Received: 19 December 2022 | Revised: 26 February 2023 | Accepted: 3 March 2023

DOI: 10.1002/ppp3.10370

RESEARCH ARTICLE

Plants People Planet PPP Open Access

The seed transcriptome of *Rafflesia* reveals horizontal gene transfer and convergent evolution: Implications for conserving the world's largest flower

Jeanmaire Molina[1,2] | Adhityo Wicaksono[3,4] | Todd P. Michael[5,6] | Su-Hwan Kwak[2] | Ronniel D. Pedales[7] | Zoé Joly-Lopez[8] | Semar Petrus[6] | Allen Mamerto[5] | Brian Tomek[2] | Sumaya Ahmed[2] | Venkatasivasankar Maddu[2] | Kristina Yakubova[2] | Danilo Tandang[9] | Joseph W. Morin[2] | So-Yon Park[10] | Hyun-Oh Lee[10] | William McLaughlin[11] | Kyle Wallick[11] | James Adams[11] | Ari Novy[5,12] | Susan Pell[11] | Michael D. Purugganan[13]

[1]Biology Department, Pace University, New York, New York, USA

[2]Department of Life Sciences, Long Island University, Brooklyn, New York, USA

[3]Division of Biotechnology, Genbinesia Foundation, Gresik, Indonesia

[4]Study Group of Bioinformatics, Synthetic Biology Indonesia, Sumbawa, Indonesia

[5]The Plant Molecular and Cellular Biology Laboratory, The Salk Institute for Biological Studies, La Jolla, California, USA

[6]Cepheid, Sunnyvale, California, USA

[7]Institute of Biology, University of the Philippines Diliman, Quezon City, Philippines

[8]Département de Chimie, Université du Québec à Montréal, Montréal, Québec, Canada

[9]Botany and National Herbarium Division, National Museum of the Philippines, Manila, Philippines

[10]Division of Plant Science and Biotechnology, University of Missouri, Columbia, Missouri, USA

[11]U.S. Botanic Garden, Washington, District of Columbia, USA

[12]San Diego Botanic Garden, Encinitas, California, USA

[13]Center for Genomics & Systems Biology, New York University, New York, USA

Correspondence
Jeanmaire Molina, Biology Department, Pace University, New York, NY, USA.

Societal Impact Statement

Rafflesia is a genus of parasitic plants with the largest flowers in the world, unique to the threatened forest habitats of tropical Asia. Here, we report on genes that are active (the transcriptome) in *Rafflesia* seeds as part of a larger effort to understand *Rafflesia*. *Rafflesia* has never been grown successfully outside of its native range. Consequently, seed banking is not yet possible, precluding a critical management strategy for conservation. The study of *Rafflesia* seed biology is a critical step to improve its cultivation, which will educate the public about unique species and the importance of conserving their habitats.

Summary

- *Rafflesia* is of great interest as one of the only two plants known to have completely lost its chloroplast genome. *Rafflesia* is a holoparasite and an endophyte that lives inside the tissues of its host, a tropical grape vine (*Tetrastigma*), emerging only to bloom—with the largest flower of any plant. Here, we report the first *Rafflesia* seed transcriptome and compare it with those of other plants to deepen our understanding of its extraordinary life history.
- We assembled a transcriptome from RNA extracted from seeds of the Philippine endemic *Rafflesia speciosa* and compared this with those of other plants, including *Arabidopsis*, parasitic plants *Striga* and *Cuscuta*, and the mycoheterotrophic orchid *Anoectochilus*.
- Genetic and metabolic seed pathways in *Rafflesia* were generally similar to the other plant species. However, there were some notable exceptions. We found evidence of horizontal transfer of a gene potentially involved in circumventing host defenses. Moreover, we identified a possible convergence among parasitic plants because *Rafflesia*, *Striga*, and *Cuscuta* shared important similarities. We were

Email: jmolina2@pace.edu;
jeanmaire.molina@liu.edu

Funding information
National Science Foundation, Grant/Award Number: 2204938

unable to find evidence of genes involved in mycorrhizal symbiosis, suggesting that mycoheterotrophy is unlikely to play a role in *Rafflesia* parasitism.

