

RESEARCH ARTICLE

Open Access



Genome-wide identification of MST, SUT and SWEET family sugar transporters in root parasitic angiosperms and analysis of their expression during host parasitism

Vikram A. Misra¹, Eric K. Wafula², Yu Wang^{1,3}, Claude W. dePamphilis² and Michael P. Timko^{1*}

Abstract

Background: Root parasitic weeds are a major constraint to crop production worldwide causing significant yearly losses in yield and economic value. These parasites cause their destruction by attaching to their hosts with a unique organ, the haustorium, that allows them to obtain the nutrients (sugars, amino acids, etc.) needed to complete their lifecycle. Parasitic weeds differ in their nutritional requirements and degree of host dependency and the differential expression of sugar transporters is likely to be a critical component in the parasite's post-attachment survival.

Results: We identified gene families encoding monosaccharide transporters (MSTs), sucrose transporters (SUTs), and SWEETs (Sugars Will Eventually be Exported Transporters) in three root-parasitic weeds differing in host dependency: *Triphysaria versicolor* (facultative hemiparasite), *Phelipanche aegyptiaca* (holoparasite), and *Striga hermonthica* (obligate hemiparasite). The phylogenetic relationship and differential expression profiles of these genes throughout parasite development were examined to uncover differences existing among parasites with different levels of host dependence. Differences in estimated gene numbers are found among the three parasites, and orthologs within the different sugar transporter gene families are found to be either conserved among the parasites in their expression profiles throughout development, or to display parasite-specific differences in developmentally-timed expression. For example, MST genes in the pGLT clade express most highly before host connection in *Striga* and *Triphysaria* but not *Phelipanche*, whereas genes in the MST ERD6-like clade are highly expressed in the post-connection growth stages of *Phelipanche* but highest in the germination and reproduction stages in *Striga*. Whether such differences reflect changes resulting from differential host dependence levels is not known.

Conclusions: While it is tempting to speculate that differences in estimated gene numbers and expression profiles among members of MST, SUT and SWEET gene families in *Phelipanche*, *Striga* and *Triphysaria* reflect the parasites' levels of host dependence, additional evidence that altered transporter gene expression is causative versus consequential is needed. Our findings identify potential targets for directed manipulation that will allow for a better understanding of the nutrient transport process and perhaps a means for controlling the devastating effects of these parasites on crop productivity.

Keywords: Parasitic plant, Sugar transporter, MST, SUT, SWEET, Expression, *Orobanchaceae*

* Correspondence: mpt9g@virginia.edu

¹Department of Biology, University of Virginia, Gilmer Hall 044, Charlottesville, VA 22904, USA

Full list of author information is available at the end of the article



Introduction

Root-parasitic weeds in the family *Orobanchaceae* are among the major constraints to crop production worldwide [1–3], causing losses in crop yield and economic value estimated to be in excess of several billion USD annually [2, 3]. Among the hosts to these parasites are many of the world's most important food and forage grasses (corn, sorghum, millet and rice) and grain legumes (cowpea, peanut) with yield losses ranging from minimal to 100% in areas throughout their range [4–6]. Currently, measures limiting the effects and spread of these parasites are not effective. Compounding these control difficulties is the fact that the seeds of the various *Orobanchaceae* can remain viable in the soil for many years [7] making eradication difficult.

It is now well documented that progression through the parasite life-cycle is dependent upon a series of host-derived chemical cues that lead to critical developmental changes throughout the host-parasite interaction. For example, parasite seed germination only occurs after a period of after-ripening and preconditioning in the soil, and in response to the presence of host-derived chemical stimulants (e.g., strigolactones) in the rhizosphere [8, 9]. Following germination, the elongating parasite radicle then transforms into a haustorium, the unique organ for host attachment, in response to a second set of host-derived chemical cues referred to as haustorial initiation factors (HIFs) [10, 11]. Once formed, the haustorium then attaches to the host root, develops invasive tissues that transverse the host root cortex, and eventually establish connection with the host vascular system. Coincident with this is the development of specialized cells and tissues that allow the parasite to obtain and utilize host derived nutrients [10]. Most of the damage to the host plant occurs during the early pre-emergent stages of parasite growth as evidenced by a decline in host growth and vigor [1]. With the emergence of the parasite shoot from the soil, it grows to reproductive maturity, flowers and sets seed, thereby completing its life-cycle [9, 10].

Despite their enormous threat to many crops, little research was done until recently on the molecular aspects of host-parasite weed interactions relative to other well-studied disease-causing agents. Most early molecular studies of parasitic plants of the *Orobanchaceae* family focused on plastid genome evolution [12, 13], diversity analysis [14–18], and phylogenetic relationships [19–21] among family members. Subsequently, reports appeared in which investigators attempted to understand the gene expression changes in both hosts and parasites in response to attack by the parasite using both differential gene expression profiling using cDNA library analysis and sequencing of expressed sequence tags (ESTs) from both host and parasite [22–26].

Comparative transcriptomics of three parasitic species in the *Orobanchaceae* (i.e., *Triphysaria versicolor*, *Phelipanche aegyptiaca*, and *Striga hermonthica*) and several related non-parasites was carried out in order to examine the global changes that occurred during the evolution of parasitism in this group and to potentially identify changes required for the parasitic lifestyle [27–29]. Over 3 billion sequence reads from more than 30 tissue-, developmental stage-specific-, and normalized whole plant- libraries were generated and are available for analysis through the Parasitic Plant Genome Project (PPGP; <http://ppgp.huck.psu.edu/>). A first approach at analyzing these datasets conducted by Wickett et al. [30] showed that while the non-photosynthetic parasitic plant *P. aegyptiaca* had no detectable expression of genes involved in the formation of the photosystems and light harvesting proteins, the genes for chlorophyll biosynthesis were retained, transcribed, and subject to purifying selection, suggesting a function for chlorophyll independent of photosynthetic activity. Subsequently, Yang et al. [31] analyzed the transcriptomic data from *P. aegyptiaca*, *S. hermonthica*, and *T. versicolor* for differential expression and gene expression clustering in order to identify candidate parasitism genes with roles in facilitating parasitic interactions with host plants. These investigators showed that expression shifts, with or without gene duplication, was a common mechanism by which most genes involved with parasitism evolved and they suggested that either adaptive evolution or relaxed selective constraints may have been important in the evolution of haustorial genes. Based on expression data, Yang et al. [31] further suggested that the co-option of genes involved with root and floral development played a key role in the evolution of the haustorium. Using transcriptomic data generated from host-parasite interface tissue samples prepared by laser capture dissection, Honaas et al. [32] showed that parasitic plant gene expression patterns change in response to the nature of the host being attacked. Moreover, among the genes differentially expressed at the interface are various sugar transporters, indicating a likely crucial role during the host-parasite interaction.

Sugar transporters are broadly characterized into three major types: monosaccharide transporters (MSTs), sucrose transporters (SUTs), and SWEETs (Sugars Will Eventually be Exported Transporters). MSTs and SUTs function in sugar influx [33], whereas the SWEETs (also known as *Medicago truncatula* nodulin 3 (MtN3)/saliva [34]) play important roles in phloem transport [35] where they primarily function in sucrose efflux and bidirectional sugar transport [35–37].

MSTs are Sugar_tr domain (PF00083) containing members of the major facilitator superfamily (MFS) class of transporters. Structurally, MFS transporters usually

contain 12 transmembrane (TM) domains [37–39]. Seven MST clades have been defined in plants: (i) the Early Response to Dehydration 6-like (ERD6-like) clade [40] (sometimes abbreviated as ESL); (ii) the plastidic glucose translocator (pGLT) clade, which contains the SGB1 (suppressor of G protein beta 1) sub-clade [37]; (iii) the inositol transporters (INT) and (iv) the tonoplasmic monosaccharide transporters (TMT) clades [41], both of whose members localize to the tonoplast and plasma membrane [41–43] and are involved in vacuolar monosaccharide transport [37, 44]; (v) the polyol/monosaccharide transporter (PMT) clade, which is referred to as PLT in Lalonde & Frommer [41]; (vi) the vacuolar glucose transporters (VGT) clade; and (vii) the sugar transport protein (STP) clade [38, 41], whose members play a variety of roles, including interaction with symbiotic and pathogenic fungi [37], pollen development, and root development [45].

SUTs, like MSTs, are Sugar_tr domain (PF00083) members of the MFS superfamily [38], but their organization is less clear. The SUTs were originally split into five clades encompassed by SUT1–SUT5 [41, 46]. SUT1 is a dicot specific clade whose members are expressed in companion cells and sieve elements [37] and is proposed to move from companion cells to sieve elements via the endoplasmic reticulum (ER) [46, 47]. SUT1 proteins play a role in phloem loading [48, 49] and unloading [50] and in cellular interactions with symbiotic and pathogenic fungi (reviewed in [37]). Members of the SUT2 clade are found in both monocots and dicots [46] and are expressed in sink cells and, to a lesser extent, source leaves [51]. Like SUT1 clade members, SUT2 clade members can contribute to phloem loading and the transport of sucrose into various sink cells [52]. The SUT4 clade is also expressed in both monocots and dicots [46] and is localized in expression to sieve elements [53] and in source leaves [54]. They are involved in several processes including responses to dehydration and photosynthesis [54] and circadian rhythms [55], to nodule development [56]. SUT gene expression has been shown to change under different physiological conditions associated with source-to-sink sucrose transport [57]. Peng et al. [58] split the SUTs into two subfamilies: Ancient Group 1 (AG1) and Ancient Group 2 (AG2). Within AG1 are two clades, the “Type I” clade unique to dicots and the “Type III” found in both monocots and dicots. There are also two clades within AG2, the “Type II” found in both monocots and dicots and “Type IIB” which is monocot specific. According to Peng et al. [58], the SUT1 group reported by Kuhn & Grof [46] and the SUT4 group reported by Lalonde & Frommer [41] belong to the Type I and Type III clades, respectively. The Type II clade of Peng et al. [58] contains SUT2 from Lalonde & Frommer [41] and the Type IIB group contains SUT3 and SUT5 from Kuhn & Grof [46].

SWEETs are sugar transporters, which like SUTs and MSTs, undergo changes in gene expression under different physiological conditions, changes that are associated with the regulation of source-to-sink sucrose transport [57]. In contrast to MSTs and SUTs, the SWEETs are members of a superfamily of transporters containing seven TM domains [36, 37, 59]. SWEETs are characterized by an MtN3/saliva domain [60] and contain a pair of repeats that span two TM helices, connected by a loop [61]. There are four groups of SWEETs exclusive to plants, numbered I through IV.

Prior studies have established a crucial role for sugar transport in basic plant growth and development, in root and shoot elongation, growth at the meristems, and interactions with commensal and disease-causing microbes [33, 37, 45, 62]. Members of both the MSTs and SUTs have been implicated in controlling the interaction between host plants and pathogenic fungi [37, 63], and SWEETs have been shown to be exploited by pathogens such as *Xanthomonas oryzae* [64]. Apart from an analysis of sucrose transporters in *Phelipanche ramosa* [65] little is currently known about the structure and organization of sugar transporter gene families in parasitic weeds and the role they may play in mediating host-parasite interactions.

It is well documented that parasitic plants differ in their degree of host dependency ranging from facultative hemiparasitic forms (e.g., *Trypophysaria* spp.) which can photosynthesize and complete their lifecycle without host contact, to fully non-photosynthetic holoparasites, including species of *Phelipanche* (syn. *Orobanchae*) genus, which are totally dependent upon their host. Others, like the obligate hemiparasite *Striga*, are capable of photosynthesis but are host dependent at their early growth stages and remain so later in development for as yet unclear reasons [27–29]. In light of these differences and the importance of nutrient exchange from host to parasite throughout their lifecycle, we hypothesized that differences in the developmental timing and level of expression of sugar transporters may exist among parasites with differing host dependence and that these differences may be important to parasite survival. To address this question, we have characterized the sugar transporter repertoires of three parasitic plants, *Phelipanche aegyptiaca*, *Striga hermonthica*, and *Triphysaria versicolor* that differ in their degree of host dependency and examined the patterns of expression of the various family members in parasite seedlings before and following host attachment and in the early and post-emergent/reproductive stages of parasite growth. In addition, we use analyses of orthogroups to determine whether any variations in expression within a clade can be explained by the possible presence of multiple orthogroups in the same sugar transporter clade. We identify members of

the MST, SUT, and SWEET gene families with both general and developmental stage specific expression and speculate on the potential roles of these genes/gene families in controlling host interaction.

