Genre et al., 2012) and, in particular, of distinct Exo70 variants in pathogen response (Pečenková et al., 2011) shows that this indeed appears to be the case. Remarkably, one of the *Arabidopsis* Exo70 paralogs involved in pathogen response is member of the H clade, expanded specifically in the dicots, and it is thus tempting to speculate about a possible analogous role of the even more diversified monocot FX clade.

In summary, both our data and recent experimental observations show that plants do not have “an exocyst complex,” but an enormous variety of diverse exocyst complexes, and that this feature is at least as old as the land plants. Unraveling its functional significance will continue to provide interesting challenges for the plant cell biology of the near future.

ACKNOWLEDGMENTS

This work has been supported by the European Community 7th Framework Programme (FP7/2007–2013) Grant No. 238640 PLANTORIGINS, the Grant Agency of the Czech Republic (P305/11/1629), the Ministry of Education of the Czech Republic (MSM 0021620858), and the Charles University in Prague (SVV 265203/2012).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at <https://www.frontiersin.org/articles/10.3389/fpls.2012.00159/full#supplementary-material>

Cvrckova_S1.xls | List of the 392 exocyst subunit sequences analyzed, including database accession numbers, phylogenetic classification and domain composition (Microsoft Excel file).

Cvrckova_S2.zip | Protein sequences and alignments used for phylogenetic analyses (compressed Zip file containing protein sequences in text format – *.txt and alignment sequences in FASTA format – *.fst).

REFERENCES

- Anisimova, M., and Gascuel, O. (2006). Approximate likelihood ratio test for branches: a fast, accurate and powerful alternative. *Syst. Biol.* 55, 539–552.
- Awashi, S., Palmer, R., Castro, M., Mobarak, C. D., and Ruby, S. W. (2001). New roles for the Snp1 and Exo84 proteins in yeast pre-mRNA splicing. *J. Biol. Chem.* 276, 31004–31015.
- Benson, D. A., Karsch-Mizrachi, I., Clark, K., Lipman, D. J., Ostell, J., and Sayers, E. W. (2012). GenBank. *Nucleic Acids Res.* 40, D48–D53.
- Bombarely, A., Menda, N., Buels, R. M., Strickler, S., Fischer-York, T., Pujar, A., Leto, J., Gosselin, J., and Mueller, L. A. (2011). The sol genomics network (solgenomics.net): growing tomatoes using perl. *Nucleic Acids Res.* 39, D1149–D1155.
- Chong, Y. T., Gidda, S. K., Sanford, C., Parkinson, J., Mullen, R. T., and Goring, D. R. (2010). Characterization of the *Arabidopsis thaliana* exocyst complex gene families by phylogenetic, expression profiling, and subcellular localization studies. *New Phytol.* 185, 401–419.
- Cole, R. A., Synek, L., Žárský, V., and Fowler, J. E. (2005). SEC8, a subunit of the putative *Arabidopsis* exocyst complex, facilitates pollen germination and competitive pollen tube growth. *Plant Physiol.* 138, 2005–2018.
- Croteau, N. J., Furgason, M. L. M., Devos, D., and Munson, M. (2009). Conservation of helical bundle structure between the exocyst subunits. *PLoS ONE* 4, e4443. doi:10.1371/journal.pone.0004443
- Dellago, H., Löscher, M., Ajuh, P., Ryder, U., Kaisermayer, C., Grillari-Voglauer, R., Fortschegger, K., Gross, S., Gstraunthaler, A., Borth, N., Eisenhaber, F., Lamond, A. I., and Grillari, J. (2011). Exo70, a subunit of the exocyst complex, interacts with SNEV(hPrp19/hPso4) and is involved in pre-mRNA splicing. *Biochem. J.* 438, 81–91.
- Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J. F., Guindon, S., Lefort, V., Lescot, M., Claverie, J. M., and Gascuel, O. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 36, W465–W469.
- Eliáš, M., Drdová, E., Žiak, D., Bavlnka, B., Hála, M., Cvrčková, F., Soukupová, H., and Žárský, V. (2003). The exocyst complex in plants. *Cell Biol. Int.* 27, 199–201.
- Fendrych, M., Synek, L., Pečenková, T., Toupalová, H., Cole, N., Drdová, E., Nebesářová, J., Šedinová, M., Hála, M., Fowler, J. E., and Žárský, V. (2010). The *Arabidopsis* exocyst complex is involved in cytokinesis and cell plate maturation. *Plant Cell* 22, 3053–3065.
- Genre, A., Ivanov, S., Fendrych, M., Faccio, A., Žárský, V., Bisseling, T., and Bonfante, P. (2012). Multiple exocytotic markers accumulate at the sites of perifungal membrane biogenesis in arbuscular mycorrhizas. *Plant Cell Physiol.* 53, 244–255.
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten,