- To date, ex situ propagation of *Rafflesia* by seed has been mostly unsuccessful. Our research is a bold step forward in understanding the fundamentals of *Rafflesia* seed biology that will inform the continued propagation and seed-banking efforts concerning this recalcitrant plant. We discuss our findings in the broader context of the conservation of a genus in peril.

KEYWORDS
endophyte, parasite, Rafflesiaceae, seed development, *Tetrastigma*

1 | INTRODUCTION

Since first described in the 18th century from Java, Indonesia, the enormous blooms (~1 m) of *Rafflesia* that smell like rotting meat have mesmerized scientists and lay naturalists alike. Unlike most plants, *Rafflesia* (Rafflesiaceae) is a parasite bereft of stems, roots, and leaves. It lives as an endophyte, growing within the tissues of its host vine, the genus *Tetrastigma* (Vitaceae), and only emerges from its host's exterior to flower (Davis et al., 2007). Genomic studies have demonstrated an association between *Rafflesia* and *Tetrastigma* predating the origin of *Tetrastigma* itself, suggesting other taxa in the Vitaceae may have even served as hosts in the past (Cai et al., 2021). An association dating back to the Cretaceous may have afforded *Rafflesia*, and the closely-related *Sapria*, to become the only plants we know of to completely discard their plastid genomes (Cai et al., 2021; Molina et al., 2014) while retaining their plastid compartments for amino acid and lipid synthesis (Ng et al., 2018). The loss of the plastid genome is so far unique in Rafflesiaceae, that a science news article even asked, "When is a plant no longer a plant? (Pennisi, 2014)." However, its massive flowers attest to its bona fide "plant status."

There are about 40+ *Rafflesia* spp. endemic to the dwindling forests of tropical Asia with several species considered critically endangered (Barcelona et al., 2009; Pelser et al., 2019). Dubbed panda of the plant world, *Rafflesia* is a charismatic icon of plant conservation, but unlike the panda, efforts to conserve and cultivate *Rafflesia* out of its natural habitat have been limited and challenging (Molina et al., 2017; Wicaksono et al., 2016, 2021). This is perhaps a direct consequence of the little that we know of its biology, especially of its seeds. Small mammals, like shrews previously observed to gnaw on the hard covering of the fruits, are believed to facilitate seed dispersal (Bänziger, 2004; Nais, 2001), though ants have also been seen transporting the seeds (Pelser et al., 2013, 2016). How these seeds germinate and infect the host is unknown (Wicaksono et al., 2021), but after the seed germinates in the host, it seems that the *Rafflesia* endophyte can persist in the vegetative stage for years (M. Gabin, personal communication, July 28, 2016; Bascos et al., 2021).

Molina et al. (2017) inoculated *Rafflesia* seeds in uninfected *Tetrastigma* host species growing in the US Botanic Garden, following techniques by Marius Gabin, a *Rafflesia* grower in Malaysia, but the emergence of the bud has yet to be seen. Ex situ propagation attempts from *Rafflesia* seed have been unsuccessful except for one—a blooming flower of *Rafflesia arnoldii* was reported in 2022 in Bogor Botanic Garden, Indonesia, presumably a result of a seed inoculation experiment several years prior (S. Mursidawati and D. Latifah, personal communication, January 5, 2023). However, grafting *Rafflesia*-infected *Tetrastigma* to an uninfected rootstock seems to be the most viable method as implemented in Bogor Botanic Garden, producing multiple blooms for public display (Mursidawati et al., 2015; Thorogood et al., 2021), and exemplifies the importance of these emblematic species in promoting plant conservation awareness. To date, only one species, *Rafflesia patma*, has been propagated ex situ (Mursidawati et al., 2015). It is estimated that 76% of parasitic plant diversity is missing in botanic gardens (Thorogood et al., 2021), including rare holoparasites like *Rafflesia*, translating to lost opportunity to raise awareness and appreciation for these evolutionary marvels and their threatened habitats.