Results

We have identified members of the MST, SUT and SWEET sugar transporter gene families in three parasitic plants, the facultative hemiparasite *Triphysaria versicolor*, the holoparasitic *Phelipanche aegyptiaca*, and the obligate hemiparasite *Striga hermonthica*, and the non-parasite *Mimulus guttatus* (syn. *Erythranthe guttata*) a member of the traditional Scrophulariaceae (now Phrymaceae) [66], using a novel gene discovery pipeline. This discovery pipeline (see Materials and Methods) uses a strict set of rules based upon domain structure to define gene family membership. To identify members of the SUT families and to accommodate the previous phylogenetic organizations of Kuhn & Grof [46] and Lalonde & Frommer [41]. In the present study, we used representative sequences from the rice SUT3 and SUT5

clades and the Arabidopsis SUT1, SUT2, and SUT4 clades to query the parasite datasets. Clades SUT1 – SUT5 are given in the figures presented, and for consistency with prior studies of Peng et al. [58] we treat the SUT1 clade as being synonymous with Type I clade, SUT2 as being synonymous with Type II clade, SUT4 as being synonymous with Type III, and SUT3 and SUT5 clades as being synonymous with Type IIB. As search queries to find parasitic plant SWEET sequences we employed the conserved SWEET domains of 33 rice SWEET gene sequences (21 from the *Japonica* sub-species and 12 from the *Indica* sub-species) and 17 sequences from *A. thaliana*. Table 1 shows the estimated sizes of the MST, SUT and SWEET gene families based on the three parasitic plants transcriptomes studied and the *M. guttatus* genome. These are representative genes in which multiple isoforms are excluded. The obligate hemiparasite *S. hermonthica* has the largest number of expressed MST representative genes, just slightly greater than that found in the facultative parasite *T. versicolor*. This is almost double of that found in the holoparasite

Table 1 Estimated numbers of sugar transporters in three parasitic plant transcriptomes and the *Mimulus* genome. For each group of sugar transporters, the percentage of sequences is given for each taxon on the phylogenetic trees in Figs. 1, 2, 3, 4 and 5 and Additional files 1, 2, 3, 4 and 5. Please note that the percentages calculated in this table are the percentages of only the sequences found in the phylogenetic trees. In the table, “outliers” refer to sequences that were eliminated during the part of the curation phase in which preliminary neighbor-joining trees were made to detect sequences that were on branches longer than 0.3 when related branches had length of under 0.1; these outlier sequences were on branches longer than 0.3

	<i>Phelipanche aegyptiaca</i>		<i>Striga hermonthica</i>		<i>Triphysaria versicolor</i>		<i>Mimulus guttatus</i>	
	#	% of Total on Tree	#	% of Total on Tree	#	% of Total on Tree	#	% of Total on Tree
ERD6	5	14.71%	8	12.12%	6	9.52%	6	13.04%
pGLT_SGB	2	5.88%	2	3.03%	3	4.76%	4	8.70%
INT	3	8.82%	7	10.61%	3	4.76%	6	13.04%
TMT	6	17.65%	2	3.03%	3	4.76%	4	8.70%
PMT	7	20.59%	20	30.30%	23	36.51%	7	15.22%
VGT	2	5.88%	2	3.03%	3	4.76%	2	4.35%
STP	9	26.47%	25	37.88%	22	34.92%	17	36.96%
Total MST on Tree	34		66		63		46	
MST Isoforms and Outliers	80		162		174		23	
SUT1	2	50.00%	5	55.56%	3	60.00%	2	50.00%
SUT2	1	25.00%	1	11.11%	1	20.00%	1	25.00%
SUT4	1	25.00%	3	33.33%	1	20.00%	1	25.00%
Total SUT on Tree	4		9		5		4	
SUT Isoforms and Outliers	7		11		10		0	
SWEET Clade I	8	57.14%	8	32.00%	5	38.46%	10	34.48%
SWEET Clade II	0	0.00%	5	20.00%	3	23.08%	3	10.34%
SWEET Clade III	6	42.86%	10	40.00%	5	38.46%	14	48.28%
SWEET Clade IV	0	0.00%	2	8.00%	0	0.00%	2	6.90%
Total SWEET on Tree	14		25		13		29	
Other SWEET Isoforms and Outliers	3		6		1		5	

P. aegyptiaca. *S. hermonthica* also has the largest number of expressed SUTs and SWEETs. The *M. guttatus* genome contains 46 representative MST genes, 4 SUT genes, and 29 representative SWEET genes (Table 1). By comparison, there are 53 MSTs [45] and 9 SUTs [58] in the *Arabidopsis thaliana* genome, while the rice genome has 65 MSTs [67] and 5 SUTs [58]. Yuan & Wang [68] reported that there are 21 SWEETs in the rice genome and 17 in the *A. thaliana* genome. Two points are worth noting here. First, due to the fragmentary nature of transcriptome de novo assemblies, our copy numbers for the parasite species refer to estimated numbers since accurate copy numbers can only be determined when complete genome assemblies and mature annotations become available. Second, the sequence set at the nucleotide and amino acid level was examined for possible isoforms using the following criteria: sequences with attributes such as high similarity to other sequences, similar domain structure, and short (< 0.01) distances on a neighbor joining phylogram. These criteria were used to remove possible isoforms. However, the expression values from isoform-rich and single-isoform filtered datasets were studied.

Phylogenetic analysis of sugar transporter gene families

We next examined the phylogenetic relationships of the various members of the MST, SUT and SWEET gene families within and among the parasitic species in order to determine whether similarities and differences in particular clades account for the observed variation in estimated gene numbers. Significant differences in MST family structure were observed among the three parasitic species (Figs. 1, 2, 3, 4 and 5, Table 1, Additional files 1, 2, 3, 4 and 5). For example, the PMT and STP clades were the largest of the clades comprising the MST family in all three parasites, with membership in the ERD6-like clade also being a significant proportion of the total family size (Table 1). *Phelipanche* had the greatest number of TMT genes, whereas *Striga* had the greatest number of ERD6-like, INT, and STP genes. *Triphysaria* had the greatest number of PMT and VGT genes, while *Mimulus* had the greatest number of pGLT/SGB genes.

In general, the three parasites have comparable numbers of members in the three clades comprising the SUT family (i.e., SUT1, SUT2 and SUT4) of these species. The only exceptions are the expansion of SUT1 and SUT4 genes in *S. hermonthica* relative to its counterparts (Fig. 4, Table 1). *S. hermonthica* has the largest number of SWEET genes among the three parasites (Fig. 5, Table 1) and almost twice in the total number of SWEET genes found in *T. versicolor* and *P. aegyptiaca*. Expansion in *S. hermonthica* appears to principally be in SWEET Clades II, and III. Overall, in comparison to their non-parasitic relative *M. guttatus*, the three

parasitic plants are over-represented in clade II and under-represented in clade III SWEETS. Of the parasites, the number and distribution of SWEETS within clades in *S. hermonthica* is most similar to that of *M. guttatus*. Only *Striga* and *Mimulus* had SWEET genes in clade IV (Table 1, Fig. 5).

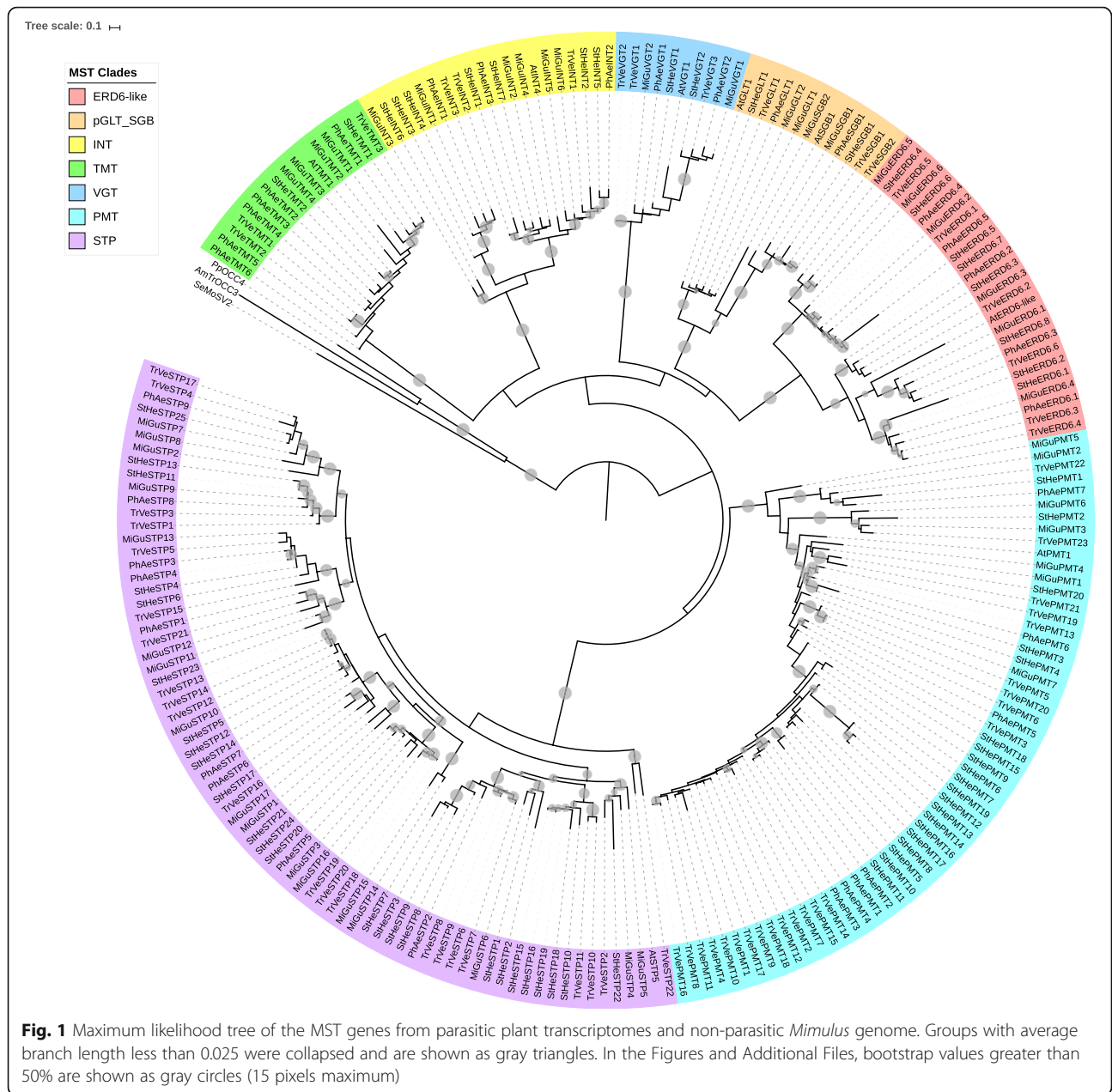
Expression of sugar transporters in three parasitic plants

To uncover differences in sugar transport gene expression that may exist among parasites with different levels of host dependency, we first examined the levels of expression (based upon FPKM data) of the various MST, SUT and SWEET family members at the different developmental stages in each of the parasites. We paid particular attention to any notable difference in expression levels before and following host contact. The results of this analysis are discussed below for each transporter family. Descriptions of each stage of the parasite life cycle can be found in the Materials and Methods. In this analysis, in order to avoid underestimating the expression levels of the various sugar transporters, the gene-level expression was initially considered rather than that of each isoform (as shown in Figs. 2, 3, 4 and 5 and Additional files 1, 2, 3, 4 and 5). However, since the possibility exists that different isoforms are expressed during different stages of the life cycle we also examined both the gene-level and isoform-level expression separately. These values are given in Additional files 7, 8 and 9.

