Weed Science

www.cambridge.org/wsc

Transcriptomics of host-specific interactions in natural populations of the parasitic plant purple witchweed (*Striga hermonthica*)

Lua Lopez[1], Emily S. Bellis[2], Eric Wafula[3], Sarah J. Hearne[4,5], Loren Honaas[6], Paula E. Ralph[7], Michael P. Timko[8], Nnanna Unachukwu[9], Claude W. dePamphilis[10] and Jesse R. Lasky[11]

[1]Postdoctoral Research Associate, The Pennsylvania State University, University Park, PA, USA; Research Assistant Professor, Binghamton University, State University of New York, Binghamton, NY, USA; [2]Postdoctoral Research Fellow, The Pennsylvania State University, University Park, PA, USA; [3]Bioinformatics Programmer, The Pennsylvania State University, University Park, PA, USA; [4]Research Scientist, International Institute of Tropical Agriculture, Ibadan, Nigeria; [5]Principal Research Scientist, International Maize and Wheat Improvement Center, Texcoco, Mexico; [6]Research Molecular Biologist, Physiology and Pathology of Tree Fruits Research, USDA-ARS, Wenatchee, WA, USA; [7]Lab Manager, The Pennsylvania State University, University Park, PA, USA; [8]Professor, University of Virginia, Charlottesville, VA, USA; [9]Senior Research Supervisor, International Institute of Tropical Agriculture, Ibadan, Nigeria; [10]Professor, The Pennsylvania State University, University Park, PA, USA and [11]Assistant Professor, The Pennsylvania State University, University Park, PA, USA

Research Article

Cite this article: Lopez L, Bellis ES, Wafula E, Hearne SJ, Honaas L, Ralph PE, Timko MP, Unachukwu N, dePamphilis CW, Lasky JR (2019) Transcriptomics of host-specific interactions in natural populations of the parasitic plant purple witchweed (*Striga hermonthica*). Weed Sci. doi: 10.1017/wsc.2019.20

Received: 25 January 2019
Revised: 29 March 2019
Accepted: 13 April 2019

Associate Editor:
Sara Martin, Agriculture and Agri-Food Canada

Keywords:
Allele frequency; gene expression; host–parasite interactions; parasitic plant; RNAseq

Author for correspondence:
Lua Lopez, Email: lopezperez.lua@gmail.com

Abstract

Host-specific interactions can maintain genetic and phenotypic diversity in parasites that attack multiple host species. Host diversity, in turn, may promote parasite diversity by selection for genetic divergence or plastic responses to host type. The parasitic weed purple witchweed [*Striga hermonthica* (Delile) Benth.] causes devastating crop losses in sub-Saharan Africa and is capable of infesting a wide range of grass hosts. Despite some evidence for host adaptation and host-by-*Striga* genotype interactions, little is known about intraspecific *Striga* genomic diversity. Here we present a study of transcriptomic diversity in populations of *S. hermonthica* growing on different hosts (maize [*Zea mays* L.] vs. grain sorghum [*Sorghum bicolor* (L.) Moench]). We examined gene expression variation and differences in allelic frequency in expressed genes of aboveground tissues from populations in western Nigeria parasitizing each host. Despite low levels of host-based genome-wide differentiation, we identified a set of parasite transcripts specifically associated with each host. Parasite genes in several different functional categories implicated as important in host–parasite interactions differed in expression level and allele on different hosts, including genes involved in nutrient transport, defense and pathogenesis, and plant hormone response. Overall, we provide a set of candidate transcripts that demonstrate host-specific interactions in vegetative tissues of the emerged parasite *S. hermonthica*. Our study shows how signals of host-specific processes can be detected aboveground, expanding the focus of host–parasite interactions beyond the haustorial connection.

Introduction

The dependence of parasites on their hosts and the high virulence of many parasites make the host–parasite relationship one of the strongest biotic interactions. Due to the diverse physiologies and molecular phenotypes of different hosts, these intimate relationships pose a challenge for generalist parasite species that attack multiple host species. Parasites with broad host ranges often display high levels of intraspecific variation (Archie and Ezenwa 2011; Kaci-Chaouch et al. 2008). This may be partly maintained by balancing selection within populations or local adaptation across populations on different hosts (Frank 1993). Within-host competitive interaction can also act as a maintaining force of intraspecific genetic diversity (Bashey 2015). Alternatively, in generalist parasites, plastic responses to hosts can be key to successful host generalism (Lajeunesse and Forbes 2002; Leggett et al. 2013). While gene-for-gene and simple genetic architectures of host–parasite interactions have been frequently studied (Guo et al. 2009; Watson 1970), we know less about host–parasite interactions with a complex genetic basis, such as those between parasitic plants and their hosts. As many parasitic plants are major agricultural pests, an important goal is to understand the genetics of parasitism across diverse host crops (Honaas et al. 2013; Yang et al. 2015).

Purple witchweed [*Striga hermonthica* (Delile) Bent.] (Orobanchaceae) is an obligate root hemiparasite of grasses in sub-Saharan Africa causing billions of dollars in crop loss annually (Spallek et al. 2013). *Striga hermonthica* commonly attacks grain sorghum [*Sorghum bicolor* (L.) Moench], maize (*Zea mays* L.), and pearl millet [*Pennisetum glaucum* (L.) R. Br.], with

increasing incidence of parasitism now being observed in sugarcane (*Saccharum officinarum* L.) and upland rice (*Oryza sativa* L.) fields (Spallek et al. 2013). Crop infestations of parasitic weeds have devastating effects on crop yield around the world, especially in Africa (Rodenburg et al. 2016). Farmers have reported losses between 20% and 80%, and high-density infestations might lead to the total loss of the crop and abandonment of a field (Atera et al. 2012). Particularly for *Striga* spp., weed control is challenging, because 1) the severe reaction of hosts to parasitism initiates almost immediately after haustorial attachment, negating the value of weeding of aboveground parasites as a crop-loss mitigation strategy; 2) mature plants produce tens to hundreds of thousands of tiny seeds that are readily dispersed by wind, water, contaminated grain, livestock, and people; and 3) *Striga* seeds are long-lived, capable of establishing a persistent seedbank. *Striga hermonthica* exhibits genetic variation in interactions both among and within crops with variability in host preference and virulence, hindering the development of resistant cultivars (Bozkurt et al. 2015; Koyama 2000; Olivier et al. 1998; Unachukwu et al. 2017). Reports of tolerant and/or resistant genotypes vary among crop species. For example, several grain sorghum genotypes have been shown to be tolerant or resistant to *Striga* (Haussmann et al. 2004; Hess et al. 1992; Maiti et al. 1984; Vogler et al. 1996). Meanwhile, tolerance or resistance in maize has rarely been reported (Amusan et al. 2008; Lane et al. 1997; Mutinda et al. 2018).

Genetic and phenotypic responses to host variation in parasitic plants of the Orobanchaceae family are likely complex, given the multiple developmental stages to interactions and mechanisms of host defenses that also vary among host species. In parasitic plants like *S. hermonthica*, researchers are only beginning to uncover the genetic and phenotypic basis of host-specific interactions (Ejeta 2007; Honaas et al. 2013; Swarbrick et al. 2008). Expression profiling of genes expressed at the host–parasite interface suggest that one possible strategy for parasitic plants is to express a set of shared parasitic genes irrespective of the host and a different set of genes that are expressed on specific hosts (Honaas et al. 2013). Yet, other studies have found limited evidence of genetic differentiation between *S. hermonthica* from different hosts of origin (Unachukwu et al. 2017; Welsh and Mohamed 2011). The dynamic and complex nature of variation in host-specific interactions in natural *S. hermonthica* populations challenges our understanding of its molecular basis.

Transcriptome profiling offers the opportunity to rapidly study functional variation in the genome and to connect this variation to specific genetic loci, evolutionary processes, and environmental conditions. By studying transcriptomic variation in individuals from natural populations, we can build a more mechanistic understanding of host–parasite interactions and their impacts on populations and communities (McGill et al. 2006). Recent examples demonstrate the utility of transcriptomic studies of wild organisms in natural systems (Swenson et al. 2017). Furthermore, gene expression analyses can help reveal complex mechanisms of host–parasite interactions (Ichihashi et al. 2015). Gene expression analysis in the parasite yellowbeak owl's-clover (*Triphysaria versicolor* Fisch. & C.A. Mey.) showed how this parasite relies on a generalist response in which functionally overlapping but distinct gene sets are expressed based on the parasitized host (Honaas et al. 2013). In a comprehensive comparative transcriptomic study of individuals in the Orobanchaceae family, Yang et al. (2015) identified a set of core parasitism genes used by members of three parasite species with different extents of nutritional dependence. The core parasitism genes belonged to specific members of gene families encoding cell wall–modifying enzymes and peroxidases, proteins known to be part of the parasitic invasion process (Antonova and Terborg 1996; Losner-Goshen et al. 1998; Pérez-de-Luque 2013; Singh and Singh 1993).

Despite prior efforts aimed at characterizing host-specific responses of *S. hermonthica*, genomic studies have been limited by the lack of functional genomic data (i.e., beyond reduced-representation sequencing). Transcriptomes of *S. hermonthica* have been studied across developmental stages and tissues on a single host under laboratory conditions (Yang et al. 2015; Parasitic Plant Genome Project, http://ppgp.huck.psu.edu). However, parasite transcriptomic profiling across hosts may be used to build a functional understanding of host specificity. Furthermore, growth under controlled laboratory conditions may not fully represent functional diversity of parasite response in natural environments. Here, we leverage transcriptomic data obtained from natural populations of *S. hermonthica* parasitizing different host species to investigate gene expression variation across hosts. Specifically, we analyze transcriptomes from *S. hermonthica* growing in the field on maize and grain sorghum to address two central questions about host–parasite interactions: 1) to what extent are *S. hermonthica* populations on different hosts divergent in genotype and expression phenotypes, and 2) which, if any, parasite genes exhibit host-specific expression?

Materials and methods

Study site and sampling

Field populations of *S. hermonthica* were sampled from farms located around the town of Mokwa in Niger State (Supplementary Figure 1; Supplementary Table 1). Mokwa is located at an elevation of 160 m above sea level and has a tropical savanna climate (Aw on the Koppen climate classification) (Geiger 1954). The average temperature and precipitation are 27.4 C and 1,056 mm yr^{-1}, respectively (Karger et al. 2017). More than 95% of annual rain tends to fall during a rainy season from April to October, which is followed by a dry period from November to March.

In November 2009, we sampled 14 populations of *S. hermonthica* from 14 fields: 3 growing on maize and 11 growing on grain sorghum (Supplementary Figure 1; Supplementary Table 1). From each population, we pooled young leaves from 10 to 20 individuals and, separately, we pooled flower buds from the same individuals. The tissues were immediately flash frozen in liquid nitrogen in the field.