In vitro experiments to induce *Rafflesia* seed germination using various plant growth regulators such as gibberellins, ethylene, brassinolide, and cytokinin have also been unsuccessful (Molina et al., 2017). Even the synthetic strigol, GR24, known to stimulate seed germination in plant parasites of Orobanchaceae, did not induce germination. Transcriptome sequencing of germinating seeds of the Orobanchaceae parasite, *Phelipanche aegyptiaca* exposed to GR24 revealed numerous differentially expressed genes related to DNA, RNA, protein repair and biosynthesis, carbohydrate, and energy metabolism. Upregulation of abscisic acid (ABA), gibberellin (GA), and ethylene-associated genes were also observed, with seed germination related to the increasing concentration of GA and ethylene and concomitant reduction in ABA (Yao et al., 2016). In fact, these plant hormone interactions during seed germination are not unique to plant parasites and generally apply to many plants, including the model species *Arabidopsis* (Carrera-Castano et al., 2020).

Ranjan et al. (2014) profiled the transcriptomes of different developmental stages of the holoparasite *Cuscuta pentagona* (Convolvulaceae), and in the pre-haustoria seedling stage, they found an enrichment of genes associated with responses to stress, biotic, and endogenous stimuli, as well as transporter, kinase, and catalytic activities. The haustorium is the specialized feeding structure in many

plant parasites that regulates transfer of nutrients from host to parasite (Teixeira-Costa, 2021; Yang et al., 2015). Yang et al. (2015) studied the genetic basis of haustorial development in plant parasites of Orobanchaceae, namely, *Phelipanche*, *Tryphysaria*, and *Striga*, characterizing the shared upregulated genes as "parasitism genes," having enhanced expression in the haustoria. These included genes for proteolysis such as serine carboxypeptidase and aspartyl protease, cell wall modification genes (e.g., pectate lyase), and peroxidase genes. They discovered that orthologs of these haustorial genes were expressed significantly in root and pollen tissues in nonparasitic model plant species suggesting that haustorium genes evolved from these nonparasitic but similarly intrusive plant structures.

Unlike the seeds of Orobanchaceae which are responsive to host-derived strigolactones (Wang et al., 2017), it is yet unknown what triggers germination and development of the *Rafflesia* seed. In this study, we seek a better understanding of *Rafflesia* seed biology by comparing a newly generated transcriptome of *Rafflesia speciosa* seeds to published seed transcriptomes of *Arabidopsis* (Dorone et al., 2020), *Striga* (Yang et al., 2015), *Cuscuta* (Ranjan et al., 2014), and *Anoectochilus* (Liu et al., 2015). The orchid *Anoectochilus* is initially mycoheterotrophic during seed germination, reliant on mycorrhizal fungi due to lack of nutritive reserves in the minute seeds until leafy structures begin to photosynthesize (Liu et al., 2015). *Rafflesia* seeds are comparatively tiny (0.7–0.9 mm; Figure 1) with scant oily endosperm enveloping the tiny undifferentiated embryo (Bouman & Meijer, 1994). JM (unpublished) has characterized diverse endophytic microbiota in *R. speciosa* seeds with plant-growth promoting properties in the literature (Khalaf & Raizada, 2016). It is possible that *Rafflesia* may be similarly dependent on microbial symbionts during germination, similar to *Anoectochilus* and other mycorrhizal-dependent orchids. To our knowledge, this study presents the first seed transcriptome of *Rafflesia*. Here, we present the seed genetics of this unique holoparasite and