MST gene family

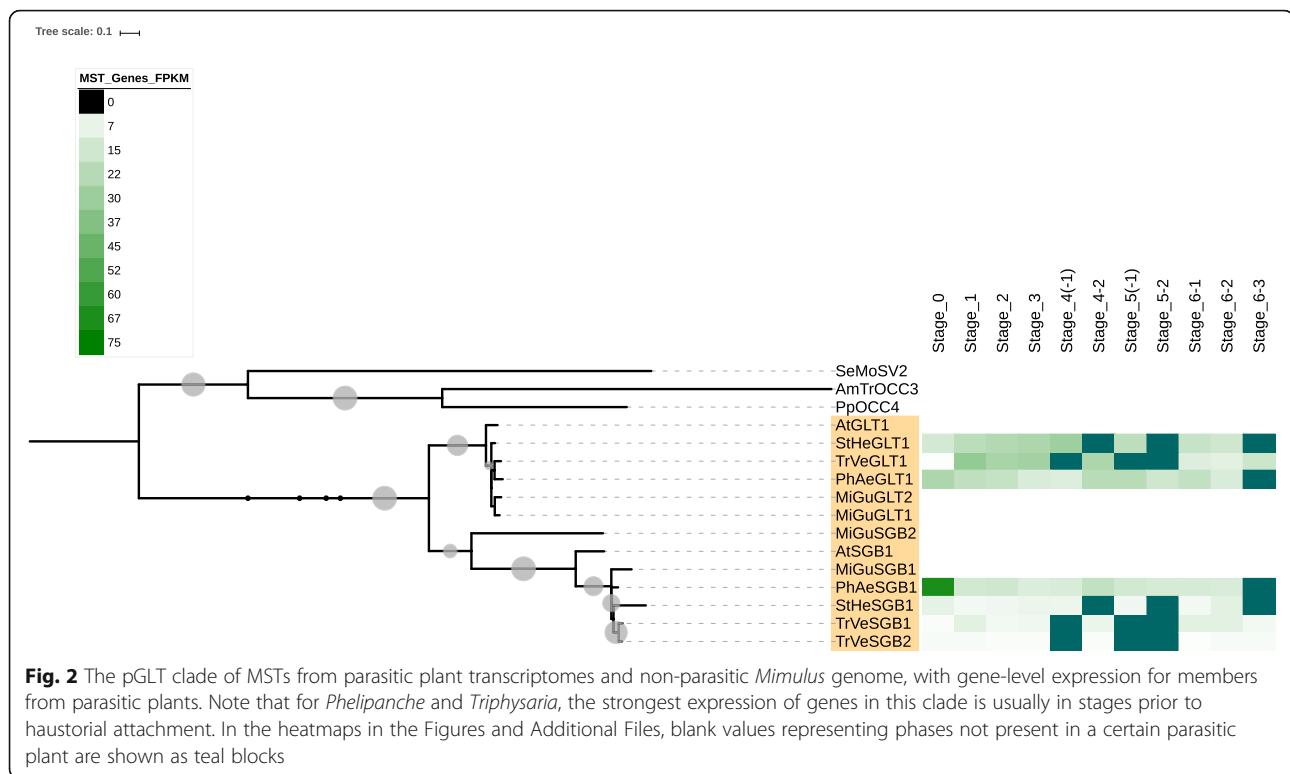
In general, expression of MSTs in the various parasitic plants increase from germination to pre-attachment root growth (stages 1 and 2), decreases during haustorial connection, and then increases again during pre- and post-emergence growth and development, anthesis and reproduction (Figs. 2, 3, 4 and 5 and Additional files 1, 2, 3, 4 and 5). Among the MSTs, increased expression is most evident in the PMT and STP clades (Additional file 6). However, the sharp increase in MST expression in the post-emergence growth and reproductive stages evident in *Striga* and *Triphysaria* is not seen in *Phelipanche* (Additional file 6). In *Triphysaria*, MST expression is also high during germination and pre-attachment root growth (stages 1 and 2).

Among the three parasitic plants, the expression profiles of genes in the ERD6-like clade show the most differences (Additional files 1 and 6). For example, in *Phelipanche*, the ERD6-like genes are expressed highly during post-attachment root growth (stage 5–1) and stem and leaf growth stages (stage 6–1) whereas in *Striga*, most ERD6-like genes are strongly expressed during the germination and reproductive phases (Additional files 1 and 6).



Members of the pGLT clade of MSTs similarly show differential expression during the parasite life-cycle (Fig. 2, Additional file 6). In *Phelipanche*, pGLT members tended to be most highly expressed in germination. In *Striga*, orthologous pGLT members are expressed the time of haustorial connection (stage 4–1) and in floral bud development in reproduction (stage 6–2), and in *Triphysaria*, sequences in the pGLT clade expressed most strongly during pre-haustorial root growth. In *Triphysaria*, expression of genes in the INT clade was higher in leaf and stem growth (6–1). In *Striga*, INT genes appear to be more highly expressed during pre-haustorial root growth (stage 1) and reproduction (stage 6–2) (Additional files 2 and 6).

Of the plants studied, *Mimulus* had the highest proportion of INT annotated genes, and *Triphysaria*, the lowest. Expression of representative genes in the TMT clade was highest in the pre- and post-emergent growth stages in *Phelipanche*, and post-emergent growth stages in *Striga* and *Triphysaria* (Additional files 3 and 6). Most PMTs were expressed most strongly during post-emergent growth in all three parasites (Additional file 6). However, in *Triphysaria*, PMTs also expressed strongly in pre-haustorial root growth (Fig. 3 and Additional file 6). In *Striga* and *Triphysaria*, the VGT clade members were found to be more highly expressed during post-emergent tissue growth and reproduction (Additional file 6).



However, in *Phelipanche*, expression of VGTs was greatest during haustorial connection and post-emergent root growth (Additional files 4 and 6). Members of the STP clade in all three parasitic plants tended to express most strongly during post-emergent growth and reproduction. In *Triphysaria*, STP expression is also strong in pre-haustorial growth (Additional files 5 and 6), and in *Striga*, STP expression is strongest in pre-emergent growth.

SUT gene family

The expression of SUT family members appears to be highest in post-emergent growth stages of the life-cycles of the three parasites (Fig. 4, Additional file 6). This is especially evident in *Striga* and *Triphysaria*. However, significant levels of SUT expression can be observed during germination and pre-haustorial growth in *Striga* and *Triphysaria* (Fig. 4, Additional file 6).

Members in the SUT1 clade exhibit high expression during post-emergent growth stages, with highest levels occurring as plants progressed towards reproduction (stage 6–2). In *Phelipanche*, for example, SUT1 genes expressed most strongly in stages 6–1 and 6–2 (above ground stem and leaf growth and reproduction, respectively) whereas in *Striga* and *Triphysaria*, most SUT1 genes were strongly expressed during reproduction (stage 6–2) (Fig. 4, Additional file 6). SUT2 clade members are expressed during various stages throughout the life-cycle (Fig. 4, Additional file 6), with no strong preference for

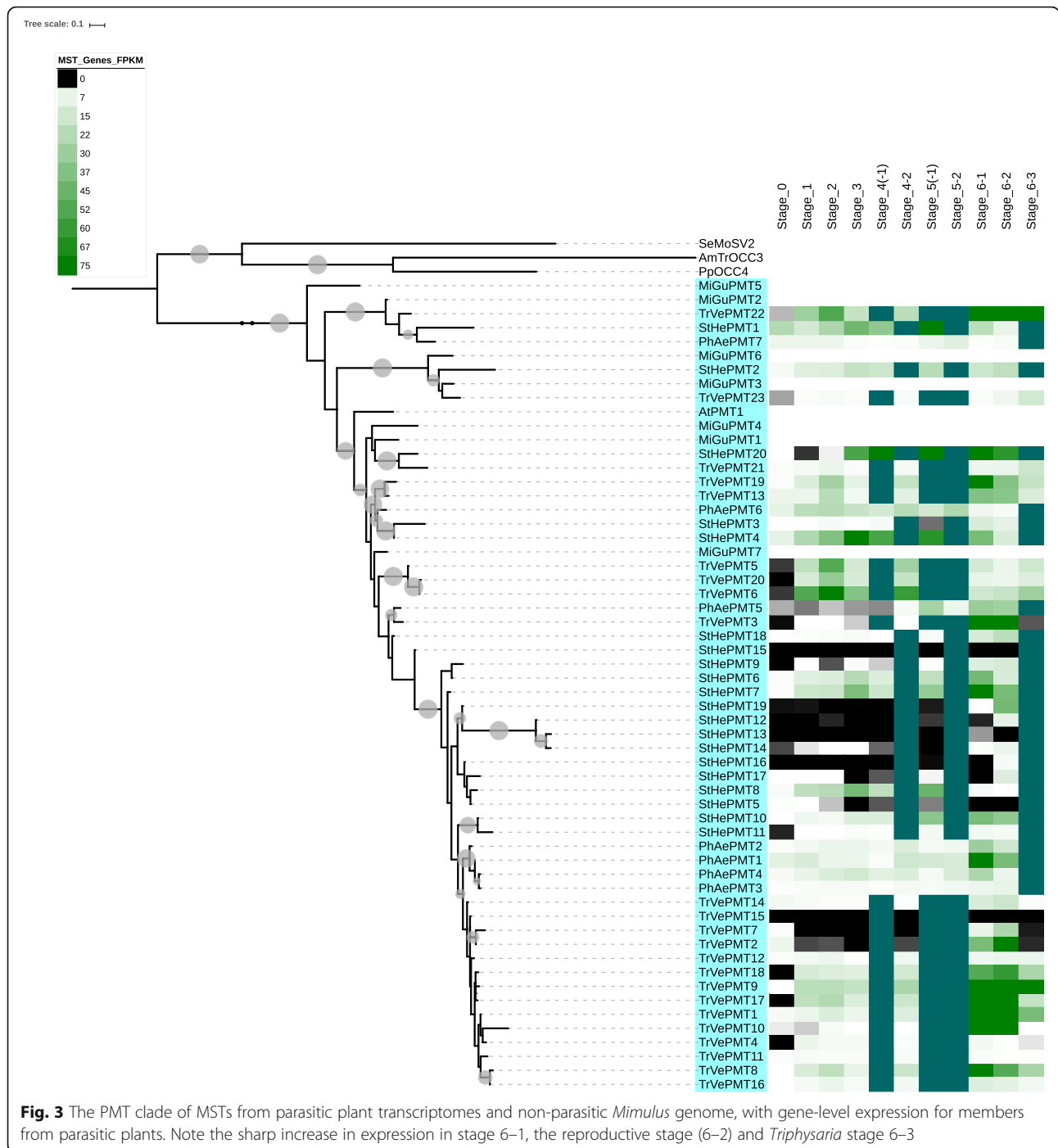
any one developmental time point. Members of the SUT4 clade appear to be expressed to a greater extent in the pre-haustorial root growth stage 3 and haustorial connection phase 4–2 in *Phelipanche*. In *Triphysaria*, SUT4 expression was strongest during germination (stage 0).

SWEET gene family

The expression profiles of the members of the SWEET family are shown in Fig. 5. In general, we observed that members of the SWEET gene families from the three related parasites tended to have similar patterns of expression throughout the parasite life-cycle with *Triphysaria* and *Striga* SWEET family members being most similar (Fig. 5, Additional file 6). Overall, members of the SWEET family are most highly expressed during reproduction and floral bud development (stage 6.2) of parasite development (Fig. 5, Additional file 6). In *Striga* and *Triphysaria*, there are also several SWEET genes that express more strongly during germination and pre-attachment root growth (Fig. 5, Additional file 6).