- U., Putnam, N., and Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178–D1186.
- Grunt, M., Žárský, V., and Cvrčková, F. (2008). Roots of angiosperm formis: the evolutionary history of plant FH2 domain-containing proteins. *BMC Evol. Biol.* 8, 115. doi:10.1186/1471-2148-8-115
- Guindon, S., and Gascuel, O. (2003). A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704.
- Guo, W., Grant, A., and Novick, P. (1999). Exo84p is an exocyst protein essential for secretion. *J. Biol. Chem.* 274, 23558–23564.
- Guo, W., Roth, D., Gatti, E., and Novick, P. (1997). Identification and characterization of homologues of the exocyst component Sec10p. *FEBS Lett.* 404, 135–139.
- Hála, M., Cole, R. A., Synek, L., Drdová, E., Pečenková, T., Nordheim, A., Lamkemeyer, T., Madlung, J., Hochholdinger, F., Fowler, J. E., and Žárský, V. (2008). An exocyst complex functions in plant cell growth in *Arabidopsis* and tobacco. *Plant Cell* 20, 1330–1345.
- Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* 41, 95–98.
- He, B., and Guo, W. (2009). The exocyst complex in polarized exocytosis. *Curr. Opin. Cell Biol.* 21, 537–542.
- He, B., Xi, F., Zhang, X., Zhang, J., and Guo, W. (2007). Exo70 interacts with phospholipids and mediates the targeting of the exocyst to the plasma membrane. *EMBO J.* 26, 4053–4065.
- Heider, M. R., and Munson, M. (2012). Exorcising the exocyst complex. *Traffic* 13, 898–907.
- Hertzog, M., and Chavrier, P. (2011). Cell polarity during motile processes: keeping on track with the exocyst complex. *Biochem. J.* 433, 403–409.
- Jiao, Y., Wickett, N. J., Ayyampalayam, S., Chandrabali, A. S., Landherr, L., Ralph, P. E., Tomsho, L. P., Hu, Y., Liang, H., Soltis, P. S., Soltis, D. E., Clifton, S. W., Schlarbaum, S. E., Schuster, S. C., Ma, H., Leebens-Mack, J., and de Pamphilis, C. W. (2011). Ancestral polyploidy in seed plants and angiosperms. *Nature* 473, 97–100.
- Juan, D., Pazos, F., and Valencia, A. (2008). Co-evolution and co-adaptation in protein networks. *FEBS Lett.* 582, 1225–1230.
- Kee, Y., Yoo, J. S., Hazuka, C. D., Peterson, K. E., Hsu, S. C., and Scheller, R. H. (1997). Subunit structure of the mammalian exocyst complex. *Proc. Natl. Acad. Sci. U.S.A.* 94, 14438–14443.
- Kitashiba, H., Liu, P., Nishio, T., Nasrallah, J. B., and Nasrallah, M. E. (2011). Functional test of *Brassica* self-incompatibility modifiers in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18173–18178.
- Koch, M. A., Haubold, B., and Mitchell-Olds, T. (2000). Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabidopsis*, and related genera (Brassicaceae). *Mol. Biol. Evol.* 17, 1483–1498.
- Koumandou, V. L., Dacks, J. B., Coulson, R. M., and Field, M. C. (2007). Control systems for membrane fusion in the ancestral eukaryote; evolution of tethering complexes and SM proteins. *BMC Evol. Biol.* 7, 29. doi:10.1186/1471-2148-7-29
- Kulich, I., Cole, R. A., Drdová, E., Cvrčková, F., Soukup, A., Fowler, J. E., and Žárský, V. (2010). *Arabidopsis* exocyst subunits SEC8 and EXO70A1 and exocyst interactor ROH1 are involved in the localized deposition of seed coat pectin. *New Phytol.* 188, 615–625.
- Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D. L., Garcia-Hernandez, M., Karthikeyan, A. S., Lee, C. H., Nelson, W. D., Ploetz, L., Singh, S., Wensel, A., and Huala, E. (2012). The *Arabidopsis* information resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* 40, D1202–D1210.
- Lang, D., Eisinger, J., Reski, R., and Rensing, S. (2005). Representation and high-quality annotation of the *Physcomitrella patens* transcriptome demonstrates a high proportion of proteins involved in metabolism among mosses. *Plant Biol.* 7, 238–250.
- Lassmann, T., and Sonnhammer, E. L. (2006). Kalign, Kalignvu and Mumsa: web servers for multiple sequence alignment. *Nucleic Acids Res.* 34, W596–W599.
- Lawrence, C. E., Altschul, S. F., Boguski, M. S., Liu, J. S., Neuwald, A. F., and Wootton, J. C. (1993). Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. *Science* 262, 208–214.
- Lovell, S. C., and Robertson, D. L. (2010). An integrated view of molecular co-evolution in protein-protein interactions. *Mol. Biol. Evol.* 27, 2567–2575.
- McGinnis, S., and Madden, T. L. (2004). BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res.* 32, W20–W25.
- Munson, M., and Novick, P. (2006). The exocyst defrocked, a framework of rods revealed. *Nat. Struct. Mol. Biol.* 13, 577–581.
- Pečenková, T., Hála, M., Kulich, I., Kocourková, D., Drdová, E., Fendrych, M., Toupalová, H., and Žárský, V. (2011). The role for the exocyst complex subunits Exo70B2 and Exo70H1 in the plant-pathogen interaction. *J. Exp. Bot.* 62, 2107–2116.