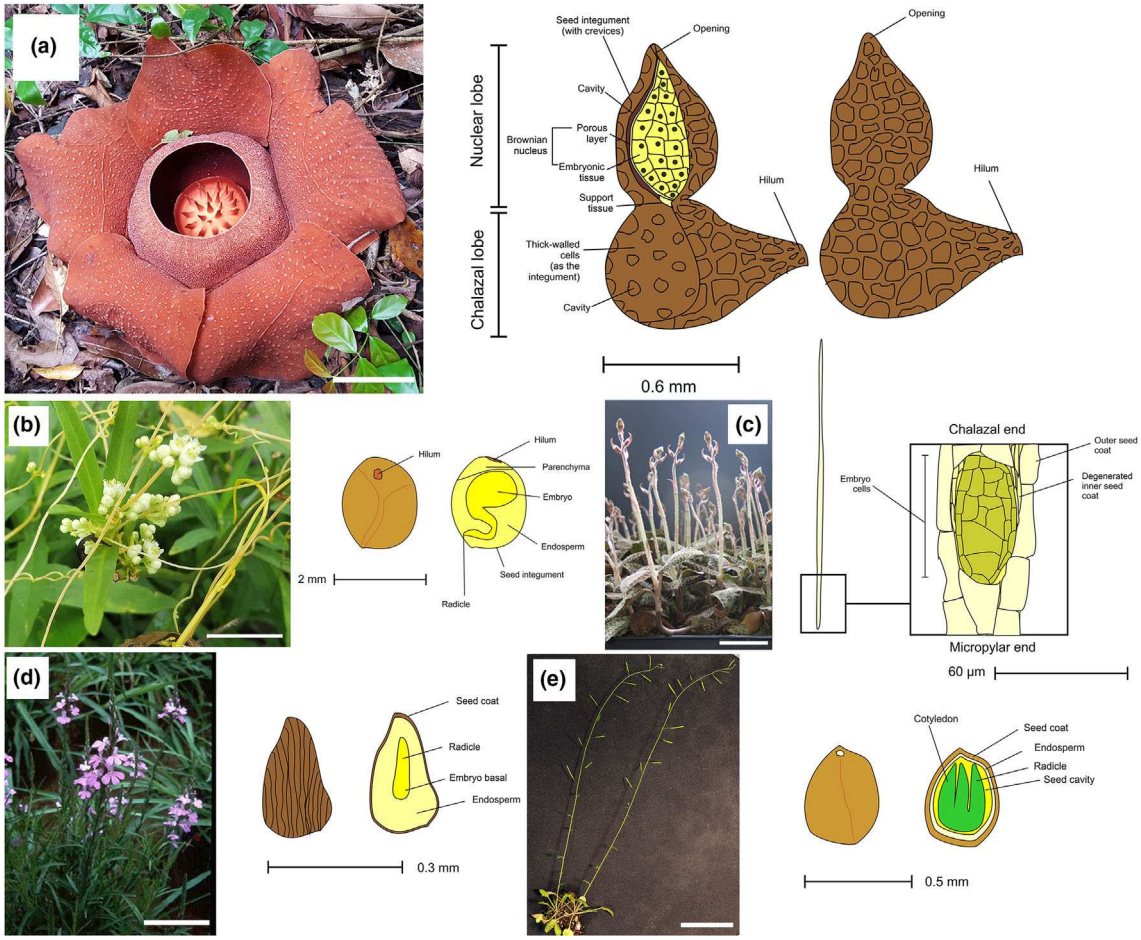


FIGURE 1 Plants in this study and their corresponding seed anatomy. (a) *Rafflesia speciosa* (photo by J. Molina, scale bar = 10 cm), seed anatomy modified from Wicaksono et al. (2021) to reveal the internal tissue of the embryo based on Ng et al. (2018); (b) *Cuscuta pentagona* (photo by Mason Brock, scale bar = 1 cm), seed anatomy inspired by Olszewski et al. (2020); (c) *Anoectochilus roxburghii* (image from Ye et al., 2017, *J Ethnopharm* 209: 184–202, with permission, scale = 5 cm), seed anatomy inspired by Li et al. (2019); (d) *Striga hermonthica* (photo from USDA ARS, public domain via Wikimedia Commons, scale = 5 cm), seed anatomy inspired by Seed Identification Guide (2018), www.idseed.org; and (e) *Arabidopsis thaliana* (photo by SH Kwak, scale = 5 cm), seed anatomy inspired by Haughn and Chaudhury (2005). The images of *C. pentagona* and *S. hermonthica* were taken by Mason Brock and the United States Department of Agriculture: Agricultural Research Service, respectively and are available in Wikimedia public domain with no copyright. The image of *Anoectochilus roxburghii* is added by permission from Elsevier.

describe its similarities and differences to other plants, including other parasites and a mycoheterotrophic orchid. We discuss how insights from the *Rafflesia* seed transcriptome can benefit future ex situ propagation and conservation efforts of the world's largest flower.