Conserved motifs and Orthogroup analyses

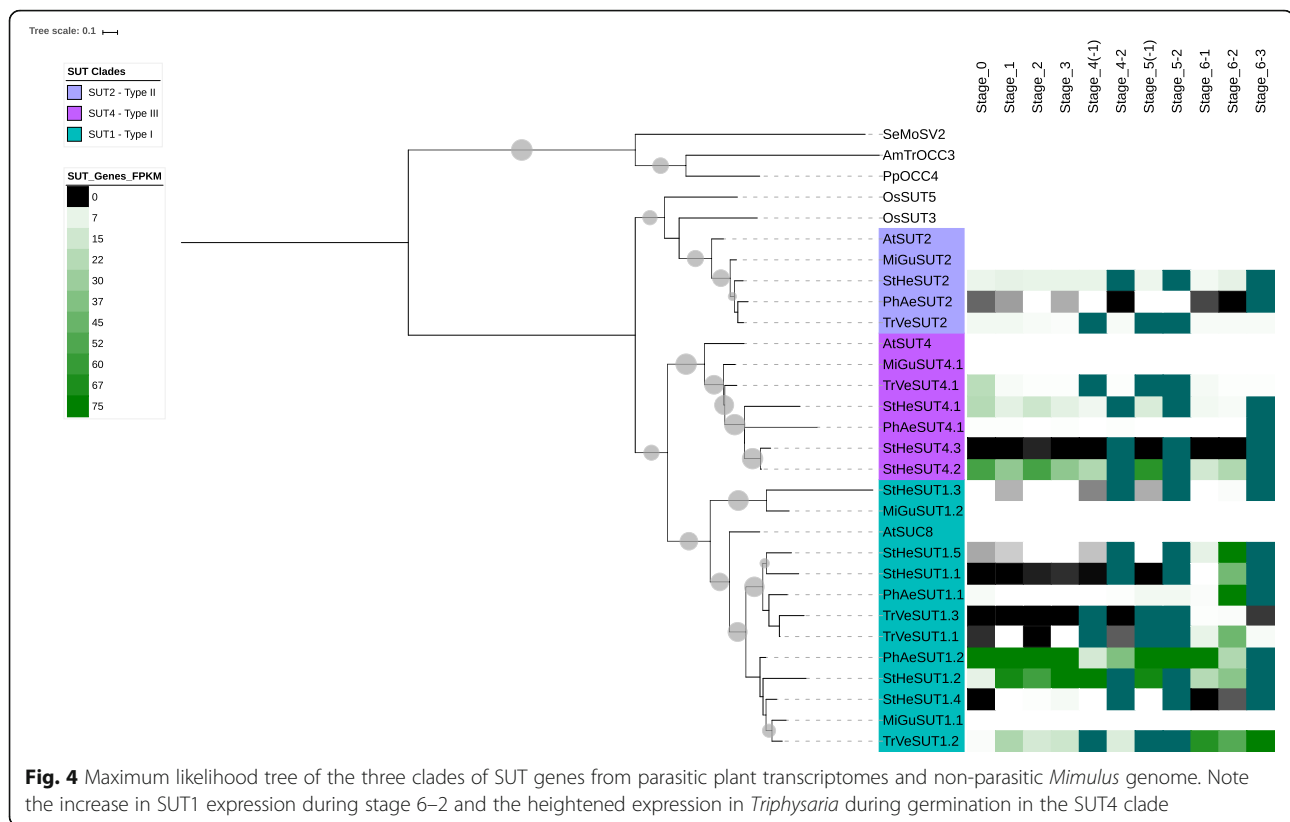
We next examined the nature of the conserved motifs defining members of these various sugar transporters using MEME analysis (Additional files 7, 8, 9, 10, 11, 12 and 13) with the orthogroup most commonly associated with each motif and the sequence for each motif found in the MST, SUT and SWEET gene families. This was



done to determine if variations within each clade of MST, SUT or SWEET are associated with differences in expression within these clades, and whether there was a one-to-one correspondence between the gene family definition based on our homology searches and orthogroup definitions. The statistics on the motifs whose presence correlated with certain clades, plants and stages are shown in Additional files 10, 11 and 12; these data are summarized in Additional file 13. Overall, within each

clade of sugar transporters, there existed a significant number of motifs that were either exclusive to or strongly represented in certain clades of the MST family.

An orthogroup analysis in OrthoFinder [69] (Additional files 14, 15, 16 and 17) on the Galaxy public server at usegalaxy.org [70] revealed that many orthogroups were unique to one clade, but that certain clades had more than one orthogroup. Here it is important to note that while the orthogroups in the MST and SUT families



all contained a Sugar_tr PFAM domain and the SWEET orthogroups all had an MtN3_slv domain, that members of different clades (and therefore orthogroups as well) had different conserved MEME domain architectures (Additional files 10, 11 and 12).

In the MST family, the ERD6-like, pGLT, INT and PMT clades each had two orthogroups that were unique to them. The STP clade had five orthogroups. In the SUT family, orthogroup 1431 contained all members of the SUT1 and SUT4 clades, whereas the SUT2 clade contained only orthogroup 3807. In the SWEET family, clade I was almost entirely represented by three orthogroups (with one *Mimulus* SWEET in a fourth orthogroup, 5750), all of which were unique to that clade, all clade II members were part of orthogroup 769, clade III members were part of orthogroup 429, and clade IV members belonged to orthogroup 2089.

In several instances different orthogroups within the same clade were found to have their highest levels of expression at different developmental stages. An example where this can be seen is the five orthogroups of the STP clade (Additional file 17). This may be explained by the fact that different STP clade members can have varying affinities for different monosaccharides [45].

Discussion

Parasitic weeds rely upon their host to different extents for their nutritional needs and therefore the organization and expression of nutrient transporters would likely need to vary both developmentally and temporally during their life cycle. Our findings indicate that members of the *Orobanchaceae*, while sharing sugar transporter genes in common, may adjust the expression of these genes in response to their life-styles. *P. aegyptiaca* is a holoparasite and *S. hermonthica* is an obligate hemiparasite [27]. Since holoparasites have lost anatomical and molecular features associated with photosynthesis [27, 30, 70, 71], it is expected that sugar transporters associated with photosynthesis would be lost in *Phelipanche*. The difference in repertoire size between *Phelipanche* and *Striga* (i.e., *Phelipanche* has fewer MSTs, SUTs and SWEETs than *Striga*) could reflect the differential state of host dependency in the parasites since a number of the sugar transporters associated with photosynthesis may be lost in *Phelipanche*. Differences in the composition of the MST, SUT and SWEET families among the parasites might also reflect differences in nutrient preference between holoparasitic, obligate and facultative hemiparasitic forms [72].

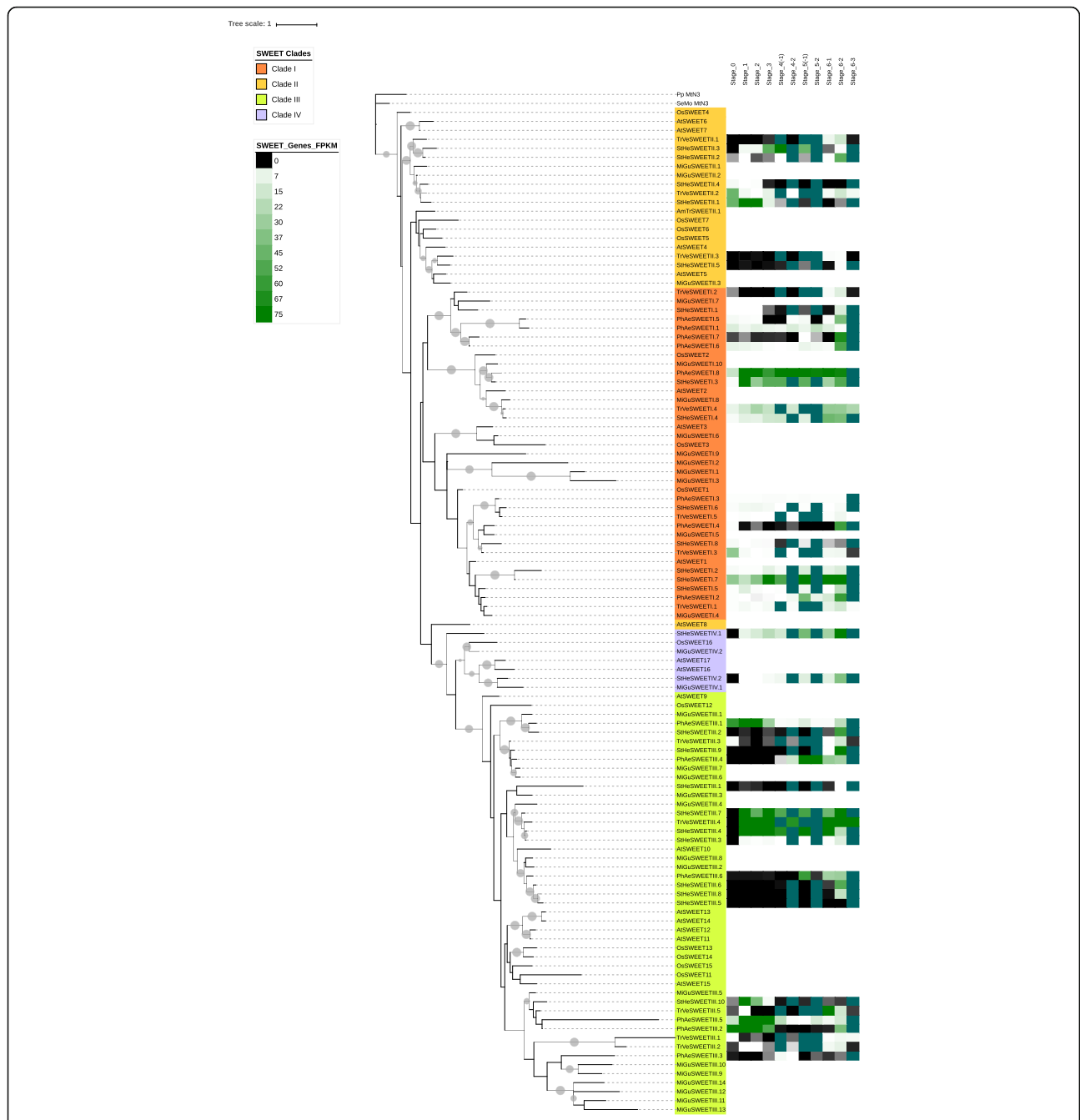


Fig. 5 Maximum likelihood tree of the four clades of SWEET genes from parasitic plant transcriptomes and non-parasitic *Mimulus* genome. Note the apparent similarity of the SWEET expression profiles to those of the MST family, as well as the lack of *Phelipanche* sequences in clades II and IV

Phylogenetic Analysis of Sugar Transporter Gene Families
 The three parasitic plants transcriptomes examined contained representatives of all seven clades of MSTs present in plants (i.e., ERD6-like, pGLT/SGB, INT, TMT, PMT, VGT, STP), three SUT family clades associated with dicots (i.e., SUT1, SUT2 and SUT4), and the four clades of SWEETs previously identified. We queried the datasets for all three parasites to determine whether any

orthologs of the two monocot specific clades of SUTs defined by Kuhn & Grof [46] were not present since there is a possibility of their presence resulting from horizontal gene transfer (HGT) as described in *S. hermonthica* and other plants in the *Orobanchaceae* family [73–76]. However, we did not find any SUTs that were nested within with members of SUT3 or SUT5 (see Fig. 4 and Additional file 8).

The parasitic plant transcriptomes studied had a lower proportion of members in the ERD6-like clade than observed in the *A. thaliana* genome, which has 19 members accounting for 35.8% of all *Arabidopsis* MSTs [77]. Since ERD6-like proteins transport glucose from vacuole to cytoplasm in non-parasitic plants [78], and taking into account that parasitic plants have been suggested to be strong sugar sinks [65], it is possible that ERD6-mediated glucose transport may be limited in parasitic plants. Since there is evidence that some STPs function preferentially in root and pollen development [45], it is not unreasonable to speculate that differences in the STP clade sizes among the three parasitic plants could reflect differential needs in each plant.

It has been previously reported that SWEET13 homologs in maize has stronger expression in leaf vasculature and thus may be important in photosynthetic carbon movement and that SWEET13a, b, c triple-knockout mutants have impaired photosynthesis [76]. In this study, we observed that expression of genes in *Phelipanche* SWEET Clade III is stronger in stages of growth that do not require photosynthesis (i.e., Stage 5–2 and before pre-emergence shoot growth) (see Additional file 6). In contrast, *Striga* and *Triphysaria*, show stronger expression of SWEET clade III in the post-emergent stages 6–1 (leaves and stems) and reproductive Stage 6–2, and floral structure development in stage 6–3. Because *Striga* and *Triphysaria* are capable of photosynthesis, unlike the holoparasitic *Phelipanche*, the difference in expression levels between *Phelipanche* and the hemiparasitic *Striga* and *Triphysaria* appear to be an example of differences in sugar transporter activity between parasitic plants correlated with their different degree of host dependence.

We also found that *Phelipanche* had the highest proportion of clade I SWEETs (57.14% of SWEETs in *Phelipanche*) (Table 1), that *Phelipanche* and *Triphysaria* lacked any clade IV SWEETs, and that parasitic plants smaller proportion of clade III SWEETs than in *Mimulus*, while *Striga* and *Triphysaria* have a greater proportion of clade II SWEETs than *Mimulus*. Together, these results suggest that overall; the SWEET family of sugar transporters in the *Orobanchaceae* family of parasitic plants may be organized differently from other plants. However, it must be noted that the transcriptomes studied contain genes that have sufficiently high transcription levels. There may be genes in these sugar transporter families that are not transcribed at high enough levels to be included in the transcriptomes studied, but may be found should full genome assemblies and annotations become available.

Expression of sugar transporters in three parasitic plants

According to Figs. 2 and 3 and Additional files 1, 2, 3, 4 and 5 and 7, all parasitic plants in this study in general

increase in MST expression during pre-haustorial root growth phases, decrease during Haustorial attachment and penetration, and increase significantly during pre- and post-emergent developmental phases. This result is consistent with most damage being done to the host plant during post-connection growth stages [1]. Given the findings that carbohydrates accumulate in pre-emergent shoots in *P. aegyptiaca* [79], it is possible that sugar transporters could be involved in mobilizing the storage of these carbohydrates during the flowering phase.