RNA extraction, library preparation, and sequencing

For each pooled collection (leaf or flower bud tissue), the entire collection was emptied into a prechilled mortar and thoroughly mixed in a pregrinding step before sampling (and further grinding) the amount needed for single RNA isolations. RNA isolations were carried out using Ambion's RNAqueous™ Midi Kit with Plant Isolation Aid (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's recommendations. The quality of the resulting RNA was assessed using an Agilent Bioanalyzer (RNA integrity number [RIN] values $\geq$ 7.4, based on a scale of 1 to 10, with 1 being most degraded and 10 most intact). The RNA was precipitated (with 0.1 volumes sodium acetate and 3 volumes of 100% ethanol) to concentrate the RNA. Samples were DNase treated, column purified, and evaluated on the Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA) for quality and quantity. Finally, leaf and flower bud RNA samples were pooled in equimolar

https://doi.org/10.1017/wsc.2019.20

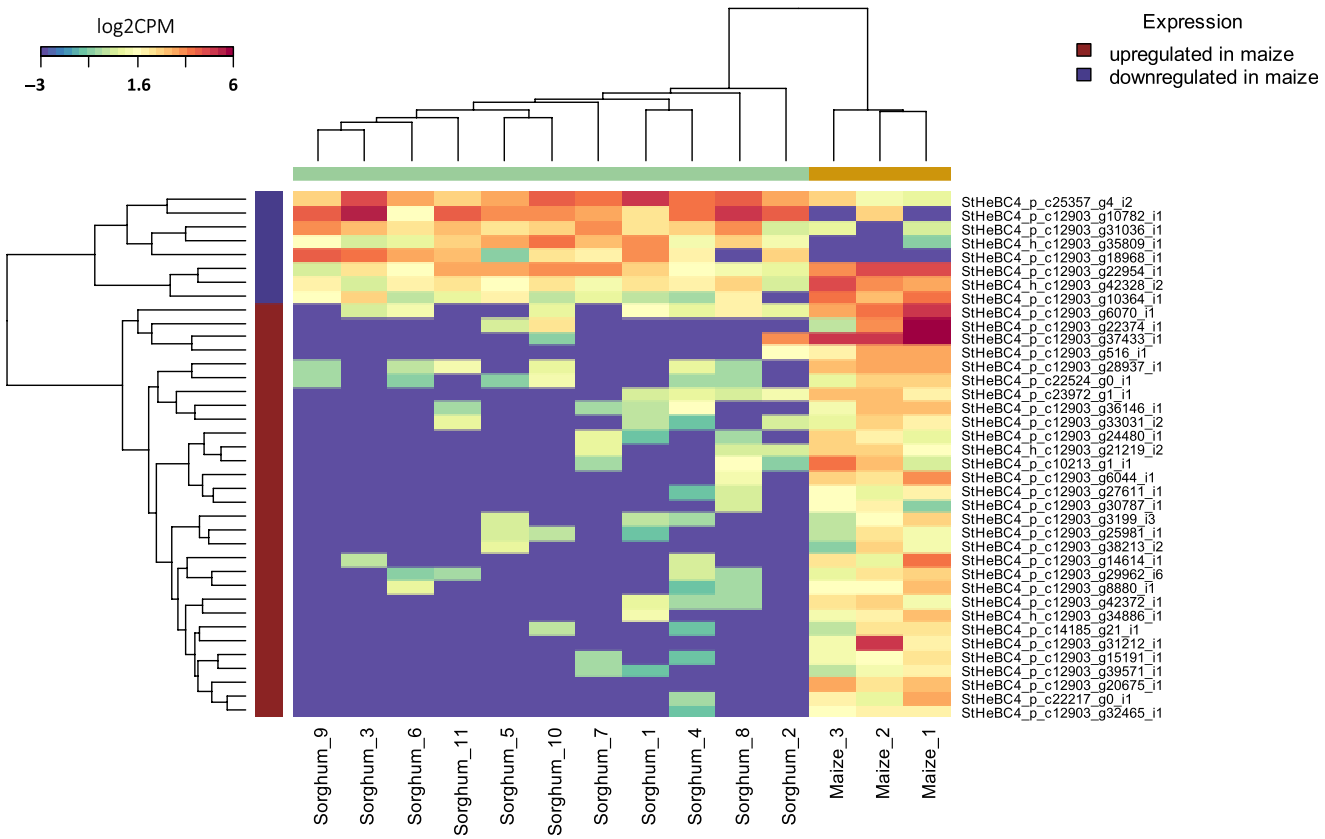

Figure 1. Heat map showing the expression pattern and clustering of the 38 differentially expressed transcripts ($P < 0.01$ and log fold change > 2) with at least 60% of their variation explained by host in *Striga hermonthica*. The dendogram shows that the two main branches cluster samples by host (sorghum vs. maize). Colors correspond to log2CPM (counts per million).

amounts for each of the 14 field populations. To make a cDNA library, 10 µg of each RNA pool was used.

Indexed cDNA libraries for paired-end sequencing were made using Illumina's mRNA Seq Sample Prep Kit and Multiplexing Sample Prep Oligonucleotide Kit (Illumina, San Diego, CA), with the size selection targeting an approximately 185-bp insert size. Because library concentrations were low after the final gel purification step, 3 to 12 additional cycles of amplification (with effective increases in library concentrations ranging from 10X to 100X) were performed to reach the target concentration of 10 nM before sequencing. Finally, libraries were validated by Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA) assay and qPCR. The qPCR validation was performed using KAPA SYBR® FAST qPCR Master Mix (F. Hoffmann-La Roche AG, Basel, Switzerland) and Illumina PhiX control DNA (PhiX control kit v. 2.0, CT-901-2001, Illumina, San Diego, CA) with a size-correction calculation according to the Kapa Biosystem manual. Libraries were sequenced as 75-bp PE (paired-end sequencing lanes) in two lanes (eight libraries per lane) on the Illumina GAIIx platform (Illumina, San Diego, CA). Each library was sequenced at the same sequencing depth in two lanes, and the resulting files were combined.

Data preprocessing and de novo assembly

Raw paired-end reads were trimmed and filtered using Trimmomatic v. 0.36 to remove adapter sequences, short reads, and low-quality reads from raw sequence data (Bolger et al. 2014). To remove adapter sequences, we looked for seed matches with a maximum of three mismatches and an initial length of 10. These seeds were extended and clipped when paired-end reads had a score of 30 or lower. We removed leading and trailing bases of low quality (Phred quality score <20) or N's. Likewise, we cut and removed the 3′ end of reads if the quality of a 4-base-wide sliding window dropped below 15. Finally, we discarded any reads remaining that were shorter than 50 bases long. Filtered reads were inspected using FastQC v. 0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects) to confirm that they met our quality standards.

We used the *S. hermonthica* de novo transcriptome assembly (StHeBC4) as a reference for both expression and single-nucleotide polymorphism (SNP) discovery analyses from the Parasitic Plant Genome Project (PPGP) website (http://ppgp.huck.psu.edu/download.php; http://bigdata.bx.psu.edu/PPGP_II_data). In brief, the reference de novo assembly was performed with Trinity (Grabherr et al. 2011) using RNA-Seq data from the following parasite developmental stages: roots of germinated seedlings after exposure to GR24, roots of germinated seedlings after exposure to host roots, expression in haustoria attached to host roots (prevascular connection), haustoria attached to host roots after vascular connection, late postattachment from belowground plants, stem and leaf tissue, and flowers (Westwood et al. 2012). The assembled contigs were postprocessed into nonredundant sets including predicted coding sequences and their corresponding translations using the PlantTribes pipeline (https://github.com/dePamphilis/PlantTribes). They were also screened against the grain sorghum genome and the National Center for Biotechnology Information (NCBI) nonredundant proteins database (nr) to remove host and other contaminant sequences. Detailed information about the PPGP parasitic plants reference transcriptome assemblies is provided on the PPGP website and described in Yang et al. (2015).

https://doi.org/10.1017/wsc.2019.20

Table 1. Top 1% of *Striga hermonthica* single-nucleotide polymorphisms (SNPs) with highest F_{ST} between hosts

Transcript[a]	Position[b]	RefAvFS[c]	RefAgFM[d]	F_{ST}[e]
StHeBC4_h_c12903_g20328_i2	688	0.798	0.431	0.112
StHeBC4_h_c17987_g1_i1	519	0.977	0.744	0.136
StHeBC4_p_c11301_g0_i1	552	1	0.826	0.142
StHeBC4_p_c12903_g12054_i1	577*	0.957	0.646	0.166
StHeBC4_p_c12903_g13725_i1	796*	1	0.633	0.313
StHeBC4_p_c12903_g1495_i2	367*	1	0.810	0.156
StHeBC4_p_c12903_g1513_i2	1373	0.959	0.717	0.118
StHeBC4_p_c12903_g16665_i2	1654	0.995	0.792	0.150
StHeBC4_p_c12903_g182_i1	238*	0.996	0.846	0.107
StHeBC4_p_c12903_g18522_i1	219	0.952	0.667	0.140
	156	0.934	0.667	0.111
StHeBC4_p_c12903_g19726_i1	484	0.821	0.483	0.102
StHeBC4_p_c12903_g223_i3	301*	0.936	0.594	0.167
	323*	0.995	0.833	0.117
StHeBC4_p_c12903_g25706_i3	633	0.928	0.613	0.139
StHeBC4_p_c12903_g2606_i2	786	1	0.875	0.101
StHeBC4_p_c12903_g26210_i3	1852	0.991	0.833	0.102
StHeBC4_p_c12903_g30530_i1	1592*	0.101	0.521	0.192
	1605	0.114	0.454	0.128
StHeBC4_p_c12903_g3163_i1	841	0.778	0.417	0.105
StHeBC4_p_c12903_g36387_i1	1628*	0.981	0.782	0.115
StHeBC4_p_c12903_g38045_i1	1514	1	0.844	0.127
StHeBC4_p_c12903_g491_i2	1608	0.962	0.666	0.162
StHeBC4_p_c12903_g6670_i1	405	0.089	0.527	0.216
StHeBC4_p_c12903_g9097_i1	322	0.989	0.784	0.135
StHeBC4_p_c15203_g0_i3	551	0.757	0.395	0.101
	1097	0.932	0.578	0.171
	1173*	0.974	0.667	0.191
StHeBC4_p_c15560_g4_i2	722	0.135	0.589	0.194
StHeBC4_p_c17241_g1_i1	2362	0.999	0.750	0.206
StHeBC4_p_c17257_g0_i1	193*	0.898	0.542	0.145
	696	0.438	0.837	0.108
StHeBC4_p_c17662_g0_i1	474	0.957	0.692	0.132
StHeBC4_p_c17982_g0_i1	435	0.104	0.421	0.119
StHeBC4_p_c18022_g0_i1	850	0.945	0.667	0.129
StHeBC4_p_c18111_g0_i2	1552*	1	0.778	0.183
StHeBC4_p_c18261_g0_i1	799*	0.987	0.778	0.135
StHeBC4_p_c18413_g0_i2	962*	1	0.812	0.154
StHeBC4_p_c18675_g1_i2	260	1	0.718	0.236
StHeBC4_p_c20635_g1_i1	1244	1	0.833	0.136
StHeBC4_p_c23303_g1_i1	651	0.883	0.559	0.117
StHeBC4_p_c23608_g0_i1	570	1	0.791	0.172
StHeBC4_p_c24035_g2_i1	643	1	0.873	0.102
StHeBC4_p_c24289_g0_i1	267*	1	0.758	0.201
StHeBC4_p_c24517_g0_i1	683	0.831	0.438	0.137
StHeBC4_p_c25018_g1_i1	1029	1	0.870	0.105
StHeBC4_p_c25267_g0_i1	77*	0.918	0.556	0.165
StHeBC4_p_c25745_g16_i3	1531	0.029	0.373	0.216
StHeBC4_p_c26238_g1_i1	283*	1	0.794	0.170
StHeBC4_p_c26242_g25_i1	2801	1	0.836	0.133
StHeBC4_p_c26384_g1_i2	659	0.962	0.705	0.132
StHeBC4_p_c26622_g0_i1	619	0.964	0.738	0.110
StHeBC4_p_c26760_g0_i6	1269	1	0.750	0.208
StHeBC4_p_c26874_g2_i2	215	0.964	0.750	0.102
StHeBC4_p_c26940_g5_i2	286*	1	0.810	0.155
StHeBC4_p_c27060_g3_i5	161*	1	0.764	0.195

[a] Transcripts with nonconservative amino acid changes are indicated in bold.
[b] Position of the SNPs within the transcript. Positions with a nonsynonymous amino acid change are indicated with an asterisk (*).
[c] RefAvFS, averaged frequency of the reference allele in a population with sorghum host.
[d] RefAgFM, averaged frequency of the reference allele in population with maize host.
[e] F_{ST} is a measure of structure in natural populations based on expected heterozygosity for each host-specific population relative to the total population.