2 | MATERIALS AND METHODS

2.1 | RNA extraction and library preparation

Seeds (0.5 g) from one fruit of *R. speciosa* were collected in August 2016 at Miagao, Iloilo, Philippines with permission from the Philippine government (Gratuitous Permit 242). The seeds were scraped from the dehisced fruit and air dried for transport to the United States with the appropriate transport and export permits. The dried seeds were ground in liquid nitrogen with total RNA extracted using Qiagen RNeasy Plant Mini Kit (Cat. No. 74904) following manufacturer's protocol resulting in 6.24 ng/μL of extracted RNA. Eluted RNA was treated with Ambion TURBO DNA-free kit (Thermo Fisher Scientific, Cat. No. AM1907) and then quantified by Qubit RNA BR assay. Ribo-Zero rRNA removal procedure was performed according to the manufacturer's instructions. rRNA-depleted RNA was cleaned up using the Agencourt RNAClean XP Kit (Beckman Coulter, Cat. No. A63987). ScriptSeq Complete Seed/Root kit (Illumina, Cat. No. BSR1206) was used for library prep following the protocol for "Low Input" RNA. The RNA-seq library was purified using the AMPure XP system. Final library quantification and quality control were performed using the Agilent Tape Station. Libraries were sequenced using Illumina protocols for 2 × 100-bp reads on an Illumina HiSeq 2500 instrument at the New York University GeneCore facility.

2.2 | *Rafflesia* seed transcriptome assembly

Illumina reads were filtered and trimmed for quality using Trimmomatic (v0.27). The resulting high-quality reads were assembled with Trinity (v2.8.5) into contigs. For comparison, raw data files for seed transcriptomes of *Striga hermonthica* (StHe0GB1, imbibed seed [Yang et al., 2015]), *C. pentagona* (SRR966236, seed condition whether dry/imbibed not stated [Ranjan et al., 2014]), *Anoectochilus roxburghii* (SRR2962491, dry seeds [Liu et al., 2015]), and *Arabidopsis thaliana* (GSM5100805, dry seeds [Dorone et al., 2020]) were downloaded and independently do novo assembled using Velvet 1.2.10 in Geneious Prime (Biomatters Ltd.) with default parameters.

The fasta files of the assembled contigs of all five taxa were each imported into OmicsBox (Biobam Bioinformatics, Valencia, Spain), blasted against Magnoliopsida nonredundant protein sequences using blastx-fast algorithm with e-val $1.0e^{-100}$. For *Rafflesia*, additional blasting against fungi was performed, and blast results were compared with those obtained for Magnoliopsida. Transcripts that had higher similarity values for fungal genes versus plant genes were eliminated in further analyses as these transcripts may be from endophytic fungi within *Rafflesia* seeds. After filtering, Salmon (v1.6.0) was used to map the reads to the transcriptome followed by transcript quantification for quality control. Gene ontology (GO) mapping (Götz et al., 2008) using the latest database version was performed on all protein blast hits and then annotated with Blast2GO. Statistics related to Blast hits (top hit species), annotation (direct GO count biological process, molecular function, and cell component), and enzyme code mapping were generated for each transcriptome.

3 | RESULTS

The read set (158,907,700) had an overall mapping rate of 92.1% and determined to be suitable for downstream analysis. Ninety-four percent (135,913,373) of our mapped reads were associated with highly expressed transcripts (TPM > 1000). Of the ~90K assembled *Rafflesia* seed contigs, over 7K had blast hits using Blastx-fast algorithm ($1.0e^{-100}$) in OmicsBox against nonredundant protein sequences of Magnoliopsida. Ninety-one percent (91%) of these had GO annotations (Table 1). Comparable statistics for other plant taxa are also given in Table 1. Table 2 provides the proportion of contigs in the *Rafflesia* seed transcriptome. The top 10 blast hits for *Rafflesia*

TABLE 1 Descriptive statistics of the de novo assembled transcripts for taxa in this study including *Rafflesia*.