Some ERD6-like proteins have been shown to be involved in the transport of sugars out of vacuoles [78]. During the post-emergent growth and reproduction stages of the parasites there is likely a greater need for host nutrients and therefore it might be expected that as the parasitic develops, some ERD6-like proteins would be differentially required based on individual parasite needs. It has been previously reported that INTs are localized in the tonoplast and play important roles in root development [43]. In this current study, the non-parasite *Mimulus* was found to have a high proportion of INTs compared to its parasitic relatives, and the facultative *Triphysaria* had an especially low proportion of INTs, suggesting that parasitic plants may have different requirements for INTs than their non-parasitic counterparts. In *Striga*, TMT expression was more likely to express strongly during vegetative growth stage 6–1, and in *Phelipanche*, pre-emergent root growth showed the greatest likelihood for strong TMT expression (Additional files 6 and 7), suggesting higher mobilization of glucose and fructose to the vacuolar lumen in these plants, similar to what is seen in *Arabidopsis thaliana* [80]. The profile of polyols is host dependent [81]. For example, *P. aegyptiaca* accumulates mannitol when on a tomato host [82]. PMTs, along with VGTs, are known for long-distance transport and phloem loading [83]. The parasites investigated in the present study were all grown on different hosts and, therefore, it is not surprising that the PMT genes have different expression profiles in the different parasitic plants. If one assumes that plants metabolize different polyols, and accumulate different levels of various polyols in various tissues [84] then one might expect varied levels of transporters throughout the lifecycle. The findings in this study (see Additional file 6) suggest that members of the PMT clade were most likely to exhibit high level of expression in stages 6–1 through 6–3. In *Arabidopsis*, VGT proteins are mainly expressed in above ground tissues [85]. The observed high expression of VGT genes in post-emergent growth stages, especially vegetative growth (stage 6–1) in *Striga* and *Triphysaria* is consistent with these findings. The expression of groups of STPs in growth and reproductive stages, especially in

Striga and *Triphysaria*, is consistent with STPs in *Arabidopsis* being expressed during pollen development and root development [45].

Weise et al. [53] reported that SUT4 is associated with movement of photosynthates whereas Frost et al. [54] reported that SUT4 is involved in transport process associated with water stress. In the present study, SUT2 was the weakest expressed member of the SUT family in the parasitic plants (Fig. 4, Additional file 8) consistent with the results of Peron et al. [65], who reported that SUT2 transcripts accumulate at low levels throughout the *P. ramosa* life cycle. In contrast, expression of *PrSUT1* and *PrSUT3*, (members of the SUT1 clade in *P. ramosa*), were expressed highest after emergence of the flowering shoot [65]. We similarly observed SUT1 expression to be strongest during reproduction in the three parasites.

In non-parasitic plants, SUT1 is associated with phloem loading [51, 52, 86] and unloading [50] and SUT2 has been proposed to be a sucrose sensor [52]. Thus, it is possible that in *Triphysaria*, a facultative hemiparasite, the mechanisms for nutrient movement in the free living state are similar to those for nutrient movement when attached to a host. It is also possible that SUT1s are involved in retrieval of sucrose during transport in sieve elements and unloading in sink organs such as roots and flower structures. The expression data in Fig. 4 and Additional file 8 suggest sucrose unloading into flower buds. To fully understand whether this is the case, though, studies on regulation of and by SUT1 and SUT2 in parasitic plants would have to be done.

In this study, SWEET expression, especially in *Striga* and *Triphysaria*, were strongest during post-connection growth and especially reproductive stages. These results are thus consistent with the finding that SWEETs have been shown to efflux sucrose out of parenchyma cells into the phloem [35]. However, there were members of SWEET clades II and III in *Triphysaria* that express more strongly in pre-host attachment stages. Given the growth of roots and haustoria in these stages [28] and given the role of clade III SWEETs AtSWEET11 and AtSWEET12 in root growth [35, 64], it is possible that SWEETs have roles in pre-haustorial root growth in parasitic plants such as *Phelipanche* or *Triphysaria*.

Fig. 2, 3, 4 and 5 and Additional files 6, 7, 8 and 9 suggest the presence of few instances of high expression during haustorial attachment (stage 3–4). While it is possible that expression shifting among the members of the sugar transporter families could be significant, it is important to note that were it available, careful annotation of a full genomic assembly could identify additional undetected orthologous genes or nearly identical paralogous genes that are not distinguishable in de novo transcriptome assemblies

The movement of RNAs (both transcripts and small RNAs) has been reported to occur between hosts and parasitic plants [87, 88]. Therefore, it may be possible to use transgenic expression of RNAi to reduce or eliminate specific SWEETs necessary for post-attachment development of host-parasite connections required for parasite nutrient acquisition. The use of RNAi to better defend host plants from parasitic plants has been proposed [89], and suggestions that this approach may be successful have appeared see for example refs. [90, 91]. Whether or not such approaches would eventually also succumb over time allowing the parasite to bypass the need for specific targeted transporters remains an open question [92].

Conserved motifs and Orthogroup analyses

In this study, MEME analyses [93] identified conserved motifs in common and unique to proteins encoded by members of genes in different clades, such as the ERD6-like and PMT clades (see Additional files 10, 11, 12 and 13). These conserved motifs may be essential to the different functions of the encoded proteins (e.g., their cellular targeting or substrate utilization) corresponding to the differential developmental and temporal expression of genes within these clades. In addition, there exists a significant number of motifs in the SUT and SWEET families that were exclusive to *Arabidopsis*, rice and the outgroup genes in this study. These results suggest that parasitic plant sugar transporters are likely to use different motifs from those in non-parasitic plants suggesting that protein structural difference may have evolved to fulfill specialized functions in the parasite.

In addition, consistent with expectations, clade definition and orthogroup definition were different. While some clades only had one orthogroup, other clades consisted of multiple orthogroups. In the MST family, while there were no instances of one orthogroup being in multiple clades, there were instances of one clade having more than one orthogroup (Additional file 16).

In the SUT family of parasitic plants, all SUT1 and SUT4 sequences are associated with the same orthogroup, a result consistent with phylogeny proposed by Peng et al. [58] in which SUT1 and SUT4 are classified as members of the same Ancient Group (AG1). The orthogroup composition of the sugar transporter gene families was used to investigate whether orthogroups could explain variations in expression within a clade. As expected, different orthogroups within a larger clade showed differential expression preference for different stages of the life cycle (e.g., in the STP clade). However, some of this variation may be explained by members of different orthogroups being more frequent in different species (Additional file 17). Of the seven orthogroups identified in the SWEET family, only clade II

(orthogroup, group 769) and clade IV (orthogroup 2089) had single orthogroup membership (Additional file 16). Clade I and II however shared some orthogroup overlap indicating a possible relatedness between these clades.

The above results underscore the need for further studies to determine how the members of the different clades of MST, SUT and SWEET transporters are used during parasite interactions with the host. While examining conserved protein motifs represented in the various MST, SUT and SWEET transporters may serve as one level of analysis, it might also be enlightening to look at the conserved regulatory motifs in the promoters of these various genes responsible for their differential expression. However, at the present time only transcriptomic data are available for the parasitic plants studied here, and therefore, such an analysis cannot be done until whole-genome data becomes available.

Conclusions

In this study, we identified members of the MST, SUT and SWEET families of sugar transporters in three parasitic weeds, *P. aegyptiaca*, *S. hermonthica*, and *T. versicolor*, and investigated the phylogenetic relationships among the various gene family members and their differential expression during parasite growth and development. We showed that members of different clades of MSTs, SUTs and SWEETs were differentially expressed throughout the parasitic plant life cycle and expression profiles were dependent on the parasitic plant. Our observations indicate that that different parasitic plants are regulating sugar transporter expression differently and speculate that some of these differences may be due to the differences in host species and degree of host dependency exhibited by the parasitic plant species. This is consistent with prior suggestions that parasitic plants may change the expression of genes within their regulatory networks to most effectively parasitize a host. How this is accomplished remains unknown, but could involve the differential expression or activation of members of transcription factor and transcriptionally-active protein gene families that regulate sugar transporter gene expressions as previously proposed [32, 94]. These studies do however identify potential targets for directed manipulation that will allow for a better understanding of the nutrient transport process and perhaps a means for controlling the devastating effects of these parasites on crop productivity.

Materials and methods

Tissues, transcriptomic library preparation and sequence analysis

Detailed descriptions of parasite growth and the collection of biological materials from the parasite developmental stages (which extend from imbibed seed, germination and haustorial development, attachment to above ground tissues) can be found in Westwood et al.

[28]. Transcriptome sequencing was previously described in Yang et al. [31] and encompassed multiple stages of parasite development from the species *T. versicolor*, *S. hermonthica*, and *P. aegyptiaca* within *Orobanchaceae* as described therein.

Data sources

Transcriptome assemblies and expression data for *P. aegyptiaca*, *S. hermonthica*, and *T. versicolor* were obtained from the PPGP II datasets from the PPGP Website (<http://ppgp.huck.psu.edu/download.php>). The de novo assemblies were performed with Trinity [95] and post processed into non-redundant sets including predicted coding sequences and their corresponding using the PlantTribes pipeline (<https://github.com/dePamphilis/PlantTribes>) [31]. Cleaned reads were mapped to post-processed assembled transcripts and expression abundance of the parasite developmental stages [28] estimated using the RSEM pipeline [96] with the Bowtie2 [97] read aligner option. A detailed description of the PPGP II datasets is available on the PPGP website.

Known *Arabidopsis thaliana* and rice SWEET sequences were retrieved from the NCBI Protein, Nucleotide, and EST Databases [98–100]. The *Arabidopsis* MSTs and SUTs used for each clade (from [38, 41]) were also found on those same databases. The *Mimulus guttatus* genome version 2.0 [101, 102] was retrieved from Phytozome version 12.1 [103, 104].

Identification of sugar transporters

The *Arabidopsis* sugar transporter sequences and rice SWEET sequences that were found as described above were used as queries to search for potential sugar transporters in the assembled transcriptomes of the parasitic plants *P. aegyptiaca*, *S. hermonthica*, and *T. versicolor*, as well as the annotated genome of the related non-parasitic plant *Mimulus guttatus*. In searching for sugar transporter sequences, the FASTA, FASTX and TFASTX methods [105, 106] from FASTA version 36.3.8e (September 30, 2016, accessed from https://fasta.bioch.virginia.edu/fasta_www2/fasta_list2.shtml [107]) was used with an E-value cutoff of 1e-3. The potential sequences were retrieved using custom Perl scripts that incorporated modules from BioPerl 1.7.1 [108, 109]. Any coding sequence (CDS) or nucleotide (NT) sequences were translated into amino acid sequences across six reading frames using a custom BioPerl script [110]. The sequences were searched using SANSParallel [111, 112] against the UniprotKB database [113], with 50 hit sequences displayed per query, and the “very slow” setting. For all BLAST results, inspection of best matches for each query sequence was used to determine if a sequence was a plant MST, SUT or SWEET. Only

sequences whose best matches had an E-value of $1e-3$ or smaller were kept for further analyses.

Curation, multiple sequence alignment, and phylogenetic analysis

MEME Suite 4.12.0 [93, 114] command line was used to discover motifs in all sugar transporter sequences studied, with any number of repetitions expected, 100 motifs maximum, and each motif having a width ranging from 25 to 250 amino acids long, and maximum iterations at 250. All species were analyzed in the same MEME run in order to accurately determine which motifs sugar transporter sequences had in common, and which motifs, if any, may be essential to the function of members of certain clades of the MST, SUT or SWEET families of sugar transporters. In addition, the discovered motifs were used to curate all sequences analyzed and eliminate false positives in the sequence sets, thus ensuring more accurate downstream analyses. MEME was used twice in this study, once as a tool for curation, and the second time as a tool to compare sequences within a gene family (described below).

Multiple sequence alignment was done using MAFFT L-ins-i version 7.312 [115, 116].

MSTs and SUTs are members of the MFS superfamily and contain Sugar_tr or MFS_1 PFAM domains, and SWEETs contain the MtN3_slv PFAM domain. Therefore, sequences with the appropriate PFAM domains (Sugar_tr and MFS_1 for both MST and SUT, GPH_sucrose for SUT, and MtN3_slv for SWEET) were used to aid in rooting sugar transporter phylogenetic trees; these sequences were from the *Physcomitrella patens* genome assembly v3.3 [103], the *Selaginella moellendorffii* assembly v1.0 [117], and the genome of the basal angiosperm *Amborella trichopoda* v1.0 [118]. For MSTs and SUTs, a keyword search using the term “MFS” in Phytozome was used. For the SWEETs, rooting sequences were found by using BLASTP against the Phytozome Database [103] with an E-value cutoff of $1e-5$, using *Physcomitrella patens* sequences from the OrthoMCL database (ppat|e_gw1.127.40.1, orthogroup OG5_127038) [119]; the best hits from each of the three proteomes searched were kept as sequences to be used as outgroups for the SWEETs in this study.