SNP discovery and allelic frequency analyses

We further sought to understand host-specific differences in *S. hermonthica* parasitizing different hosts by characterizing SNPs in aboveground expressed transcripts. Trimmed reads were mapped to the de novo assembled transcriptome using *aln* and *sampe* commands in BWA, recommended for 75-bp paired-end reads (Li and Durbin 2009). We converted the SAM files containing the mapped reads to BAM with SAMtools v. 1.5 (Li et al. 2009) and used Picard Tools v. 2.8.2 (https://github.com/broadinstitute/picard) to sort and remove duplicate mappings. We filtered the sorted and deduplicated BAM files with SAMtools v. 1.5 and kept only proper pairs with map quality greater than 30. We used the

https://doi.org/10.1017/wsc.2019.20

BAM files containing high-quality pair-ended mappings to produce an mpileup file in SAMtools with minimum base quality of 20.

We next identified host-associated allelic variation. Allelic variation in transcriptomes can be due to host-associated allelic frequency or allele-specific expression. Both processes are of interest for identifying loci associated with host specificity. It has been shown that SNPs can be accurately called from pooled transcriptome data using conservative parameters to select sites for allelic frequency estimations (Konczal et al. 2013). Thus, we called genotypes in BCFtools for biallelic SNPs with a minimum coverage of 5X and removed SNPs with coverage less than 5X in more than one sample. To investigate how the individual samples clustered based on their genotype, we performed a principal components analysis (PCA) with the R package SNPRelate using the default options (Zheng et al. 2012a). For allelic frequency estimation, we only used SNPs that were present in all samples. Allele frequencies at the remaining polymorphic sites were calculated for each library with VCFtools based on counts of reference and alternate alleles. We grouped libraries into two populations: a population for host maize ($n = 3$) and a population for host grain sorghum ($n = 11$). For each SNP, we computed fixation index (F_{ST}) based on expected heterozygosity for each host-specific population relative to the total population using a custom perl script (Nei 1977). We selected SNPs in the upper 1% tail (outliers) of F_{ST} between host populations. To gauge whether differentiation of these "outlier" SNPs exceeded that expected by chance, we calculated F_{ST} values for three grain sorghum populations located nearest the maize populations (Sorghum_1, Sorghum_5, and Sorghum_11) versus the eight remaining sorghum populations. The sorghum versus sorghum comparison average F_{ST} was 0.017, and the maize versus sorghum comparison average F_{ST} was 0.016. The distribution of F_{ST} values for this sorghum versus sorghum analysis was compared with the distribution of F_{ST} values obtained for the maize populations versus the eight grain sorghum populations (excluding Sorghum_1, Sorghum_5, and Sorghum_11).

We annotated candidate outlier SNPs following the same steps used for the differentially expressed transcripts and further characterized them with respect to their predicted impact on coding sequence. We aligned the transcripts containing the outlier SNPs to their predicted coding sequence (CDS) using Geneious 9.1.8 to determine whether the SNP was located within the coding region and whether it coded for a synonymous or nonsynonymous mutation. Nonsynonymous changes were categorized as conservative or nonconservative (see "Results and Discussion"), and we used the score from the substitution matrix BLOSUM-62 to indicate the structural similarity between the amino acids.

Differential expression analysis

We used the alignment-free package Salmon to perform differential expression (DE) analysis (Patro et al. 2017). Salmon is a tool for expression quantification that couples the concept of quasi-mapping with a two-phase inference procedure. To characterize host-related expression patterns, the output matrix containing count estimates from Salmon were analyzed with edgeR (Robinson et al. 2010). edgeR identifies differentially expressed transcripts based on the assumption that the number of reads produced by each transcript is proportional to its abundance. Because Salmon produces estimates instead of total counts, we followed the guidelines and rounded the estimates for downstream analysis with edgeR. Populations of *S. hermonthica* parasitizing the same host crop (maize or grain sorghum) were considered biological replicates, with 3 replicates for maize and 11 for grain sorghum. Before performing the DE analysis, we excluded transcripts with very low expression values using a cut-off of 1 CPM (counts per million) in at least three samples (number of biological replicates for libraries with maize as host crop), and we used the trimmed mean of M-values method for normalizing the counts. We used an exact test implemented in edgeR, using host as group, in the transcripts remaining after filtering. We focused on those transcripts that were differentially expressed with a log fold change (logFC) significantly greater than 2 and a P-value < 0.01.

To estimate the importance of host-associated expression genome-wide, we implemented variation partitioning in the R package Vegan 2.3-5 (Oksanen et al. 2013). Variation partitioning can be used to estimate the proportion of variance explained by covariates (host) with multivariate responses (transcriptomes). We also tested whether the effect of host (proportion of transcriptomic variation explained) was significant genome-wide using permutation tests that randomized host identity across libraries. Finally, we used the R package VariancePartition (Hoffman and Schadt 2016) to quantify the contribution of host to variance explained at each level of individual transcript.

To focus on transcripts with a significant and large effect size of host on expression, we combined the results from the DE and VariancePartition analysis. We selected significantly differentially expressed transcripts ($P < 0.01$ and $logFC > 2$) with at least 60% of the transcript expression variation explained by host for annotation with Blast2GO (Conesa and Götz 2008; Conesa et al. 2005). We annotated these focal transcripts using the Basic Local Alignment Search Tool (blastx) (e-value cutoff 1.0×10^{-5}) against the nr database from the NCBI (Pruitt et al. 2007) and InterPro from the European Bioinformatics Institute (Finn et al. 2017). We ran an enzyme code search and a Kyoto Encyclopedia of Genes and Genomes pathway analysis (https://www.genome.jp/kegg; Kanehisa and Goto 2000). Then, we retrieved gene ontology (GO) terms (http://www.geneontology.org; Ashburner et al. 2000; Gene Ontology Consortium 2017) from the blastx, InterPro, and enzyme hits. We extracted functional gene categories for each transcript, comparing its translated sequence with the Embryophyta orthologue database in OrthoDB v. 9.1 (Zdobnov et al. 2017). Finally, we searched for noncoding RNA families and other structured RNA elements using the Rfam database (Griffiths-Jones et al. 2003).

Our libraries captured differentially expressed transcripts in aboveground tissue of parasites but did not contain information on other tissues relevant for host–parasite interactions. To investigate tissue specificity of the focal DE transcripts, we used *S. hermonthica* RNA-Seq expression data across parasite developmental stages obtained from the PPGP v. 2 (PPGP II), a large database of replicated expression data for species of Orobanchaceae, including *S. hermonthica* collected in Borno State (Nigeria) that was growing on grain sorghum (http://bigdata.bx.psu.edu/PPGP_II_data). The developmental stages represented in this data are: imbibed seeds (stage 0), roots of germinated seedlings after exposure to GR24 (stage 1), roots of germinated seedlings after exposure to host roots (stage 2), expression in haustoria attached to host roots (prevascular connection) (stage 3), haustoria attached to host roots after vascular connection (stage 4), late postattachment from belowground plants (stage 5), stem and leaf tissue (stage 6.1), and flower tissue (stage 6.2) (Westwood et al. 2012).

https://doi.org/10.1017/wsc.2019.20

Table 2. Annotation of *Striga hermonthica* transcripts with highest F_{ST} between hosts containing nonsynonymous single-nucleotide polymorphisms

Transcript[a]	Description[b]	Functional gene category[c]
StHeBC4_p_c12903_g12054_i1*	NA	• Posttranslational modification, protein turnover, chaperones • Cell wall/membrane/envelope biogenesis • Amino acid transport and metabolism • Carbohydrate transport and metabolism
StHeBC4_p_c12903_g13725_i1*	3-Isopropylmalate dehydratase large chloroplastic	• Coenzyme transport and metabolism
StHeBC4_p_c12903_g1495_i2	WRKY protein	• Mobilome: prophages and transposons • Replication, recombination, and repair • Translation, ribosomal structure, and biogenesis
StHeBC4_p_c12903_g182_i1*	ABC transporter C family member 10, ATP-energized glutathione *S*-conjugate pump 14, glutathione *S*-conjugate-transporting ATPase 14, multidrug resistance-associated 14	NA
StHeBC4_p_c12903_g223_i3*	UDP-glycosyltransferase 76F1	• Cell wall/membrane/envelope biogenesis • Carbohydrate transport and metabolism
StHeBC4_p_c12903_g30530_i1*	E3 ubiquitin ligase DRIP2, DREB2A-interacting 2, RING-type E3 ubiquitin transferase DRIP2	• Signal transduction mechanisms
StHeBC4_p_c12903_g36387_i1	NADH dehydrogenase	NA
StHeBC4_p_c15203_g0_i3*	Primary amine oxidase	• Amino acid transport and metabolism • Secondary metabolite biosynthesis, transport, and catabolism
StHeBC4_p_c17257_g0_i1	50S ribosomal chloroplastic ame, CL1	• Translation, ribosomal structure, and biogenesis
StHeBC4_p_c18111_g0_i2*	Probable monogalactosyldiacylglycerol chloroplastic, MGDG synthase type A	• Carbohydrate transport and metabolism • Secondary metabolite biosynthesis, transport, and catabolism • Cell wall/membrane/envelope biogenesis
StHeBC4_p_c18261_g0_i1*	Bidirectional sugar transporter SWEET2	• Signal transduction mechanisms • Coenzyme transport and metabolism • Secondary metabolite biosynthesis, transport, and catabolism • Nucleotide transport and metabolism • Inorganic ion transport and metabolism
StHeBC4_p_c18413_g0_i2*	Serine acetyltransferase chloroplastic	• Signal transduction mechanisms • Intracellular trafficking, secretion, and vesicular transport • Energy production and conversion Posttranslational modification, protein turnover, and chaperones • Amino acid transport and metabolism
StHeBC4_p_c24289_g0_i1	Glycerate dehydrogenase peroxisomal, NADH-dependent hydroxypyruvate reductase 1	NA
StHeBC4_p_c25267_g0_i1	DAG chloroplastic	NA
StHeBC4_p_c26238_g1_i1*	dnaJ homologue subfamily B member, heat shock Hsp40-3	• Intracellular trafficking, secretion, and vesicular transport • Translation, ribosomal structure, and biogenesis • Energy production and conversion • Lipid transport and metabolism • Transcription • Posttranslational modification, protein turnover, and chaperones
StHeBC4_p_c26940_g5_i2	Proteasome subunit beta type-6, proteasome delta chain, tobacco cryptogein-induced 7	• Transcription • Posttranslational modification, protein turnover, and chaperones
StHeBC4_p_c27060_g3_i5	Oxygen-evolving enhancer 2- chloroplastic, 23-kDa subunit of oxygen evolving system of photosystem II, 23-kDa thylakoid membrane, OEC 23-kDa subunit	• Cell motility • Extracellular structures • Signal transduction mechanisms