Plant taxon and reference	De novo assembled contigs	Number of BLAST hits	Percentage with GO annotations from number of BLAST hits
Arabidopsis (dry, GSM5100805; Dorone et al., 2020)	20,236	10,995	93.4
Anoechtochilus (dry, SRR2962491; Liu et al., 2015)	63,274	5367	76.7
Cuscuta (unknown seed condition, SRR966236; Ranjan et al., 2014)	29,037	2120	83.2
Rafflesia (air-dried, this study)	89,950	7025	91.2
Striga (imbibed, StHe0GB1; Yang et al., 2015)	13,872	5584	98.5

Note: Published raw data sets for *Anoechtochilus*, *Arabidopsis*, *Cuscuta*, and *Striga* were downloaded and de novo assembled producing contigs (>300 bp) that were blasted and annotated.
Abbreviation: GO, gene ontology.

TABLE 2 Contig statistics for the *Rafflesia* seed transcriptome.

Length range (bp)	Count	Percentage
300–600	121	1.7
601–1000	1978	28.2
1001–2000	4011	57.1
2001–3000	681	9.7
3001–4000	167	2.4
4000+	67	0.9
Total	7025	100

Note: The proportion of contigs with a certain length, from those with BLAST hits, is given.
Abbreviation: bp, base pairs.

(Figure 2) included Malpighiales taxa (*Hevea*, *Manihot*, *Jatropha*, *Ricinus*, and *Populus*, total 49% transcripts) to which *Rafflesia* belongs, though 2% transcripts were most similar to *Vitis* of Vitaceae, the host family. We do not think this is contamination from the host because the *Rafflesia* seeds were enclosed in the partially dehisced fruit and the seeds were not touching the host root.

Because of the differences in sequence coverage across datasets resulting in discrepant counts for assembled contigs, we compared the proportions for GO counts for biological process (BP), molecular function, and cellular component (Figure 3) for each taxon (#GO for a specific category / total #GO for the taxon, e.g., 228 *Rafflesia* transcripts for protein phosphorylation divided by all *Rafflesia* transcripts annotated under BP). The most enriched GO functions for BP in the *Rafflesia* seed were "protein phosphorylation," "proteolysis," and "phosphorylation," though still proportionally lower compared with other taxa (Figure 3a). As for molecular function GO, *Rafflesia* seed transcripts were enriched for "ATP binding," "metal ion binding," and "ATP hydrolysis activity," like other taxa (Figure 3b). However, for cellular component (Figure 3c), *Rafflesia* seed values were below average for "nucleus" but higher for "cytoplasm" which may be related to the higher activity of cytoplasmic-associated Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways such as "protein processing in endoplasmic reticulum (ER)," "glycolysis/gluconeogenesis," "pyruvate metabolism," "oxidative phosphorylation," "peroxisome," and "fatty acid degradation." "mRNA surveillance pathway and beta-alanine metabolism" were also increased (Figure 4). The *Rafflesia* seed transcriptome was proportionally higher for fatty acid metabolism, compared with the seed transcriptomes of the other taxa studied here, based on the heatmap of homologous oil metabolism genes (Figure 5) involved in seed germination (Miray et al., 2021).

Sixty-seven percent (67%) of *Rafflesia* transcripts were annotated as enzymes. Similar proportional trends for various enzyme classes were apparent across taxa (Figure 6) with transferases as the most abundant enzyme class in all seed transcriptomes examined (>35%), and "transferring phosphorus-containing groups" is the top function in all taxa (results not shown). Hydrolases composed about 30% of the enzymes present in all seed transcriptomes. Oxidoreductases made up over 10% of seed transcripts.

Figure 7, which was adapted from Carrera-Castano et al. (2020), depicts the interplay of various phytohormones in germination and dormancy. We discuss below potential interaction of phytohormonal response proteins and germination-dormancy molecular regulators detected in the *Rafflesia* seed, as shown in Figure 8, in regulating its development. The *Rafflesia* seed had transcriptionally active genes for GA, brassinosteroid, and salicylic acid, which were all active in *Striga* and *Arabidopsis*, the latter having additional hormone biosynthesis genes present. Interestingly, the *Rafflesia* seed had higher proportion of GA synthesis genes compared with other taxa (Figure 9).