Sequence sets for proteins and transcripts of the genes studied were analyzed. First, transcript sequences were aligned and examined for isoforms, the method for which is described below. Once isoforms were removed from the set, the names of the remaining transcript sequences were used to retrieve their protein counterparts. The proteins were then aligned and examined for isoforms. Isoforms of the same gene for *Mimulus*, *Arabidopsis* and rice sequences were removed prior to multiple sequence alignment. Parasitic plant sequences that were potentially isoforms of the same gene were eliminated from the sequence set as described below. To

ensure an accurate phylogenetic tree, trimAl 1.4.1 [120] was used to first trim alignments to remove columns with less than 10% occupancy, and then eliminate any sequences that covered less than 50% of the alignment ($-resoverlap\ 0.50 -seqoverlap\ 50$).

The alignment process at both the transcript and the protein levels consisted of using a BioPerl script we developed to eliminate redundant sequences, running MAFFT, and then running trimAl. This process was repeated until all sequences covered at least 50% of the alignment. A similar procedure has been used in Yang et al. (2015) [31]. In addition, to find potential outliers in the phylogenetic dataset, a preliminary neighbor-joining (NJ) phylogenetic tree was made with ClustalW2 program [121], using uncorrected p-distance and no gap exclusion. If any non-outgroup branch length for a sequence was shown to be unusually long (i.e., branch length > 0.3 when related branches had lengths of under 0.1), the unaligned, non-trimmed version of that sequence was retrieved from the unedited assembly and then manually curated using information from SANSParallel, MEME or the Conserved Domain Database [60, 93, 122]. After that, the newly edited version of the sequence replaced the non-edited version in the set of sequences to align. Isoforms of the same gene were found and eliminated from the sequence set in the preliminary NJ tree mentioned above for sequences with one or more of the following criteria: high sequence similarity (indicated by branch lengths < 0.01), length of sequences (too short and the sequence was suspected to be a truncated version of the gene), and similarity of MEME protein motif architecture. Then the sequences were re-aligned. The process of detecting long branches, curating sequences on long branches, removing potential isoforms, and re-aligning was repeated until the presence of very long branches and isoforms, which could mislead phylogenies, was minimized or eliminated. The resulting curated set of sequences was used in phylogenetic analysis. Phylogenetic trees were made using RAxML 8.2.11 [123, 124], using 1000 bootstrap replicates and using the PROTGAMMAAUTO option, which automatically includes model testing, with five gamma categories, and corrected Akaike information criterion (AICc) [125]. The names of the sequences on the phylogenetic trees in the Figures and Additional files are shorthand names, and are listed in Additional file 18 alongside their corresponding names from the FASTA files.

Analysis of conserved motifs and Orthogroups

The names of the sequences that were phylogenetically analyzed were used to retrieve the unedited versions of the sequences (i.e., the versions of the sequences prior to curation and alignment). These unedited versions of the sequences were run through MEME (using the parameters described above). The resulting motifs were included in Additional files 7, 8 and 9. To avoid the presence of

identical gene copies, which would have made the analyses less accurate, the sequences that were eliminated during curation and alignment were not included in this analysis. The MEME domains were linked to the sequences to which they belonged in order to determine which clade, life cycle stage, and species with which the MEME motifs were most strongly associated.

The set of unedited versions of the sequences used in the phylogenetic analysis were also used in an OrthoFinder [68] analysis on the Galaxy server [70], using both blastp and hmmscan as protein classifiers, and a minimum e-value of $1e^{-5}$.

Analyses of differential gene expression

Expression values for the parasitic weeds upon infecting a host [31] were based on read counts expressed as FPKM (fragments per kilobase of unigene length per million reads), which were based on the results of read mapping for every gene in the library for each species, after parasitic weeds were allowed to infect host plants. For each gene, each FPKM value in the data studied here represents the sum of FPKM values from all isoforms for a given gene. That is, after *Phelipanche* infected *Arabidopsis thaliana* and tobacco; after *Striga* infected sorghum; and after *Triphysaria* infected *Medicago truncatula* [28, 32].

The names of each developmental stage in the parasite are described in detail in Westwood et al. (2012) [28] and illustrated in Yang et al., (2015) [31]. In the present study, stage 0 represents germination, 1 and 2 represent radicle elongation and haustorial differentiation, respectively [28]. Stage 3 represents the phases in which a parasitic plant uses its haustoria to attach to the host plant, and in stage 4, the parasitic plant connects to its vascular system [28]. Stage 5 is the phase of the growth of pre-emergent tissues, roots (5.1) and shoots (5.2). Stage 6.1 represents the post-emergent growth of leaves and stems, and stages 6.2 represented reproduction, the development of floral buds, while stage 6.3 represented further floral maturation in *Triphysaria* [28].

The expression values for each sugar transporter sequence were placed into a text file using a custom Perl script that we developed. The phylogeny, expression data and conserved MEME domains were visualized using Interactive Tree of Life (iTOL) version 4 [126, 127].

Additional files

Additional file 1: The ERD6-like clade of MSTs from parasitic plant transcriptomes and non-parasitic *Mimulus* genome. Please note the *Triphysaria* ERD6-like clade members with strongest expression in stage 2, in contrast with *Phelipanche* and *Striga*. (PDF 30 kb)

Additional file 2: The INT clade of MSTs from parasitic plant transcriptomes and non-parasitic *Mimulus* genome. Note that, as with most other parasitic plant sugar transporter families, there is an increase in expression in pre-haustorial growth stages, a decrease in expression

during haustorial connection, and an increase in expression in pre- and post-emergence growth and reproduction. (PDF 26 kb)

Additional file 3: The TMT clade of MSTs from parasitic plant transcriptomes and non-parasitic *Mimulus* genome. Note the slight increase in TMT expression during pre-connection and haustorial connection phases. (PDF 24 kb)

Additional file 4: The VGT clade of MSTs from parasitic plant transcriptomes and non-parasitic *Mimulus* genome. Note the tendency to express more strongly in stages 6–1 and 6–2. (PDF 23 kb)

Additional file 5: The STP clade of MSTs from parasitic plant transcriptomes and non-parasitic *Mimulus* genome. Note the increase in expression during post-attachment stages in *Triphysaria*. (PDF 44 kb)

Additional file 6: Numbers of sugar transporters and their stages of strongest expression in *Phelipanche aegyptiaca*. The stages of strongest expression are shown for each clade of MSTs, SUTs and SWEETS, as well as for all MSTs, SUTs and SWEETS, in the transcriptomes of *P. aegyptiaca*, *S. hermonthica*, *T. versicolor*. (XLSX 13 kb)

Additional file 7: Maximum likelihood tree of the MST genes, with heat map and domain architectures. Note the general increase in expression with each successive stage of the life cycle. (PDF 158 kb)

Additional file 8: Maximum likelihood tree of the SUT genes, with heat map and domain architectures. Note that in the SUT1 clade, there is a tendency for genes to express more strongly in later life cycle stages, culminating in a sharp increase in expression in stage 6–2. (PDF 43 kb)

Additional file 9: Maximum likelihood tree of the SWEET genes, with heat map and domain architectures. Note the strong tendency to show higher levels of expression during the reproductive stage 6–2, with some exceptions in clades II and III. (PDF 73 kb)

Additional file 10: MEME motifs and their associations with stages of the life cycle. These associations were determined by finding the genes to which these motifs belonged, and determined the stage in which these genes are most strongly expressed. This was done for (A) MSTs, (B) SUTs, and (C) SWEETS. The motifs most strongly associated with given stages were determined by finding the stage with the highest raw number of occurrences of the motif in that stage, as well as the representation of that motif as a percentage of motifs present in genes that showed strongest expression during that stage. (XLSX 47 kb)

Additional file 11: MEME motifs and their associations with plants studied. These associations were determined by finding the genes that contain a given motif. This was done for (A) MSTs, (B) SUTs, and (C) SWEETS. The plant species with which motifs are most strongly associated was determined by finding the plant with the highest raw number of occurrences of the motif, as well as the representation of that motif as a percentage of instances of that motif in the MST, SUT or SWEET family. (XLSX 33 kb)

Additional file 12: MEME motifs and their associations with clades of MSTs, SUTs or SWEETS. These associations were determined by finding which genes to which these motifs belonged, and determined the clade in which these genes belonged. This was done for (A) MSTs, (B) SUTs, and (C) SWEETS. The clade with which motifs are most strongly associated was determined by finding the clade with the highest raw number of occurrences of the motif in that stage, as well as the representation of that motif as a percentage of motifs present in genes that showed strongest expression during that clade. In addition, MEME motifs unique to MST, SUT or SWEET clades were found and counted. (XLSX 43 kb)

Additional file 13: MEME motifs in the (A) MST, (B) SUT and (C) SWEET families of sugar transporters. For each motif, the stage in which the genes containing these motifs express most strongly is shown, as well as the plant and clade in which genes containing these motifs are most commonly found, and the consensus sequence for each motif. (XLSX 24 kb)

Additional file 14: Orthogroups in each stage of three *Orobanchaceae* life cycles. The data suggest different orthogroups being expressed in a variety of stages throughout the life cycle, especially in the MST family. (XLSX 21 kb)

Additional file 15: The species to which orthogroups are most frequently associated. Note the differences between the raw number of members in a certain orthogroup versus the percentage of sequences from a species that are members of that orthogroup. (XLSX 13 kb)

Additional file 16: The clades with which each orthogroup is most strongly associated. These orthogroups were identified in the A) MST, B) SUT and C) SWEET families. Note that many sugar transporter clades have at least one orthogroup that is unique to them. In the SUT family, one orthogroup has all SUT1 and SUT4 sequences studied as members. (XLSX 14 kb)

Additional file 17: Every sugar transporter gene, their strongest stage of expression, clade and orthogroup. The clade, stage of strongest expression and orthogroup are listed for genes in the A) MST, B) SUT and C) SWEET families. (XLSX 10 kb)

Additional file 18: The names of the genes on the phylogenetic trees, and their FASTA file aliases. The names as they appear in Figs. 1, 2, 3, 4 and 5 and in Additional files 1, 2, 3, 4, 5 and 7, 8, 9 are shorthand labels that show the species (e.g., *Striga hermonthica* is shortened to StHe), the clade, and an arbitrarily given number. (XLSX 17 kb)

Abbreviations

AG1: Ancient Group 1; AG2: Ancient Group 2; CDD: Conserved Domain Database; CDS: Coding sequence; CDS: Coding sequence; EBI: European Bioinformatics Institute; ERD6-like: Early Response to Dehydration 6-like; ESL: Early Response to Dehydration 6-like; EST: Expressed sequence tag; FPKM: Fragments per kilobase of unigenes length per million reads; GO: Gene Ontology; HGT: Horizontal gene transfer; INT: Inositol transporter; iTOL: Interactive Tree of Life; MAFFT: Multiple Alignment using Fast Fourier Transform; MEME: Multiple Em for Motif Elicitation; MFS: Major facilitator protein; MST: Monosaccharide transporter; MtN3: *Medicago truncatula* nodulin 3; NT: Nucleotide; NT: Nucleotide; pGLT: Plastidic glucose translocator; PLT: Polyol transporter; PMT: Polyol/monosaccharide transporter; PPGP: Parasitic Plant Genome Project; RAxML: Randomized Axelerated Maximum Likelihood; RNAi: RNA interference; SANSParallel: Suffix Array Neighborhood Search Parallel; STP: Sugar transport protein; Sugar_tr: Sugar transporter; SUT: Sucrose transporter; SWEET: Sugars will eventually be exported transporters; TM: Transmembrane; TMT: Tonoplast monosaccharide transporter; TST: Tonoplast sugar transporter; VGT: Vacuolar glucose transporter

Acknowledgments

Many thanks to Zhenzhen Yang for suggestions regarding analyses, and to the PPGP group for generating the transcriptomic data used in this study.

Funding

This research and the participation of VAM, EKW, YW, CWD, and MPT were supported by grants from the National Science Foundation Plant Genome Research Program (DBI-0701748 and IOS-1238057) awarded to James Westwood, CWD, John Yoder and MPT, with additional funding and support for VAM and MPT coming from a NSF grant (IOS-1213059) awarded to MPT.