[a] Transcripts with nonconservative amino acid changes are indicated in bold, and transcripts with gene functions that match the core parasitism gene categories (Yang et al. 2015) are noted with an asterisk (*).
[b] Gene description obtained from the best match (based on the e-value) in the Basic Local Alignment Search Tool (blastx).
[c] Functional gene categories obtained from the orthologue comparison in the OrthoDB.

https://doi.org/10.1017/wsc.2019.20

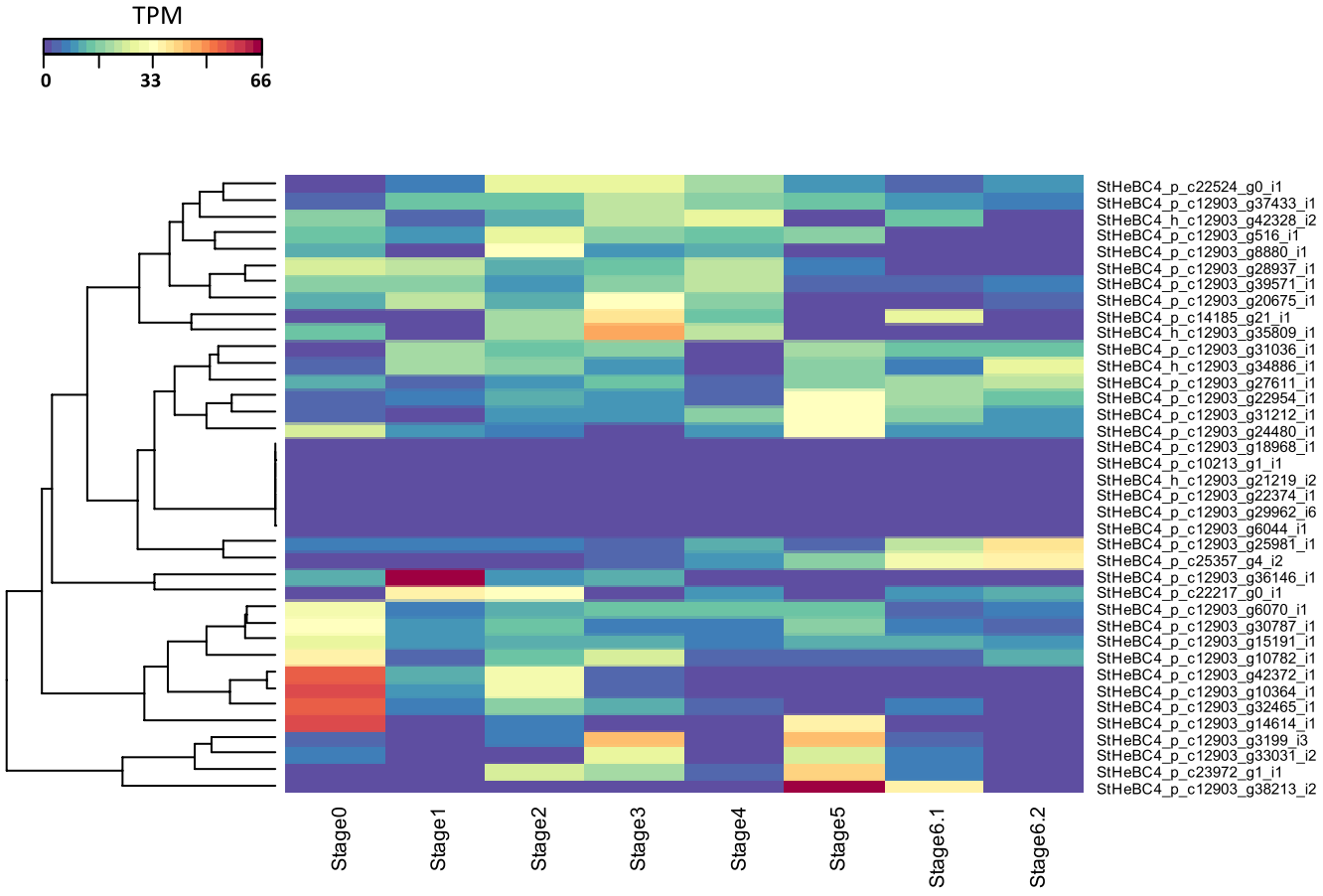

Figure 2. Heat map showing expression in transcripts per million (TPM) from tissue-specific data in different developmental stages of *Striga hermonthica*. Data were extracted from the Parasitic Plant Genome Project II (http://ppgp.huck.psu.edu/). Stage 0, expression in imbibed seeds; Stage 1, expression in roots of germinated seedlings after exposure to GR24; Stage 2, expression in roots of germinated seedlings after exposure to host roots; Stage 3, expression in haustoria attached to host roots (prevascular connection); Stage 4, expression in haustoria attached to host roots after vascular connection; Stage 5, late postattachment from belowground plants; Stage 6.1, expression in stem and leaf tissue; Stage 6.2, expression in flower tissue. Values correspond to the averaged relative expression from three replicates in each stage.

Results and discussion

Data preprocessing and de novo assembly

The total number of raw reads across the 14 libraries of *S. hermonthica* ranged from 2,837,888 in sample St_He_NG_12 to 7,749,848 in St_He_NG_2 (Supplementary Table 1) (raw sequencing data available at http://bigdata.bx.psu.edu/PPGP_II_data/StHeBC4 and in the Dryad repository https://doi.org/10.5061/dryad.s98q453). Likewise, the minimum and maximum number of high-quality clean reads were found in St_He_NG_12 and St_He_NG_2 (2,694,360 and 7,381,916, respectively) (Supplementary Table 1). Quality filtering of individual reads retained an average of 93.11% of read pairs.

After assembly of the quality trimmed reads with Trinity (Grabherr et al. 2011), a series of postprocessing steps were used to reduce the assembly to a single transcript per putative locus (see Yang et al. 2015; Supplementary Table 2). Out of 94,969 postprocessed *S. hermonthica* de novo assembly transcripts, we selected 59,243 putative locus representatives to use as the reference transcriptome for subsequent SNP discovery and expression analyses.

Host-associated Allelic differentiation analysis: low genome-wide divergence in expressed allelic frequency between hosts

Before SNP calling, we mapped clean reads to the reference transcriptome with an average mapping rate of 83.67% (ranging from 79.66% to 90.19%). The initial SNP calling contained 401,544 SNPs, but after filtering for biallelic SNPs and minimum coverage of 5X, we retained 4,806 SNPs distributed across 899 transcripts. A low SNP retention rate primarily resulted from coverage of transcripts expressed at low levels. The set of filtered SNPs was used to investigate how the individual samples clustered based on their genotype using a PCA. The first six eigenvectors captured 51.4% of the data variation, with each eigenvector explaining a similar amount of variation (9.7%, 8.8%, 8.4%, 8.3%, 8.1%, and 7.9%). The PCA clustering did not mirror geographic distribution of *S. hermonthica* samples, and clustering with respect to host was not observed (Supplementary Figure 2). After removing nonpolymorphic SNPs and SNPs with missing values, we calculated pairwise F_{ST} values for each SNP. We focused on the top 1% of sites with highest differentiation between groups from different host species (Supplementary Figure 3), which yielded 56 SNPs distributed in 50 transcripts (Table 1). For the 56 SNPs, difference in reference allelic frequency between populations from maize versus grain sorghum ranged from 12.5% to 45.4%.

The low F_{ST} between hosts averaged across SNPs found in expressed transcripts (0.012) indicated minimal genome-wide divergence across the study region (~180 km^2). Insect-pollinated species like *S. hermonthica* (Safa et al. 1984) can display high gene flow among populations (Ahmed et al. 2009). This and the high spatial heterogeneity in available hosts (adjacent farm fields with different crops) are likely to promote extensive gene flow between

https://doi.org/10.1017/wsc.2019.20

S. hermonthica parasites on different hosts. Seed distribution from neighboring fields via wind and water and distribution from near and more distant locations via contaminated crop seed, animals, equipment, and farmers would also contribute to lower population differentiation. Moreover, in the study region, farmers frequently rotate crop species between growing seasons, facilitating gene flow of *S. hermonthica* across hosts. Eight of 14 populations in this study were growing with another crop species (intercrop) when sampled (Supplementary Table 1). Of these eight, two populations were from fields where maize and grain sorghum were intercropped, and only one population was from a field planted to the same crop (grain sorghum) for the 3 yr before sampling. Altogether, we anticipated low genome-wide divergence between *S. hermonthica* populations from different hosts. Although F_{ST} values calculated for individual SNPs between host groups were generally low, we also found SNPs with moderate to high differentiation (i.e., exceeding ~0.2). We cannot rule out the role of neutral processes underlying high levels of differentiation among these transcripts (Supplementary Figure 4). However, high host-associated allelic differentiation could indicate host-specific selection over generations, causing strong differentiation at loci important for host specificity. Because parasites initiate the infection process belowground, it could be expected that selection on host-specific responses would be strongest during initial stages of parasite germination, development, and attachment (Conn et al. 2015). Additionally, allelic differentiation could result if seeds of *S. hermonthica* genotypes germinate in response to certain host genotypes. Following attachment, different physiological/metabolic environments presented by different hosts may also impose selection on loci important for physiological integration at the vegetative stage.

To understand the role that the transcripts containing F_{ST} outlier loci had in *S. hermonthica*, we predicted the impact that the change of allele had on the coding sequence and annotated the coding sequences. Six out of the 56 outlier SNPs were located outside a predicted CDS (Table 1). For the 50 SNPs located within a CDS, we identified 32 synonymous and 18 nonsynonymous substitutions (Table 1). Among the nonsynonymous substitutions, 9 were conservative and 9 were nonconservative changes. While all nonsynonymous mutations lead to an amino acid change in the resulting protein, the properties of the amino acid remain unchanged in conservative substitutions, while the properties of the resulting amino acid differ in nonconservative substitutions, potentially resulting in protein functional changes. We retrieved annotation for 45 out of the 50 transcripts containing highly differentiated SNPs (Table 2; Supplementary Table 3). To identify relevant genes for the host–parasite interaction, we focused on the functional gene categories and the GO terms extracted from the annotation. The three most common GO terms (level 3) identified during the annotation process for the category biological processes were cellular metabolic process, organic substance metabolic process, and nitrogen compound metabolic process. In the category molecular function, we found the terms protein binding, ion binding, and organic cyclic compound binding. In the cellular component category, the three most common GO terms were intracellular part, intracellular organelle, and membrane-bounded organelle. The functional annotations showed two core parasitic gene categories previously reported as enriched among core parasitism genes and formation of haustoria (Yang et al. 2015): transporters and cell wall biogenesis (Supplementary Table 3). The most abundant category was transporters, with five transcripts containing nonconservative substitutions (Table 2). Finally, the enzyme code search indicated that the most common categories were oxidoreductases and lyases (Supplementary Table 3). When we checked pathway annotations for transcripts from the F_{ST} analyses, we found 34 pathways that included 22 enzymes, the most common ones were purine and thiamine metabolism (Supplementary Table 4).