The transcripts of regulatory factors such as TRANSPARENT TESTA GLABRA1 (TTG1), HOMEODOMAIN GLABRA2, and GLABRA2-EXPRESSION MODULATOR, as well as other metabolic genes that are involved in phenylpropanoid synthesis pathway, were identified in the *Rafflesia* seed transcriptome (Table 3), including cinnamoyl-CoA reductase—a key enzyme in lignin biosynthesis (Table 3). Our results also corroborate previous studies on the absence of a plastome (Cai et al., 2021; Molina et al., 2014). We did not detect any plastid-encoded mRNA in *Rafflesia* seeds. However, *psbA*, which encodes for the D1 protein of Photosystem II, was present in the other four taxa studied including the plant parasites *Striga* and *Cuscuta* as well as the mycorrhizae-dependent seed *Anoectochilus* (Figure 10). Over 2% of the transcripts had *Vitis* as the top hit (vs. a member of Malpighiales) and were presumably horizontally transferred (HT; Table 4). Many transcripts are involved in "DNA integration" including transposon integration perhaps mediated by "nucleic acid phosphodiester bond hydrolysis," another enriched GO. Gene trees for SERK, RHM1, thiC, and glutamyl-tRNA (Gln) amidotransferase subunit A demonstrate putative horizontal transfers from the host family (Figure 11).

4 | DISCUSSION

4.1 | Seed comparison

To better understand the genetics of the *Rafflesia* seed, we generated its transcriptome and compared it with de novo assembled seed transcriptomes from the plant parasites, *Cuscuta* and *Striga*; the mycorrhizal orchid, *Anoectochilus*; and the model photosynthetic plant, *Arabidopsis* (Table 1). The seeds of the different taxa compared here are all tiny with an embryo and scant endosperm surrounded by a seed coat (Figure 1). *Rafflesia* seed has a distinctive bilobed shape with tiny crevices all over its woody seed coat. Within the seed, the oily endosperm is one layered and envelopes the tiny embryo, with 14–22 cells and undifferentiated cotyledons (Baskin & Baskin, 2022). The mycorrhizal-dependent *Anoectochilus roxburgii* seed is twice the length (~2 mm) of the *Rafflesia* seed. Proteins and lipids are stored within the embryo proper. There is some lignin in the seed coat (Li et al., 2019). *Striga* seed is even more miniscule, at 0.2–0.35 mm in length, with an endosperm that can only support the embryo for a few days (Runo & Kuria, 2018). *Cuscuta* seed size (length) ranges from 0.7 to 3 mm with an embryo surrounded by hard endosperm but no