Availability of data and materials

All data and material are available from the authors upon request.

Authors' contributions

MPT and CWD conceived of the project and directed all research activities. VAM performed the computational analyses of the sugar transporters with the assistance of YW. EKW created meta-assemblies and software and performed the transcriptome-wide expression analyses on which the computational analyses of the sugar transporters were based. All authors have assisted in the writing of the manuscript and have read and approved the final submitted version of the manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Biology, University of Virginia, Gilmer Hall 044, Charlottesville, VA 22904, USA. ²Department of Biology, Penn State University, University Park, PA 16802, USA. ³Present Address: Center for Quantitative Sciences, Vanderbilt University, 2220 Pierce Avenue, 571 Preston Research Building, Nashville, TN 37232-6848, USA.

Received: 12 September 2018 Accepted: 17 April 2019

Published online: 14 May 2019

References

- Ejeta G. The *Striga* scourge in Africa: a growing pandemic. In: Ejeta G, Gressel J, editors. Integrating new technologies for *Striga* control: towards ending the witch hunt. Singapore: World Scientific; 2007. p. 3–16.
- Scholes JD, Press MC. *Striga* infestation of cereal crops – an unsolved problem in resource limited agriculture. *Curr Opin Plant Biol.* 2008;11(2):180–6.
- Spallek T, Mutuku M, Shirasu K. The genus *Striga*: a witch profile. *Mol Plant Pathol.* 2013;14(9):861–9.
- Timko MP, Singh BB. Cowpea, a multifunctional legume. In: Moore PH, Ming R, editors. Genomics of tropical crop plants. New York: Springer New York; 2008. p. 227–58.
- Matvienko M, Torres MJ, Yoder JI. Transcriptional responses in the hemiparasitic plant *Triphysaria versicolor* to host plant signals. *Plant Physiol.* 2001;127(1):272–82.
- Cardwell KF, Lane JA. Effect of soils, cropping system and host phenotype on incidence and severity of *Striga gesnerioides* on cowpea in West Africa. *Agric Ecosyst Environ.* 1995;53(3):253–62.
- Bebawi FF, Eplee RE, Harris CE, Norris RS. Longevity of witchweed (*Striga asiatica*) seed. *Weed Sci.* 2017;32(4):494–7.
- Yoder JI. Host-plant recognition by parasitic *Scrophulariaceae*. *Curr Opin Plant Biol.* 2001;4(4):359–65.
- Sun Z, Matusova R, Bouwmeester H. Germination of *Striga* and chemical signaling involved: a target for control methods. In: Ejeta G, Gressel J, editors. Integrating new technologies for *Striga* control: towards ending the witch-hunt. Singapore: World Scientific; 2007. p. 47–60.
- Kokla A, Melnyk CW. Developing a thief: haustoria formation in parasitic plants. *Dev Biol.* 2018;442(1):53–9.
- Kujit J. The biology of parasitic flowering plants. Berkeley: University of California Press; 1969.
- dePamphilis CW, Palmer JD. Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. *Nature.* 1990;348(6299):337–9.
- Wolfe KH, Morden CW, Palmer JD. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proc Natl Acad Sci U S A.* 1992;89(22):10648–52.
- Verkleij J, Janssen J, Pieterse A. A preliminary study on isoenzyme variation in *Orobanche crenata* and *Orobanche aegyptiaca* from Syria. In: ter Borg SJ, editor. Proceedings of a workshop on biology and control of *Orobanche*. Wageningen, Netherlands: Dept. of Vegetation Science, Plant Ecology and Weed Science, Agricultural University; 1986. p. 154–9.
- Wolfe AD, Randle CP. Relationships within and among species of the holoparasitic genus *Hyobanche* (Orobanchaceae) inferred from ISSR banding patterns and nucleotide sequences. *Syst Bot.* 2001;26(1):120–30.
- Román B. Population diversity and dynamics of parasitic weeds. In: Joel DM, Gressel J, Musselman LJ, editors. Parasitic Orobanchaceae: parasitic mechanisms and control strategies. Berlin: Springer Berlin Heidelberg; 2013. p. 345–56.
- Katzir N, Portnoy V, Tzuri G, Joel DM, Castejón-Muñoz M. Use of random amplified polymorphic DNA (RAPD) markers in the study of the parasitic weed *Orobanche*. *Theor Appl Genet.* 1996;93(3):367–72.
- Paran I, Gidoni D, Jacobsohn R. Variation between and within broomrape (*Orobanche*) species revealed by RAPD markers. *Heredity.* 1997;78:68–74.
- Young ND, Steiner KE, dePamphilis CW. Evolution of parasitism in the Scrophulariaceae/Orobanchaceae: plastid gene sequences refute an evolutionary transition series. *Ann Mo Bot Gard.* 1999;86(4):876–93.

20. Schneeweiss GM, Colwell A, Park J-M, Jang C-G, Stuessy TF. Phylogeny of holoparasitic *Orobanche* (Orobanchaceae) inferred from nuclear ITS sequences. *Mol Phylogenet Evol.* 2004;30(2):465–78.
21. Bennett JR, Mathews S. Phylogeny of the parasitic plant family Orobanchaceae inferred from phytochrome A. *Am J Bot.* 2006;93(7):1039–51.
22. Kusumoto D, Goldwasser Y, Xie X, Yoneyama K, Takeuchi Y, Yoneyama K. Resistance of red clover (*Trifolium pratense*) to the root parasitic plant *Orobanche minor* is activated by salicylate but not by jasmonate. *Ann Bot.* 2007;100(3):537–44.
23. Swarbrick PJ, Huang K, Liu G, Slate J, Press MC, Scholes JD. Global patterns of gene expression in rice cultivars undergoing a susceptible or resistant interaction with the parasitic plant *Striga hermonthica*. *New Phytol.* 2008;179(2):515–29.
24. Huang K, Mellor KE, Paul SN, Lawson MJ, Mackey AJ, Timko MP. Global changes in gene expression during compatible and incompatible interactions of cowpea (*Vigna unguiculata* L.) with the root parasitic angiosperm *Striga gesnerioides*. *BMC Genomics.* 2012;13(1):402.
25. Torres MJ, Tomilov AA, Tomilova N, Reagan RL, Yoder JI. Pscroph, a parasitic plant EST database enriched for parasite associated transcripts. *BMC Plant Biol.* 2005;5(1):24.
26. Yoshida S, Ishida JK, Kamal NM, Ali AM, Namba S, Shirasu K. A full-length enriched cDNA library and expressed sequence tag analysis of the parasitic weed, *Striga hermonthica*. *BMC Plant Biol.* 2010;10(1):55.
27. Westwood JH, Yoder JI, Timko MP, dePamphilis CW. The evolution of parasitism in plants. *Trends Plant Sci.* 2010;15(4):227–35.
28. Westwood JH, dePamphilis CW, Das M, Fernández-Aparicio M, Honaas LA, Timko MP, et al. The parasitic plant genome Project: new tools for understanding the biology of *Orobanche* and *Striga*. *Weed Sci.* 2012;60(2):295–306.
29. Westwood JH, dePamphilis CW, Timko MP, Yoder JI Parasitic Plant Genome Project <http://ppgp.huck.psu.edu/>. Accessed 12 Aug 2017.
30. Wickett NJ, Honaas LA, Wafula EK, Das M, Huang K, Wu B, et al. Transcriptomes of the parasitic plant family Orobanchaceae reveal surprising conservation of chlorophyll synthesis. *Curr Biol.* 2011;21(24):2098–104.
31. Yang Z, Wafula EK, Honaas LA, Zhang H, Das M, Fernandez-Aparicio M, et al. Comparative transcriptome analyses reveal core parasitism genes and suggest gene duplication and repurposing as sources of structural novelty. *Mol Biol Evol.* 2015;32(3):767–90.
32. Honaas LA, Wafula EK, Yang Z, Der JP, Wickett NJ, Altman NS, et al. Functional genomics of a generalist parasitic plant: laser microdissection of host-parasite interface reveals host-specific patterns of parasite gene expression. *BMC Plant Biol.* 2013;13(1):9.
33. Williams LE, Lemoine R, Sauer N. Sugar transporters in higher plants—a diversity of roles and complex regulation. *Trends Plant Sci.* 2000;5(7):283–90.
34. Gamas P, de Carvalho Niebel F, Lescure N, Cullimore JV. Use of a subtractive hybridization approach to identify new *Medicago truncatula* genes induced during root nodule development. *Mol Plant Microbe Interact.* 1996;9(4):233–42.
35. Chen L-Q, Qu X-Q, Hou B-H, Sosso D, Osorio S, Fernie AR, et al. Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. *Science.* 2012;335(6065):207–11.
36. Chen L-Q, Hou B-H, Lalonde S, Takanaga H, Hartung ML, Qu X-Q, et al. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature.* 2010;468:527–32.
37. Doody J, Grace E, Kühn C, Simon-Plas F, Casieri L, Wipf D. Sugar transporters in plants and in their interactions with fungi. *Trends Plant Sci.* 2012;17(7):413–22.
38. Reuscher S, Akiyama M, Yasuda T, Makino H, Aoki K, Shibata D, et al. The sugar transporter inventory of tomato: genome-wide identification and expression analysis. *Plant Cell Physiol.* 2014;55(6):1123–41.
39. Reddy VS, Shlykov MA, Castillo R, Sun EI, Saier MH. The major facilitator superfamily (MFS) revisited. *FEBS J.* 2012;279(11):2022–35.
40. Kiyosue T, Abe H, Yamaguchi-Shinozaki K, Shinozaki K. ERD6, a cDNA clone for an early dehydration-induced gene of *Arabidopsis*, encodes a putative sugar transporter. *Biochim Biophys Acta Biomembr.* 1998;1370(2):187–91.
41. Lalonde S, Frommer W. SUT sucrose and MST monosaccharide transporter inventory of the *Selaginella* genome. *Front Plant Sci.* 2012;3:24.
42. Schneider S, Beyhl D, Hedrich R, Sauer N. Functional and physiological characterization of *Arabidopsis* INOSITOL TRANSPORTER1, a novel tonoplast-localized TRANSPORTER for myo-inositol. *Plant Cell.* 2008;20(4):1073–87.
43. Schneider S, Schneider A, Udvardi P, Hammes U, Gramann M, Dietrich P, et al. *Arabidopsis* INOSITOL TRANSPORTER2 mediates H⁺ symport of different inositol epimers and derivatives across the plasma membrane. *Plant Physiol.* 2007;145(4):1395–407.
44. Wormit A, Trentmann O, Feifer I, Lohr C, Tjaden J, Meyer S, et al. Molecular identification and physiological characterization of a novel monosaccharide transporter from *Arabidopsis* involved in vacuolar sugar transport. *Plant Cell.* 2006;18(12):3476–90.
45. Büttner M. The *Arabidopsis* sugar transporter (AtSTP) family: an update. *Plant Biol.* 2010;12:35–41.
46. Kühn C, Grof CPL. Sucrose transporters of higher plants. *Curr Opin Plant Biol.* 2010;13(3):287–97.
47. Liesche J, Krügel U, He H, Chincinska I, Hackel A, Kühn C. Sucrose transporter regulation at the transcriptional, post-transcriptional and post-translational level. *J Plant Physiol.* 2011;168(12):1426–33.
48. Bürkle L, Hibberd JM, Quick WP, Kühn C, Hirner B, Frommer WB. The H⁺-sucrose cotransporter NtSUT1 is essential for sugar export from tobacco leaves. *Plant Physiol.* 1998;118(1):59–68.
49. Slewinski TL, Garg A, Johal GS, Braun DM. Maize SUT1 functions in phloem loading. *Plant Signal Behav.* 2010;5(6):687–90.
50. Milne RJ, Grof CPL, Patrick JW. Mechanisms of phloem unloading: shaped by cellular pathways, their conductances and sink function. *Curr Opin Plant Biol.* 2018;43:8–15.
51. Barker L, Kühn C, Weise A, Schulz A, Gebhardt C, Hirner B, et al. SUT2, a putative sucrose sensor in sieve elements. *Plant Cell.* 2000;12(7):1153–64.
52. Sauer N. Molecular physiology of higher plant sucrose transporters. *FEBS Lett.* 2007;581(12):2309–17.
53. Weise A, Barker L, Kühn C, Lalonde S, Buschmann H, Frommer WB, et al. A new subfamily of sucrose transporters, SUT4, with low affinity/high capacity localized in enucleate sieve elements of plants. *Plant Cell.* 2000;12(8):1345–55.
54. Frost CJ, Nyamdari B, Tsai C-J, Harding SA. The tonoplast-localized sucrose transporter in *Populus* (PtaSUT4) regulates whole-plant water relations, responses to water stress, and photosynthesis. *PLoS One.* 2012;7(8):e44467.
55. Chincinska I, Gier K, Krügel U, Liesche J, He H, Grimm B, et al. Photoperiodic regulation of the sucrose transporter StSUT4 affects the expression of circadian-regulated genes and ethylene production. *Front Plant Sci.* 2013;4:26.
56. Flemetakis E, Dimou M, Cotzur D, Efroze RC, Aivalakis G, Colebatch G, et al. A sucrose transporter, LjSUT4, is up-regulated during *Lotus japonicus* nodule development. *J Exp Bot.* 2003;54(388):1789–91.
57. Durand M, Mainson D, Porcheron B, Mauroussat L, Lemoine R, Pourtau N. Carbon source–sink relationship in *Arabidopsis thaliana*: the role of sucrose transporters. *Planta.* 2018;247(3):587–611.
58. Peng D, Gu X, Xue L-J, Leebens-Mack JH, Tsai C-J. Bayesian phylogeny of sucrose transporters: ancient origins, differential expansion and convergent evolution in monocots and dicots. *Front Plant Sci.* 2014;5:615.
59. Xuan YH, Hu YB, Chen L-Q, Sosso D, Ducat DC, Hou B-H, et al. Functional role of oligomerization for bacterial and plant SWEET sugar transporter family. *Proc Natl Acad Sci.* 2013;110(39):E3685–94.
60. Conserved Domains Database (CDD) and resources. National Center for biotechnology information, Bethesda. 2017. <https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>. Accessed 1 Dec 2017.
61. Ponting CP, Mott R, Bork P, Copley RR. Novel protein domains and repeats in *Drosophila melanogaster*: insights into structure, function, and evolution. *Genome Res.* 2001;11(12):1996–2008.
62. Bush DR. Sugar transporters in plant biology. *Curr Opin Plant Biol.* 1999;2(3):187–91.
63. Lingner U, Münch S, Deising HB, Sauer N. Hexose transporters of a hemibiotrophic plant pathogen: functional variations and regulatory differences at different stages of infection. *J Biol Chem.* 2011;286(23):20913–22.
64. Chen L-Q. SWEET sugar transporters for phloem transport and pathogen nutrition. *New Phytol.* 2014;201(4):1150–5.
65. Péron T, Candat A, Montiel G, Veronesi C, Macherel D, Delavault P, et al. New insights into phloem unloading and expression of sucrose transporters in vegetative sinks of the parasitic plant *Phelipanche ramosa* L. (Pomel). *Front Plant Sci.* 2017;7:2048.
66. Olmstead RG, de Pamphilis CW, Wolfe AD, Young ND, Elisons WJ, Reeves PA. Disintegration of the Scrophulariaceae. *Am J Bot.* 2001;88(2):348–61.
67. Johnson DA, Thomas MA. The monosaccharide transporter gene family in *Arabidopsis* and rice: a history of duplications, adaptive evolution, and functional divergence. *Mol Biol Evol.* 2007;24(11):2412–23.
68. Yuan M, Wang S. Rice MtN3/saliva/SWEET family genes and their homologs in cellular organisms. *Mol Plant.* 2013;6(3):665–74.
69. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 2015;16(1):157.

70. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Čech M, et al. The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 2018;46(W1):W537–44.
71. Bromham L, Cowman PF, Lanfear R. Parasitic plants have increased rates of molecular evolution across all three genomes. *BMC Evol Biol.* 2013;13(1):126.
72. Těšitel J, Plavcová L, Cameron DD. Interactions between hemiparasitic plants and their hosts. *Plant Signal Behav.* 2010;5(9):1072–6.
73. Yang Z, Zhang Y, Wafula EK, Honaas LA, Ralph PE, Jones S, et al. Horizontal gene transfer is more frequent with increased heterotrophy and contributes to parasite adaptation. *Proc Natl Acad Sci.* 2016;113(45):E7010–9.
74. Yoshida S, Maruyama S, Nozaki H, Shirasu K. Horizontal gene transfer by the parasitic plant *Striga hermonthica*. *Science.* 2010;328(5982):1128.
75. Zhang Y, Fernandez-Aparicio M, Wafula EK, Das M, Jiao Y, Wickett NJ, et al. Evolution of a horizontally acquired legume gene, albumin 1, in the parasitic plant *Phelipanche aegyptiaca* and related species. *BMC Evol Biol.* 2013;13(1):48.
76. Bezrutczyk M, Hartwig T, Horschman M, Char SN, Yang J, Yang B, et al. Impaired phloem loading in zmsweet13 a, b, c sucrose transporter triple knock-out mutants in *Zea mays*. *New Phytol.* 2018;218(2):594–603.
77. Büttner M. The monosaccharide transporter (–like) gene family in *Arabidopsis*. *FEBS Lett.* 2007;581(12):2318–24.
78. Weber A, Servaites JC, Geiger DR, Kofler H, Hille D, Gröner F, et al. Identification, purification, and molecular cloning of a putative plastidic glucose translocator. *Plant Cell.* 2000;12(5):787–801.
79. Nativ N, Hacham Y, Hersheshorn J, Dor E, Amir R. Metabolic investigation of Egyptian broomrape (*Phelipanche aegyptiaca*) reveals significant changes during developmental stages and in its different organs. *Front Plant Sci.* 2017;8:491.
80. Slewinski TL. Diverse functional roles of monosaccharide transporters and their homologs in vascular plants: a physiological perspective. *Mol Plant.* 2011;4(4):641–62.
81. Richter A, Popp M. The physiological importance of accumulation of cyclitols in *Viscum album* L*. *New Phytol.* 1992;121(3):431–8.
82. Hacham Y, Hersheshorn J, Dor E, Amir R. Primary metabolic profiling of Egyptian broomrape (*Phelipanche aegyptiaca*) compared to its host tomato roots. *J Plant Physiol.* 2016;205:11–9.
83. Noiraud N, Maurouset L, Lemoine R. Transport of polyols in higher plants. *Plant Physiol Biochem.* 2001;39(9):717–28.
84. Mäkinen KK, Söderling E. A quantitative study of mannitol, sorbitol, xylitol, and xylose in wild berries and commercial fruits. *J Food Sci.* 1980;45(2):367–71.
85. Aluri S, Büttner M. Identification and functional expression of the *Arabidopsis thaliana* vacuolar glucose transporter 1 and its role in seed germination and flowering. *Proc Natl Acad Sci.* 2007;104(7):2537–42.
86. Kühn C, Hajirezaei M-R, Fernie AR, Roessner-Tunali U, Czechowski T, Hirner B, et al. The sucrose transporter StSUT1 localizes to sieve elements in potato tuber phloem and influences tuber physiology and development. *Plant Physiol.* 2003;131(1):102–13.
87. Westwood JH, Roney JK, Khatibi PA, Stromberg VK. RNA translocation between parasitic plants and their hosts. *Pest Manag Sci.* 2009;65(5):533–9.
88. Shahid S, Kim G, Johnson NR, Wafula E, Wang F, Coruh C, et al. MicroRNAs from the parasitic plant *Cuscuta campestris* target host messenger RNAs. *Nature.* 2018;553:82–5.
89. Yoder JI, Gunathilake P, Wu B, Tomilova N, Tomilov AA. Engineering host resistance against parasitic weeds with RNA interference. *Pest Manag Sci.* 2009;65(5):460–6.
90. Frommer WB. Sucrose transporters and methods of generating pathogen-resistant plants. In: Google Patents; 2014.
91. Bock R, Jiang Z, Heckel DG, and Khan SA. "Plant protection from a pest or pathogen by expression of double-stranded RNAs in the plastid." U.S. Patent Application 15/537,410, filed November 30, 2017.
92. Fernández-Aparicio M, Reboud X, Gibot-Leclerc S. Broomrape Weeds. Underground mechanisms of parasitism and associated strategies for their control: a review. *Front Plant Sci.* 2016;7:135.
93. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 2009;37:W202–8.
94. Rolland F, Baena-Gonzalez E, Sheen J. Sugar sensing and signaling in plants: conserved and novel mechanisms. *Annu Rev Plant Biol.* 2006;57(1):675–709.
95. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 2011;29:644–52.
96. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf.* 2011;12(1):323.
97. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods.* 2012;9:357–9.
98. NCBI Protein Database. National Center for biotechnology information, Bethesda. 2017. <http://www.ncbi.nlm.nih.gov/protein>. Accessed 1 Jun 2017.
99. NCBI Nucleotide Database. National Center for biotechnology information, Bethesda. 2017. <http://www.ncbi.nlm.nih.gov/nucleotide>. Accessed 1 Jun 2017.
100. NCBI EST Database. National Center for biotechnology information, Bethesda. 2017. <http://www.ncbi.nlm.nih.gov/est>. Accessed 1 Jun 2017.
101. Hellsten U, Wright KM, Jenkins J, Shu S, Yuan Y, Wessler SR, et al. Fine-scale variation in meiotic recombination in *Mimulus* inferred from population shotgun sequencing. *Proc Natl Acad Sci.* 2013;110(48):19478–82.
102. Hartmann U, Vision T, Phillips, J. MIMULUS Project Home. www.mimulusevolution.org/ (2010). Accessed 4 May 2017.
103. Phytozome v12.1. Department of energy, Office of Science, Washington, DC. 2017. <https://phytozome.jgi.doe.gov/pz/portal.html> Accessed 12 Aug 2017.
104. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 2012;40:D1178–86.
105. Lipman D, Pearson W. Rapid and sensitive protein similarity searches. *Science.* 1985;227(4693):1435–41.
106. Pearson WR, Lipman DJ. Improved tools for biological sequence comparison. *Proc Natl Acad Sci.* 1988;85(8):2444–8.
107. Pearson WR. UVA FASTA Server. https://fasta.bioch.virginia.edu/fasta_www2/fasta_list2.shtml. Accessed 3 Jun 2017.
108. Stajich JE, Block D, Boulez K, Brenner SE, Chervitz SA, Dagdigan C, et al. The Bioperl toolkit: Perl modules for the life sciences. *Genome Res.* 2002;12(10):1611–8.
109. Fields CJ. BioPerl 1.7.1. <http://search.cpan.org/~cjfields/BioPerl-1.007001/>. Accessed 3 Dec 2016.
110. Misra VA, Wang Y, Timko MP. A compendium of transcription factor and transcriptionally active protein coding gene families in cowpea (*Vigna unguiculata* L.). *BMC Genomics.* 2017;18(1):898.
111. Somervuo P, Holm L. SANSparallel: interactive homology search against Uniprot. *Nucleic Acids Res.* 2015;43(W1):W24–9.
112. Holm L. SANSparallel. <http://ekhidna2.biocenter.helsinki.fi/sans/>. Accessed on 4 May 2017.
113. The Uniprot Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 2017;45(D1):D158–69.
114. MEME (Multiple Em for Motif Elicitation). <http://meme-suite.org/tools/meme>. Accessed on 12 Aug 2017.
115. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30(4):772–80.
116. Katoh K: MAFFT v7.312. <http://mafft.cbrc.jp/alignment/software/> (2013). Accessed on 1 Nov 2017.
117. Banks JA, Nishiyama T, Hasebe M, Bowman JL, Gribskov M, dePamphilis C, et al. The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science.* 2011;332(6032):960–3.
118. The Amborella Genome Project. The *Amborella* genome and the evolution of flowering plants. *Science.* 2013;342(6165):1241089.
119. Chen F, Mackey AJ, Stoekert CJ Jr, Roos DS. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res.* 2006;34(suppl_1):D363–8.
120. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics.* 2009;25(15):1972–3.
121. Larkin MA, Blackshields G, Brown N, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics.* 2007;23(21):2947–8.
122. Marchler-Bauer A, Derbyshire M, Gonzales N, Lu S, Chitsaz F, Geer L, Geer R, He J, Gwadz M, Hurwitz D. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 2015;43:D222–6.
123. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014;30(9):1312–3.
124. The Exelexis Lab. RAxML version 8.2.11. <https://sco.h-its.org/exelixis/web/software/raxml/index.html>. (2017) Accessed on 12 Aug 2017.
125. Hurvich CM, Tsai C-L. Regression and time series model selection in small samples. *Biometrika.* 1989;76(2):297–307.
126. Letunic I, Bork P. Interactive tree of life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics.* 2007;23(1):127–8.
127. Interactive Tree Of Life (iTOL) version 4. <http://itol.embl.de>. Accessed on 16 Dec 2017.