Genes related to nutrient transport, metabolism, and defense may be among those most important for regulating physiological integration at the vegetative stage. Among the top 1% of sites most genetically differentiated between groups were genes that reside in the functional category nutrient transport (Table 2). Two candidate transcripts with nonconservative changes encoded a predicted serine acetyltransferase and a dnaJ homologue subfamily B protein. Besides their role in nutrient transport, these transcripts are also associated with energy production and conversion. An *ABC* transporter gene associated with ABA-activated signaling and active transport was detected by the expression analysis. Parasitic Orobanchaceae have been shown to recruit plant regulatory machinery associated with multiple phytohormones to perform parasite-specific functions (Tomilov et al. 2005). Moreover, in a comparative transcriptomic study of three Orobanchaceae parasites, an ABC transporter was upregulated (Yang et al. 2015). Finally, ABA has been shown to inhibit the germination of the obligate parasite branched broomrape [*Phelipanche ramosa* (L.) Pomel] (Lechat et al. 2012; Pouvreau et al. 2013). A second ABA-associated transcript containing a conservative substitution is a homologue of a primary amine oxidase with quinone-binding activity. In other root parasitic plants, a quinone oxidoreductase was found to be necessary for haustorial initiation (Bandaranayake et al. 2010; Ngo et al. 2013). Another transcript containing a nonconservative substitution was a WRKY transcription factor associated with defense responses. This transcript encodes a promising candidate for host-specific interactions, because a member of this family, WRKY45, has been found to be associated with resistance to *S. hermonthica* infection in rice (Mutuku et al. 2015).

We also identified a set of candidate transcripts with synonymous mutations between plants parasitizing maize and grain sorghum (Table 1). Traditionally, synonymous mutations were thought to have no consequence for phenotype or fitness. However, there is growing evidence that synonymous mutations can affect the phenotype and fitness by altering mRNA stability and transcript splicing (Chamary and Hurst 2005; Kudla et al. 2009) and therefore can be subject to selection (Bailey et al. 2014; Hunt et al. 2014; Ingvarsson 2010). One transcript of potential interest was an ABA-associated, synaptosomal SNAP25-like homologous to SNAP33. In *Arabidopsis thaliana*, loss of function in SNAP25-type proteins leads to large necrotic lesions on leaves, resulting in a lethal dwarf phenotype (Heese et al. 2001). Moreover, SNAP25-like proteins are associated with defense responses, and the expression of SNAP33 increases response to pathogenic infection and mechanical stimulation (Wick et al. 2003). We also found two transcripts with synonymous mutations that were associated with response to salicylic acid (SA): salicylate 3-hydroxylase DMR6-like and a cell wall kinase receptor. When attacked by a pathogen, plants require SA for the initiation of a hypersensitive response and to trigger systemic acquired resistance (Durrant and Dong 2004; Fu and Dong 2013). Moreover, nonparasitic plants pathogens have been show to modify SA concentrations in the host for their own benefit (Cui et al. 2005; Zheng et al. 2012b). The SA-related cell wall kinase may also be an interesting candidate, because cell wall–associated receptor kinases are transmembrane proteins with an extracellular domain that binds cell wall–associated pectins. Members of this family are known

https://doi.org/10.1017/wsc.2019.20

to be involved in disease resistance, hormone signaling, legume–rhizobium symbiosis, senescence, abiotic stress, and wounding stress responses (Bethke et al. 2016; Nishiguchi et al. 2002).

Host-specific, differentially expressed transcripts in aboveground tissue—transporters, cell wall modification, and defense pathogenesis–related and oxidation–reduction processes

After we removed weakly expressed transcripts and normalized read counts, the count matrix retained 45,733 of the 59,243 transcripts in the de novo reference assembly. Across all 45,733 transcripts, host species explained 8.43% of the total variation in expression, although permutations indicated this was nonsignificant ($P = 0.379$, 999 permutations). This absence of significance for the whole set of transcripts is not surprising, because most genes were not host-specific in expression. However, the percentage of variation explained by host for each individual transcript ranged from 0% to 97.09%, indicating that there are individual transcripts with strong host-specific expression.

To reduce the influence of outlier samples within groups of small size (e.g., $n = 3$ for maize), we selected a subset of differentially expressed transcripts that were highly associated with host species (P-value < 0.01, log fold change > 2, variation explained by host of at least 60%). Using these criteria, we identified 38 transcripts, of which 33 were downregulated in grain sorghum and 5 upregulated in grain sorghum compared with maize (Figure 1). The P-values obtained in the DE analysis ranged from 5.68×10^{-9} to 0.0083, and the percentage of the variation explained by host ranged from 97.09 to 60.09 (Supplementary Table 5). Samples clustered based on host, indicating that host is the main factor associated with expression of these transcripts (Figure 1).

To further investigate the levels of these differentially expressed transcripts in other parts of the plant, we obtained expression data (transcripts per million [TPM]) for the following developmental stages in *S. hermonthica* from the PPGP II: imbibed seeds (stage 0), roots of germinated seedlings after exposure to GR24 (stage 1), roots of germinated seedlings after exposure to host roots (stage 2), expression in haustoria attached to host roots (pre-vascular connection) (stage 3), haustoria attached to host roots after vascular connection (stage 4), late postattachment from below-ground plants (stage 5), stem and leaf tissue (stage 6.1), and flower tissue (stage 6.2). The transcripts from all of these stages were isolated from parasites growing on grain sorghum as the host. Six of the differentially expressed transcripts were absent in the libraries from the tissue-specific experiments (Figure 2). Five of those six transcripts were found to be upregulated in the maize samples but downregulated in the grain sorghum samples (Figure 1). Moreover, the tissue-specific expression data (from grain sorghum hosts) contained two transcripts highly expressed in all stages except in the aboveground tissues, and those were also found to be upregulated in the maize samples and downregulated in grain sorghum.

We annotated the 38 focal transcripts, finding 34 of them had a significant hit in the blastx search (Table 3; Supplementary Table 5). Once again, we focused on the functional gene categories and the GO terms extracted from the annotation to find important transcripts for the host–parasite interaction. One-third of the focal transcripts matched gene categories reported as enriched among core parasitism genes and formation of haustoria (Yang et al. 2015). The most common group in the data was transporters, followed by cell wall modification (Table 3). Most of the transporters belonged to carbohydrate and amino acid transporters and metabolic functional gene categories. The three most common GO terms (level 3) identified during the annotation process for the category biological processes were all related to metabolism: cellular metabolic process, organic substance metabolic process, and primary metabolic process (Supplementary Table 5). In the category molecular function, we found protein binding, organic cyclic compound binding, and heterocyclic compound binding (Supplementary Table 5). Finally, for the category cellular component, the three most common GO terms were intracellular part, intracellular organelle, and membrane-bounded organelle (Supplementary Table 5). The most common enzyme categories were hydrolases, followed by oxidoreductases, transferases, and lyases (Supplementary Table 5). Because host-specific responses may be linked to important physiological mechanisms, we identified pathways associated with the focal transcripts. We found 26 pathways that included nine enzymes (Supplementary Table 6). The most common pathways were thiamine metabolism and purine metabolism, consistent with their annotation in the amino acid and metabolic functional gene category. For those transcripts with no annotation, we searched the Rfam database but found no match. Finally, although minimum coverage criteria for the F_{ST} analysis precluded detection of differentially expressed transcripts among the set of high F_{ST} transcripts, the most common pathways detected for the transcripts containing the top 1% of highly differentiated SNPs were consistent with the ones found in the DE transcripts.

If the high gene flow in *S. hermonthica* detected by our allelic frequency analyses limits adaptation to specific hosts, plastic host-specific gene expression may be key to successful parasitism. So far, expression studies looking for host-specific interactions with root parasites have focused on the haustorium and host–root interface, overlooking parasite aboveground responses to host (Honaas et al. 2013; Yang et al. 2015). Here we focused on expression of aboveground tissue in an attempt to identify a group of focal transcripts associated with host-specific, aboveground processes in this obligate hemiparasite. Of the total variation in expression, 9% was attributable to host. High genetic diversity among populations, even at limited spatial scales (Bozkurt et al. 2015; Unachukwu et al. 2017; Welsh and Mohamed 2011), together with temporal and environmental variation impossible to control in a natural setting, are both likely contribute to this variation. Still, we identified a set of transcripts exhibiting strong associations with host, supporting differential expression of a subset of genes involved in host-specific interactions and suggesting that aboveground organs also play a role in the parasitic process (Table 3).

Among differentially expressed focal transcripts, we detected several previously reported and relevant gene categories for host–parasite interactions (core parasitic genes) (Yang et al. 2015). Common functional categories detected in the analysis were genes encoding transporters, cell wall modification, oxidation–reduction process, and defense mechanisms (Table 3; Supplementary Table 5). The most abundant functional categories were amino acid and carbohydrate transport and metabolism (Table 3). Nutrient transport was also the main predicted functional gene category detected in transcripts containing F_{ST} outliers, suggesting that differences in nutrient transport in aboveground tissues may be a key characteristic of parasitism on distinct hosts. In the stem parasite field dodder (*Cuscuta pentagona* Engelm.), transcriptomic analyses revealed that after the establishment of vascular connections, amino acid and sugar transporters are upregulated (Ranjan et al. 2014). Differences in expression in these

https://doi.org/10.1017/wsc.2019.20

Table 3. Transcripts of *Striga hermonthica* detected as differentially expressed with P < 0.01 and log fold change (logFC) > 2 and variation explained by host > 60%