- Potato Genome Sequencing Consortium. (2011). Genome sequence and analysis of the tuber crop potato. *Nature* 475, 189–195.
- Samuel, M. A., Chong, Y. T., Haasen, K. E., Aldea-Brydges, M. G., Stone, S. L., and Goring, D. R. (2009). Cellular pathways regulating responses to compatible and self-incompatible pollen in *Brassica* and *Arabidopsis* stigmas intersect at Exo70A1, a putative component of the exocyst complex. *Plant Cell* 21, 2655–2671.
- Schuler, G. D., Altschul, S. F., and Lipman, D. J. (1991). A workbook for multiple alignment construction analysis. *Proteins* 9, 180–190.
- Seguí-Simarro, J. M., Austin, J. R., White, E. A., and Staehelin, L. A. (2004). Electron tomographic analysis of somatic cell plate formation in meristematic cells of *Arabidopsis* preserved by high-pressure freezing. *Plant Cell* 16, 836–856.
- Sklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P., Jensen, L. J., and von Mering, C. (2011). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.* 39, D561–D568.
- Soltis, D. E., Bell, C. D., Kim, S., and Soltis, P. S. (2008). Origin and early evolution of angiosperms. *Ann. N. Y. Acad. Sci.* 1133, 3–25.
- Songer, J. A., and Munson, M. (2009). Sec6p anchors the assembled exocyst complex at sites of secretion. *Mol. Biol. Cell* 20, 973–982.
- Stern, A., Doron-Faigenboim, A., Erez, E., Martz, E., Bacharach, E., and Pupko, T. (2007). Selecton 2007: advanced models for detecting positive and purifying selection using a Bayesian inference approach. *Nucleic Acids Res.* 35, W506–W511.
- Synek, L., Schlager, N., Eliáš, M., Quentin, M., Hauser, M. T., and Žárský, V. (2006). AtEXO70A1, a member of a family of putative exocyst subunits specifically expanded in land plants, is important for polar growth and plant development. *Plant J.* 48, 54–72.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739.
- TerBush, D. R., Maurice, T., Roth, D., and Novick, P. (1996). The exocyst is a multiprotein complex required for exocytosis in *Saccharomyces cerevisiae*. *EMBO J.* 15, 6483–6494.
- The Uniprot Consortium. (2012). Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* 40, D71–D75.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 24, 4876–4882.
- Van de Peer, Y., Maere, S., and Meyer, A. (2009). The evolutionary significance of ancient genome duplications. *Nat. Rev. Genet.* 10, 725–732.
- Wang, J., Ding, Y., Wang, J., Hillmer, S., Miao, Y., Lo, S. W., Wang, X., Robinson, D. G., and Jiang, L. (2010). EXPO, an exocyst-positive organelle distinct from multivesicular endosomes and autophagosomes, mediates cytosol to cell wall exocytosis in *Arabidopsis* and tobacco cells. *Plant Cell* 22, 4009–4030.
- Wen, T. J., Hochholdinger, F., Sauer, M., Bruce, W., and Schnable, P. S. (2005). The roothairless1 gene of maize encodes a homolog of sec3, which is involved in polar exocytosis. *Plant Physiol.* 138, 1637–1643.
- Whyte, J. R., and Munro, S. (2002). Vesicle tethering complexes in membrane traffic. *J. Cell Sci.* 115, 2627–2657.
- Woodhouse, M. R., Tang, H., and Freeling, M. (2011). Different gene families in *Arabidopsis thaliana* transposed in different epochs and at different frequencies throughout the rosids. *Plant Cell* 23, 4241–4253.
- Žárský, V., Cvrčková, F., Potocký, M., and Hála, M. (2009). Exocytosis and cell polarity in plants – exocyst and recycling domains. *New Phytol.* 183, 255–272.
- Žárský, V., and Potocký, M. (2010). Recycling domains in plant cell

- morphogenesis: small GTPase effectors, plasma membrane signalling and the exocyst. *Biochem. Soc. Trans.* 38, 723–728.
- Zhang, X., Zajac, A., Zhang, J., Wang, P., Li, M., Murray, J., TerBush, D. R., and Guo, W. (2005). The critical role of Exo84p in the organization and polarized localization of the exocyst complex. *J. Biol. Chem.* 280, 20356–20364.
- Zhang, Y., Liu, C. M., Emons, A. M. C., and Ketelaar, T. (2010). The plant exocyst. *J. Integr. Plant Biol.* 52, 138–146.
- 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.
- Received: 13 June 2012; paper pending published: 25 June 2012; accepted: 29 June 2012; published online: 18 July 2012.
- Citation: Cvrčková F, Grunt M, Bezvoda R, Hála M, Kulich I, Rawat A and Žárský V (2012) Evolution of the land plant exocyst complexes. *Front. Plant Sci.* 3:159. doi: 10.3389/fpls.2012.00159
- This article was submitted to *Frontiers in Plant Traffic and Transport*, a specialty of *Frontiers in Plant Science*.
- Copyright © 2012 Cvrčková, Grunt, Bezvoda, Hála, Kulich, Rawat and Žárský. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.