Transcript[a]	Upregulated host[b]	Description[c]	Functional gene category[d]	Tissue-specific expression[e]
StHeBC4_p_c12903_g31212_i1	Maize	Isoflavone reductase, phenylcoumaran benzylic ether reductase-like protein Fi1	NA	Late postattachment belowground tissue
StHeBC4_p_c12903_g20675_i1*	Maize	Mevalonate diphosphate decarboxylase	• Coenzyme transport and metabolism • Cell wall/membrane/envelope biogenesis • Carbohydrate transport and metabolism • Amino acid transport and metabolism	Roots prevascular connection
StHeBC4_p_c12903_g22374_i1	Maize	DUF724 domain-containing protein	NA	Absent
StHeBC4_p_c22217_g0_i1	Maize	Multidrug and toxic compound extrusion 14	• Signal transduction mechanisms • Transcription • Translation, ribosomal structure, and biogenesis	Germinated roots exposed to GR24 and host
StHeBC4_p_c12903_g6044_i1	Maize	Auxin signaling F-Box	• Defense mechanisms	Absent
StHeBC4_p_c12903_g37433_i1	Maize	Probable E3 ubiquitin ligase SUD1, suppressor of DRY2 DEFFECTS 1, RING-type E3 ubiquitin transferase SUD1, RING U-box domain-containing	NA	Haustorial prevascular connection
StHeBC4_p_c12903_g32465_i1*	Maize	NA	• Transcription • Intracellular trafficking, secretion, and vesicular transport, • Nucleotide transport and metabolism • Replication, recombination, and repair • Carbohydrate transport and metabolism	Imbibed seeds
StHeBC4_p_c12903_g14614_i1*	Maize	Serine threonine- kinase PBL10	• Carbohydrate transport and metabolism • Energy production and conversion • Amino acid transport and metabolism	Imbibed seeds
StHeBC4_p_c12903_g516_i1	Maize	*S*-adenosylmethionine decarboxylase proenzyme	NA	Germinated roots exposed to host
StHeBC4_p_c10213_g1_i1*	Maize	Pectinesterase inhibitor 58	NA	Absent
StHeBC4_p_c12903_g15191_i1	Maize	Peptidyl-prolyl *cis-trans* isomerase FKBP16-chloroplastic isoform, transcript variant	NA	Imbibed seeds
StHeBC4_h_c12903_g34886_i1*	Maize	Hexose transporter 14	NA	Flowers
StHeBC4_p_c14185_g21_i1*	Maize	Hypothetical protein	• Lipid transport and metabolism • Amino acid transport and metabolism	Haustorial prevascular connection
StHeBC4_p_c12903_g38213_i2	Maize	Transcription factor bHLH148-like	NA	Late postattachment belowground tissue
StHeBC4_p_c12903_g6070_i1*	Maize	Anion exchanger adaptor protein, Kanadaptin, contains FHA domain	• Carbohydrate transport and metabolism • Mobilome: prophages and transposons • Amino acid transport and metabolism	Imbibed seeds
StHeBC4_p_c12903_g39571_i1	Maize	ATP-dependent DNA helicase DDX11 isoform X2	• Defense mechanisms • Transcription • Replication, recombination, and repair	Haustoria postvascular connection
StHeBC4_p_c12903_g8880_i1	Maize	Uncharacterized protein LOC104229183	NA	Germinated roots exposed to host
StHeBC4_p_c12903_g30787_i1	Maize	Kinesin-related 11	• Signal transduction mechanisms	Imbibed seeds

(*Continued*)

Table 3. (*Continued*)

Transcript[a]	Upregulated host[b]	Description[c]	Functional gene category[d]	Tissue-specific expression[e]
StHeBC4_p_c12903_g24480_i1	Maize	Ureide permease 1 isoform X2	• Carbohydrate transport and metabolism	Late postattachment belowground tissue
StHeBC4_p_c12903_g28937_i1	Maize	60S acidic ribosomal P1-like	NA	Imbibed seeds, germinated roots exposed to GR24, and haustorial postvascular connection
StHeBC4_h_c12903_g21219_i2	Maize	60S ribosomal L12-like	• Translation, ribosomal structure, and biogenesis	Absent
StHeBC4_p_c12903_g29962_i6	Maize	Uncharacterized protein LOC105169457	NA	Absent
StHeBC4_p_c12903_g27611_i1	Maize	dr1-associated corepressor-like isoform X1	• Chromatin structure and dynamics	Flowers
StHeBC4_p_c12903_g42372_i1	Maize	Late embryogenesis abundant D-29-like	NA	Imbibed seeds
StHeBC4_p_c12903_g36146_i1	Maize	Aldehyde dehydrogenase	NA	Germinated roots exposed to GR24
StHeBC4_p_c12903_g3199_i3	Maize	Plant intracellular Ras-group-related LRR 6	NA	Haustorial prevascular connection and late postattachment belowground tissue
StHeBC4_p_c12903_g33031_i2	Maize	NA	NA	Haustorial prevascular connection and late postattachment belowground tissue
StHeBC4_p_c23972_g1_i1*	Maize	Organic cation carnitine transporter 4	• Replication, recombination, and repair	Late postattachment belowground tissue
StHeBC4_p_c12903_g25981_i1*	Maize	NA	• Cell wall/membrane/envelope biogenesis • Cell motility • Extracellular structures • Translation, ribosomal structure, and biogenesis • Cell cycle control, cell division, and chromosome partitioning	Flowers
StHeBC4_p_c22524_g0_i1	Maize	Non-LTR retroelement reverse transcriptase	NA	Germinated roots exposed to hosts and haustorial prevascular connection
StHeBC4_p_c12903_g10364_i1	Maize	Late embryogenesis abundant domain-containing LEA domain-containing	NA	Imbibed seeds
StHeBC4_h_c12903_g42328_i2	Maize	T-complex 1 subunit	• Posttranslational modification, protein turnover, and chaperones	Haustorial prevascular and postvascular connection
StHeBC4_p_c12903_g22954_i1	Maize	Protein CHUP1, chloroplastic	NA	Late postattachment belowground tissue
StHeBC4_p_c25357_g4_i2*	Sorghum	Caffeic acid 3-*O*-methyltransferase	• Transcription • Cell wall/membrane/envelope biogenesis • Carbohydrate transport and metabolism	Stems, leaves, and flowers
StHeBC4_p_c12903_g31036_i1	Sorghum	WVD2-like	NA	Germinated roots exposed to GR24
StHeBC4_p_c12903_g10782_i1	Sorghum	Protein TIC chloroplastic	NA	Imbibed seeds
StHeBC4_h_c12903_g35809_i1	Sorghum	NA	• Carbohydrate transport and metabolism	Haustorial prevascular connection
StHeBC4_p_c12903_g18968_i1*	Sorghum	Organic cation carnitine transporter 7	NA	Absent

[a] Transcripts not found in the tissue-specific data are highlighted in bold, and those belonging to core parasitism gene sets (Yang et al. 2015) are denoted with an asterisk (*).
[b] Host in which the transcript was identified as upregulated.
[c] Gene description obtained from the best match (based on the e-value) in the Basic Local Alignment Search Tool (blastx).
[d] Functional gene categories obtained from the orthologue comparison in the OrthoDB.
[e] Developmental stages where this transcript was found to have the highest expression in the data from the Parasitic Plant Genome Project (PPGP II).

transcripts between hosts suggests that as the parasitic process progresses, the host–parasite interaction relies increasingly on the transfer of nutrients and solutes from host to parasite. In agreement, Yang et al. (2015) found transcripts for several amino acid and sugar transporters consistently upregulated across several Orobanchaceae parasites. Moreover, *S. hermonthica* infection is associated with increased amino acid levels in the host xylem sap (Pageau et al. 2003). One strongly differentially expressed transcript between hosts (upregulated in sorghum) encoded an organic cation carnitine transmembrane transporter. Carnitine is associated with fatty acid metabolism in plants (Bourdin et al. 2007), and it is present in many seeds, where it acts as a store of energy and has a critical role in postgermination growth (Lawand et al. 2002; Panter and Mudd 1969). Another pivotal gene category for parasitic functions is cell wall modification (Honaas et al. 2013; Ranjan et al. 2014; Yang et al. 2015). To establish a successful host–parasite connection, parasites have to loosen or degrade the host cell wall and establish connections with the host vascular tissue (Lawand et al. 2002; Saucet and Shirasu 2016). A cell wall modification gene of interest as a candidate for host specificity was a pectinesterase inhibitor. This transcript was strongly differentially expressed between hosts in the field data (upregulated in maize) and was absent in the tissue-specific expression data from laboratory-grown plants. Pectinases mediate the deposition of Casparian strips in the endodermis by recruiting the lignin polymerization machinery (Roppolo et al. 2014). Pectinesterase inhibitors have been linked to enhanced resistance to fungal infection in plants (Liu et al. 2018) and to symbiotic nodulation (Young et al. 2011; Zouari et al. 2014).

Another functional gene category shown to be upregulated in parasitic plants is defense response and pathogenesis (Honaas et al. 2013; Yang et al. 2015). *Striga* manipulates host immunity for its own benefit (Saucet and Shirasu 2016) and can alter ABA, gibberellic acid, and cytokinin concentrations during host infection (Frost et al. 1997). In addition, it has been suggested that *Striga* can produce a set of effectors that subdue host defense (Runo and Kuria 2018). In this study, we found several strongly differentially expressed transcripts associated with defense and pathogenesis. One of them was a gene encoding an auxin-signaling F-box that was upregulated in the maize samples but absent from the grain sorghum host tissue-specific expression data. Finally, another relevant category is oxidation–reduction process. The latter was found to be enriched in orthogroups, with greater dN/dS in parasitic lineages compared with nonparasitic lineages (Yang et al. 2015). The most strongly differentially expressed transcript between hosts belongs to this category, a phenylcoumaran benzylic ether reductase-like protein. This protein is associated with oxidoreductase activity and flavonoids and has been previously associated with witchweed [*Striga asiatica* (L.) Kuntze] haustorial formation (Liu 2012). Certain flavonoids act as haustorium-inducing factors (Albrecht et al. 1999; Riopel and Timko 1992), while oxidoreductase activity triggers haustorial formation (Bandaranayake et al. 2010). Finally, three transcripts had no hit in the blastx search, but we found a functional gene category in the orthologue search. This was also reported in Yang et al. (2015) and Honaas et al. (2013): several parasitic-specific transcripts had nonsignificant blast alignments or had a hit to genes encoding proteins of unknown function, suggesting that many parasite-specific transcripts have a yet to be discovered function, even when evidence is based on homology comparisons.

Most of the host-associated, differentially expressed transcripts were upregulated in maize compared with grain sorghum (33 vs. 5, respectively) (Table 3). The majority of the upregulated functional gene categories and pathways were associated with nutrient metabolism (i.e., sugars, amino acids) (Table 3; Supplementary Table 6). Because nutrient uptake from the host is essential for the parasite's growth and reproduction, differences in physiological integration at the vegetative stage can be inferred by studying expression patterns of nutrient transport and metabolic transcripts. These differences in expression are likely to be caused by selection pressures imposed by a different growing environment (i.e., host crop). We hypothesize that our results might be associated with host tolerance level to the parasite infection. Tolerance to *S. hermonthica* varies greatly within and among crops (Amusan et al. 2008; Haussmann et al. 2004; Hess et al. 1992; Lane et al. 1997; Maiti et al. 1984; Vogler et al. 1996). In this regard, tolerance in grain sorghum to infection with *S. hermonthica* has been reported in many cases, but tolerance in maize is rare (Amusan et al. 2008; Lane et al. 1997; Mutinda et al. 2018). This variation could be attributed to the evolutionary history of these crops. Grain sorghum is a native African crop that has presumably coevolved with *S. hermonthica* and other *Striga* species, while maize evolution took place in the absence of the parasite. The almost absent resistance in maize might benefit the host–parasite physiological integration, facilitating nutrient uptake from the host at the root level, with the nutrients then metabolized in the parasite's aboveground tissue. Further work is required to test this hypothesis. Finally, because our data come from aboveground tissue, finding haustorium activity–related transcripts could indicate a global pattern of host-specificity expression (i.e., across all tissues). Alternatively, aboveground organs could be a source of gene expression after the parasite's germination and establishment of the haustorial connection, contributing to the maintenance of the connection with the host. Overall, the DE analysis revealed a set of predicted functional gene categories with candidate transcripts relevant to host-specific interactions (Table 3). The importance of several of these candidate transcripts in host-specific interactions was further supported by tissue-specific expression data from the PPGP II. The tissue-specific data came from *S. hermonthica* plants parasitizing grain sorghum, and interestingly, almost all of these transcripts (five out of six) were upregulated in maize and downregulated in grain sorghum (Table 3). Nevertheless, we have to consider the possibility that the absence of these transcripts in the tissue-specific expression data might be explained by methodological causes (not captured in libraries or expressed at very low levels).

In summary, the agriculturally important parasitic weed *S. hermonthica* infects multiple host species, making it a valuable system to study the complex genetic mechanisms of parasite responses to diverse hosts. Parasite interactions with different hosts likely require and result in physiological changes throughout the parasitic plant, especially in obligate parasitic weeds such as *S. hermonthica*. In addition, as host recognition acts at several levels (i.e., germination, haustorial initiation, attachment and penetration, and physiological integration), the parasite response to the host is likely to vary throughout development. After attachment, obligate parasites become a metabolic sink for carbohydrates and nitrogen from the host (Musselman 1980). It therefore remains critical to understand variation in dynamics of the host–parasite interaction beyond the more well-studied initial stages of host detection, development, and attachment of the haustorium, and establishment of vascular connections. Host transcriptomic studies have previously provided insights into *Striga* genotype–host genotype/species interactions, but so far, no study has addressed how genetic polymorphism may contribute to

https://doi.org/10.1017/wsc.2019.20

plasticity and/or host-specific genotypes. Here, we leveraged transcriptomic data from aboveground tissue to identify a set of host-associated transcripts in natural populations of *S. hermonthica* and evaluate levels of genetic polymorphism in transcripts expressed among populations parasitizing different hosts, illuminating the role that expression variation and genetic polymorphism may play in mediating host–parasite dynamics during later stages of the interaction. Despite high gene flow among populations, we identified a set of host-associated transcripts using differential expression and population genomic analysis of expressed genes. We found several functional gene categories relevant to host–parasite interactions supported by previous studies in parasitic plants. More specifically, our results highlight the role of nutrient transporters and the probable importance of gene regulation in aboveground organs in the specificity of the parasitic process. Those parasite genotypes that are not removed by selection during germination and haustorial connection may be selected later by pressures that the host, as an environment, imposes on the parasite. In addition, we found several defense- and pathogenesis- related genes together with plant hormone-response genes, supporting the idea that pathways involved in plant defense/pathogenesis and hormone response can be used by the parasite during the parasitic process. Overall, we provide a set of candidate transcripts that show host-specific interaction in aboveground tissue in the parasitic plant *S. hermonthica*. This suggests that signals of host-specific processes can be detected aboveground, expanding the focus of host–parasite interactions beyond the haustorial connection.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/wsc.2019.20.

Author ORCID. Lua Lopez https://orcid.org/0000-0002-3619-763X.

Acknowledgments. No conflicts of interest have been declared. This work was supported by research funds from the International Institute of Tropical Agriculture (IITA) Nigeria to CWD and the National Science Foundation (NSF IOS-1238057) to CWD and MPT, and by start-up funds from Penn State's Eberly College of Science to JRL. Emily S. Bellis was supported by an NSF Postdoctoral Research Fellowship in Biology (grant number 1711950). The *S. hermonthica* field sampling and sequencing was funded under a USAID linkage grant from IITA. We thank Joshua P. Der for helpful ideas and comments on the project design. We thank Rafael Guerrero for his advice with the allelic frequency analysis. Likewise, we thank two anonymous reviewers for their comments on the article.

References

Ahmed S, Stephen GC, Roger KB, Philip MG (2009) Wind-borne insects mediate directional pollen transfer between desert fig trees 160 kilometers apart. Proc Natl Acad Sci USA 106:20342–20347

Albrecht H, Yoder JI, Phillips DA (1999) Flavonoids promote haustoria formation in the root parasite *Triphysaria versicolor*. Plant Physiol 119:585–592

Amusan OI, Patrick JR, Menkir A, Housley T, Ejeta G (2008) Resistance to *Striga hermonthica* in a maize inbred line derived from *Zea diploperennis*. New Phytol 178:157–166

Antonova TS, Terborg SJ (1996) The role of peroxidase in the resistance of sunflower against *Orobanche cumana* in Russia. Weed Res 36:113–121

Archie AE, Ezenwa OV (2011) Population genetic structure and history of a generalist parasite infecting multiple sympatric host species. Int J Parasitol 41:89–98

Ashburner M, Ball AC, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, *et al.* (2000) Gene ontology: tool for the unification of biology. Nat Genet 25:25

Atera E, Itoh K, Azuma T, Ishii T (2012) Farmers' perception and constraints to the adoption of weed control option: the case of *Striga asiatica* in Malawi. J Agric Sci 4:5

Bailey FS, Hinz A, Kassen R (2014) Adaptive synonymous mutations in an experimentally evolved *Pseudomonas fluorescens* population. Nat Commun 5:4076

Bandaranayake PCG, Filappova T, Tomilov A, Tomilova NB, Jamison-McClung D, Ngo Q, Inoue K, Yoder JI (2010) A single-electron reducing quinone oxidoreductase is necessary to induce haustorium development in the root parasitic plant *Triphysaria*. Plant Cell 22:1404–1419

Bashey F (2015) Within-host competitive interactions as a mechanism for the maintenance of parasite diversity. Philos Trans R Soc Lond B Biol Sci 370:20140301

Bethke G, Thao A, Xiong G, Li B, Soltis NE, Hatsugai N, Hillmer RA, Katagiri F, Kliebenstein DJ, Pauly M, Glazebrook J (2016) Pectin biosynthesis is critical for cell wall integrity and immunity in *Arabidopsis thaliana*. Plant Cell 28:537–556

Bolger MA, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 30:2114–2120

Bourdin B, Adenier H, Perrin Y (2007) Carnitine is associated with fatty acid metabolism in plants. Plant Physiol Biochem 45:926–931

Bozkurt ML, Muth P, Parzies HK, Haussmann BIG (2015) Genetic diversity of East and West African *Striga hermonthica* populations and virulence effects on a contrasting set of sorghum cultivars. Weed Res 55:71–81

Chamary JV, Hurst LD (2005) Evidence for selection on synonymous mutations affecting stability of mRNA secondary structure in mammals. Genome Biol 6:R75

Conesa A, Götz S (2008) Blast2GO: a comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics 2008:619832

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676

Conn CE, Bythell-Douglas R, Neumann D, Yoshida S, Whittington B, Westwood JH, Shirasu K, Bond CS, Dyer KA, Nelson DC (2015) Convergent evolution of strigolactone perception enabled host detection in parasitic plants. Science 349:540–543

Cui J, Bahrami AK, Pringle EG, Hernandez-Guzman G, Bender CL, Pierce NE, Ausubel FM (2005) *Pseudomonas syringae* manipulates systemic plant defenses against pathogens and herbivores. Proc Natl Acad Sci USA 102:1791–1796

Durrant WE, Dong X (2004) Systemic acquired resistance. Annu Rev Phytopathol 42:185–209.

Ejeta G (2007) Breeding for Striga resistance in sorghum: exploitation of an intricate host–parasite biology. Crop Sci 47:216–227

Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang H-Y, Dosztányi Z, El-Gebali S, Fraser M, Gough J, Haft D, Holliday GL, Huang H, Huang X, *et al.* (2017) InterPro in 2017–beyond protein family and domain annotations. Nucleic Acids Res 45:D190–D199

Frank SA (1993) Evolution of host-parasite diversity. Evolution 47:1721–1732

Frost DL, Gurney AL, Press MC, Scholes JD (1997) *Striga hermonthica* reduces photosynthesis in sorghum: the importance of stomatal limitations and a potential role for ABA? Plant Cell Environ 20:483–492

Fu ZQ, Dong X (2013) Systemic acquired resistance: turning local infection into global defense. Annu Rev Plant Biol 64:839–863

Geiger R (1954) Klassifikation der klimate nach W. Köppen. Pages 603–607 *in* Landolt HH, Börnstein R, eds. Zahlenwerte Und Funktionen Aus Physik, Chemie, Astronomie, Geophysik Und Technik, Alte Serie. Berlin: Springer

Gene Ontology Consortium (2017) Expansion of the Gene Ontology knowledgebase and resources. Nucleic Acids Res 45:331–338

Grabherr MG, Brian JH, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnol 29:644

Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR (2003) Rfam: an RNA family database. Nucleic Acids Res 31:439–41

Guo M, Tian F, Wamboldt Y, Alfano JR (2009) The majority of the type III effector inventory of *Pseudomonas syringae* Pv. Tomato DC3000 can suppress plant immunity. Mol Plant-Microbe Interact 22:1069–1080

Haussmann BIG, Hess DE, Omanya GO, Folkertsma RT, Reddy BVS, Kayentao M, Welz HG, Geiger HH (2004) Genomic regions influencing resistance to the parasitic weed *Striga hermonthica* in two recombinant inbred populations of sorghum. Theor Appl Genet 109:1005–1016

Heese M, Gansel X, Sticher L, Wick P, Grebe M, Granier F, Jürgens G (2001) Functional characterization of the KNOLLE-interacting t-SNARE AtSNAP33 and its role in plant cytokinesis. J Cell Biol 155:239–250

Hess DE, Ejeta G, Butler LG (1992) Selecting sorghum genotypes expressing a quantitative biosynthetic trait that confers resistance to *Striga*. Phytochemistry 31:493–497

Hoffman GE, Schadt EE (2016) VariancePartition: interpreting drivers of variation in complex gene expression studies. BMC Bioinformatics 17:483

Honaas LA, Wafula EK, Yang Z, Der JP, Wickett NJ, Altman NS, Taylor CG, Yoder JI, Timko MP, Westwood JH, dePamphilis CW (2013) Functional genomics of a generalist parasitic plant: laser microdissection of host–parasite interface reveals host-specific patterns of parasite gene expression. BMC Plant Biol 13:9

Hunt RC, Simhadri VL, Iandoli M, Sauna ZE, Kimchi-Sarfaty C (2014). Exposing synonymous mutations. Trends Genet 30:308–321

Ichihashi Y, Mutuku JM, Yoshida S, Shirasu K (2015) Transcriptomics exposes the uniqueness of parasitic plants. Brief Funct Genomics 14:275–282

Ingvarsson PK (2010) Natural selection on synonymous and nonsynonymous mutations shapes patterns of polymorphism in *Populus tremula*. Mol Biol Evol 27:650–660

Kaci-Chaouch T, Verneau O, Desdevises Y (2008) Host specificity is linked to intraspecific variability in the genus *Lamellodiscus* (Monogenea). Parasitology 135:607–616

Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 28:27–30

Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder PH, Kessler M (2017) Climatologies at high resolution for the earth's land surface areas. Sci Data 4:170122

Konczal M, Koteja P, Stuglik MT, Radwan J, Babik W (2013) Accuracy of allele frequency estimation using pooled RNA-Seq. Mol Ecol Res 14:381–392

Koyama ML (2000) Genetic variability of *Striga hermonthica* and effect of resistant cultivars on Striga population dynamics. Pages 247–260 *in* RattundeHFW, GeigerHH, HaussmanBIG, HessDE, KoyamaML, GrivetL, eds. Breeding for Striga Resistance in Cereals. Weikersheim, Germany: Margraf Verlag

Kudla G, Murray AW, Tollervey D, Plotkin JB (2009) Coding-sequence determinants of gene expression in *Escherichia coli*. Science 324:255–258

Lajeunesse MJ, Forbes MR (2002) Host range and local parasite adaptation. Proc R Soc Lond B Biol Sci 269:703–710

Lane JA, Child DV, Moore THM, Arnold GM, Bailey JA (1997) Phenotypic characterization of resistance in *Zea diploperennis* to *Striga*. Maydica 42:45–51

Lawand S, Dorne AJ, Long D, Coupland G, Mache R, Carol P (2002) Arabidopsis A BOUT DE SOUFFLE, which is homologous with mammalian carnitine acyl carrier, is required for postembryonic growth in the light. Plant Cell 14:2161–2173

Lechat MM, Pouvreau JB, Péron T, Gauthier M, Montiel G, Véronési C, Todoroki Y, Le Bizec B, Monteau F, Macherel D, Simier P, Thoiron S, Delavault P (2012) *PrCYP707A1*, an ABA catabolic gene, is a key component of Phelipanche ramosa seed germination in response to the strigolactone analogue GR24. J Exp Bot 63:5311–5322

Leggett HC, Benmayor R, Hodgson DJ, Buckling A (2013) Experimental evolution of adaptive phenotypic plasticity in a parasite. Curr Biol 23:139–142

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079

Liu N, Sun Y, Pei Y, Zhang X, Wang P, Li X, Li F, Hou Y (2018) A pectin methylesterase inhibitor enhances resistance to *Verticillium* Wilt. Plant Physiol 176:2202–2220

Liu Y (2012) Comparing and contrasting two plant pathogens provide a unique window into the differences in the innate immune responses between dicots and monocots. Ph.D dissertation. Atlanta, GA: Emory University. 173 p

Losner-Goshen D, Portnoy V, Mayer A, Joel D (1998) Pectolytic activity by the haustorium of the parasitic plant *Orobanche* L. (Orobanchaceae) in host roots. Ann Bot 81:319–326

Maiti RK, Ramaiah KV, Bisen SS, Chidley VL (1984) A comparative study of the haustorial development of *Striga asiatica* (L.) Kuntze on Sorghum cultivars. Ann Bot 54:447–457

McGill BJ, Enquist BJ, Weiher E, Westoby M (2006) Rebuilding community ecology from functional traits. Trends Ecol Evol 21:178–185

Musselman LJ (1980) The biology of *Striga, Orobanche*, and other root-parasitic weeds. Annu Rev Phytopathol 18:463–489

Mutinda SM, Masanga J, Mutuku JM, Runo S, Alakonya A (2018) KSTP 94, an open-pollinated maize variety has postattachment resistance to purple witchweed (*Striga hermonthica*). Weed Sci 66:525–529

Mutuku JM, Yoshida S, Shimizu T, Ichihashi Y, Wakatake T, Takahashi A, Seo M, Shirasu K (2015) The WRKY45-dependent signaling pathway is required for resistance against *Striga hermonthica* parasitism. Plant Physiol 168:1152–1163

Nei M (1977) F-statistics and analysis of gene diversity in subdivided populations. Ann Hum Genet 41:225–233

Ngo QA, Albrecht H, Tsuchimatsu T, Grossniklaus U (2013) The differentially regulated genes TvQR1 and *TvPirin* of the parasitic plant *Triphysaria* exhibit distinctive natural allelic diversity. BMC Plant Biol 13:28

Nishiguchi M, Yoshida K, Sumizono T, Tazaki K (2002) A receptor-like protein kinase with a lectin-like domain from Lombardy poplar: gene expression in response to wounding and characterization of phosphorylation activity. Mol Genet Genomics 267:506–514

Oksanen J, Blanchet GF, Kindt R, Legendre P, Minchin P, O'Hara RB, Simpson G, Solymos P, Stevens MHH, Wagner HH (2013) Vegan: Community ecology package. R Package v. 2.0–10.

Olivier A, Glaszmann JC, Lanaud C, Leroux GD (1998) Population structure, genetic diversity and host specificity of the parasitic weed *Striga hermonthica* (Scrophulariaceae) in Sahel. Plant Syst Evol. 209:33–45

Pageau K, Simier P, Le Bizec B, Robins RJ, Fer A (2003) Characterization of nitrogen relationships between *Sorghum bicolor* and the root–hemiparasitic angiosperm *Striga hermonthica* (Del.) Benth. using K15NO3 as isotopic tracer. J Exp Bot 54:789–99

Panter RA, Mudd JB (1969) Carnitine levels in some higher plants. FEBS Lett 5:169–170

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C (2017) Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14:417–419

Pérez-de-Luque A (2013) Haustorium invasion into host tissues. Pages 75–86 *in* JoelDM, GresselJ, MusselmanLJ, eds. Parasitic Orobanchaceae: Parasitic Mechanisms and Control Strategies. Berlin: Springer

Pouvreau JB, Gaudin Z, Auger B, Lechat MM, Gauthier M, Delavault P, Simier P (2013) A high-throughput seed germination assay for root parasitic plants. Plant Methods 9:32

Pruitt KD, Tatusova T, Maglott DR (2007) NCBI Reference Sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 35:61–65

Ranjan A, Ichihashi Y, Farhi M, Zumstein K, Townsley B, David-Schwartz R, Sinha NR (2014) De novo assembly and characterization of the transcriptome of the parasitic weed dodder identifies genes associated with plant parasitism. Plant Physiol 166:1186–1199

Riopel Jl, Timko MP (1992) Signals and regulation in the development of Striga and other parasitic angiosperms. Pages 493–507 *in* Verma DPS, ed. Molecular Signals in Plant-Microbe Communications. Boca Raton, FL: CRC Press

Robinson MD, McCarthy DJ, Smyth GK (2010) EdgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140

Rodenburg J, Demont M, Zwart SJ, Bastiaans L (2016) Parasitic weed incidence and related economic losses in rice in Africa. Agric Ecosyst Environ 235:306–317

Roppolo D, Boeckmann B, Pfister A, Boutet E, Rubio MC, Dénervaud-Tendon V, Vermeer JE, Gheyselinck J, Xenarios I, Geldner N (2014) Functional and

https://doi.org/10.1017/wsc.2019.20

evolutionary analysis of the CASPARIAN STRIP MEMBRANE DOMAIN PROTEIN family. Plant Physiol 165:1709–1722
Runo S, Kuria EK (2018) Habits of a highly successful cereal killer, Striga. PLoS Pathog 14:e1006731
Safa SB, Jones BMG, Musselman LJ (1984) Mechanisms favouring outbreeding in *Striga hermonthica* (Scrofulariaceae). New Phytol 96:299–305
Saucet SB, Shirasu K (2016) Molecular parasitic plant–host interactions. PLoS Pathog 12:e1005978
Singh A, Singh M (1993) Cell wall degrading enzymes in *Orobanche aegyptiaca* and its host *Brassica campestris*. Physiol Plant 89:177–181
Spallek T, Mutuku M, Shirasu K (2013) The genus *Striga*: a witch profile. Mol Plant Pathol 14:861–869
Swarbrick PJ, Huang K, Liu G, Slate J, Press MC, Scholes JD (2008) Global patterns of gene expression in rice cultivars undergoing a susceptible or resistant interaction with the parasitic plant *Striga hermonthica*. New Phytol 179:515–529
Swenson NG, Iida Y, Howe R, Wolf A, Umaña MN, Petprakob K, Turner BL, Ma K (2017) Tree co-occurrence and transcriptomic response to drought. Nat Commun 8:1996
Tomilov AA, Tomilova NB, Abdallah I, Yoder JI (2005) Localized hormone fluxes and early haustorium development in the hemiparasitic plant *Triphysaria versicolor*. Plant Physiol 138:1469–1480
Unachukwu NN, Menkir A, Rabbi IY, Oluoch M, Muranaka S, Elzein A, Odhiambo G, Farombi EO, Gedil M (2017) Genetic diversity and population structure of *Striga hermonthica* populations from Kenya and Nigeria. Weed Res 57:293–302
Vogler RK, Ejeta G, Butler LG (1996) Inheritance of low production of *Striga* germination stimulant in sorghum. Crop Sci 36:1185–1191
Watson IA (1970) Changes in virulence and population shifts in plant pathogens. Annu Rev Phytopathol 8:209–230
Welsh B, Mohamed KI (2011) Genetic diversity of *Striga hermonthica* populations in Ethiopia: evaluating the role of geography and host specificity in shaping population structure. Int J Plant Sci 172:773–782
Westwood JH, dePamphilis CW, Das M, Fernández-Aparicio M, Honaas LA, Timko MP, Wafula EK, Wickett NJ, Yoder JI (2012) The Parasitic Plant Genome Project: new tools for understanding the biology of *Orobanche* and *Striga*. Weed Sci 60:295–306
Wick P, Gansel X, Oulevey C, Page V, Studer I, Dürst M, Sticher L (2003) The expression of the T-SNARE AtSNAP33 is induced by pathogens and mechanical stimulation. Plant Physiol 132:343–351
Yang Z, Wafula EK, Honaas LA, Zhang H, Das M, Fernandez-Aparicio M, Huang K, Bandaranayake PCG, Wu B, Der JP, Clarke CD, Ralph PE, Landherr L, Altman NS, Timko MP, *et al.* (2015). Comparative transcriptome analyses reveal core parasitism genes and suggest gene duplication and repurposing as sources of structural novelty. Mol Bio Evol 32:767–790
Young ND, Debellé F, Oldroyd GED, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KF, Gouzy J, Schoof H, Van de Peer Y, Proost S, Cook DR, Meyers BC, Spannagl M, *et al.* (2011) The *Medicago* genome provides insight into the evolution of rhizobial symbioses. Nature 480:520
Zdobnov EM, Tegenfeldt F, Kuznetsov D, Waterhouse RM, Simão FA, Ioannidis P, Seppey M, Loetscher A, Kriventseva EV (2017) OrthoDB v9.1: cataloging evolutionary and functional annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs. Nucleic Acids Res 45:744–749
Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS (2012a) A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics 28:3326–3328
Zheng XY, Spivey NW, Zeng W, Liu PP, Fu ZQ, Klessig DF, He SY, Dong X (2012b) Coronatine promotes *Pseudomonas syringae* virulence in plants by activating a signaling cascade that inhibits salicylic acid accumulation. Cell Host Microbe 11:587–596
Zouari I, Salvioli A, Chialva M, Novero M, Miozzi L, Tenore GG, Bagnaresi P, Bonfante P (2014) From root to fruit: RNA-Seq analysis shows that arbuscular mycorrhizal symbiosis may affect tomato fruit metabolism. BMC Genomics 15:221

https://doi.org/10.1017/wsc.2019.20