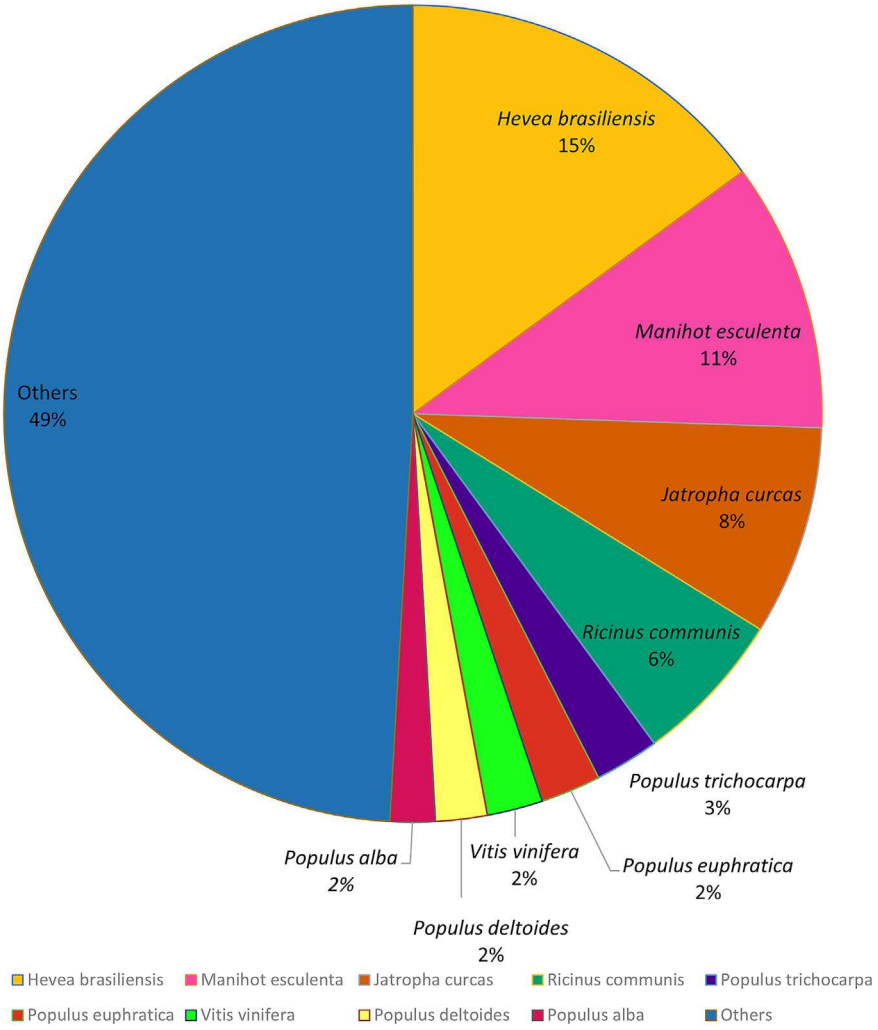


FIGURE 2 Top blast hits of the *Rafflesia* seed transcripts. The top hits included Malpighiales taxa (*Hevea*, *Manihot*, *Jatropha*, *Ricinus*, *Populus*, total 49% transcripts) to which *Rafflesia* belongs, though 2% transcripts were most similar to *Vitis* of Vitaceae, the host family.

cotyledons (Olszewski et al., 2020). Both *Cuscuta* and *Striga* can germinate in soil until they exhaust their nutritive reserves, at which point, they need to have colonized their hosts. The nonparasitic *Arabidopsis* also has a very small seed, ~0.2 mm, composed entirely of embryo filled with lipid and protein and enclosed with one layer of endosperm (Haughn & Chaudhury, 2005).

We were able to assemble >7K contigs from the *Rafflesia* seed transcriptome, with 4011 (57.1%) contigs 1001–2000 bp in length (Table 2). Forty-nine percent (49%) of the assembled *Rafflesia* seed transcripts were most similar to Malpighiales taxa to which *Rafflesia* belongs, as expected (Figure 2). At least 2% transcripts were most similar to *Vitis* of the grape family, Vitaceae, the host family, which may be a consequence of the rampant horizontal gene transfer between host and parasite given their intimate connection (Cai et al., 2021; Molina et al., 2014; Xi et al., 2012, 2013).

4.2 | GOs and KEGG pathways

We compared the assembled seed transcriptome of *Rafflesia* to those of other taxa including *Arabidopsis* as well as other parasitic plants to determine potential overlap in seed genetic pathways. When the *Rafflesia* seeds were collected, they were damp from rain and then air-dried for transport. This may have begun the process of water uptake in the seeds. Generally, to break dormancy, a seed has to imbibe water initiating a series of metabolic changes including energy metabolism, repair of DNA, degradation of dormancy-associated mRNAs, and restoration of cellular integrity including mitochondrial repair to begin cell respiration (Nelson et al., 2017). As the seed breaks dormancy, energy metabolism is increased which explains the increased cytoplasm-associated "glycolysis/gluconeogenesis" activity (Figure 3) and the high levels for the KEGG pathway "fatty acid degradation" that occurs in the "peroxisome" (Figures 4 and 5). Oxygen consumption also increases due to increased cell respiration, which may explain why "pyruvate metabolism" and "oxidative phosphorylation" were high in the *Rafflesia* seed. It is conceivable that the transported fatty acids are converted to sugars (e.g., sucrose, trehalose, xylose, fructose, and glucose) in the *Rafflesia* seed to fuel the growing embryo (Song & Zhu, 2019). Transcripts associated with the synthesis of these sugars were observed in the *Rafflesia* seed transcriptome (Data S1). There were also multiple transcripts for pyruvate dehydrogenase (PDHA; Data S1). PDHA converts glucose to acetyl-coA for

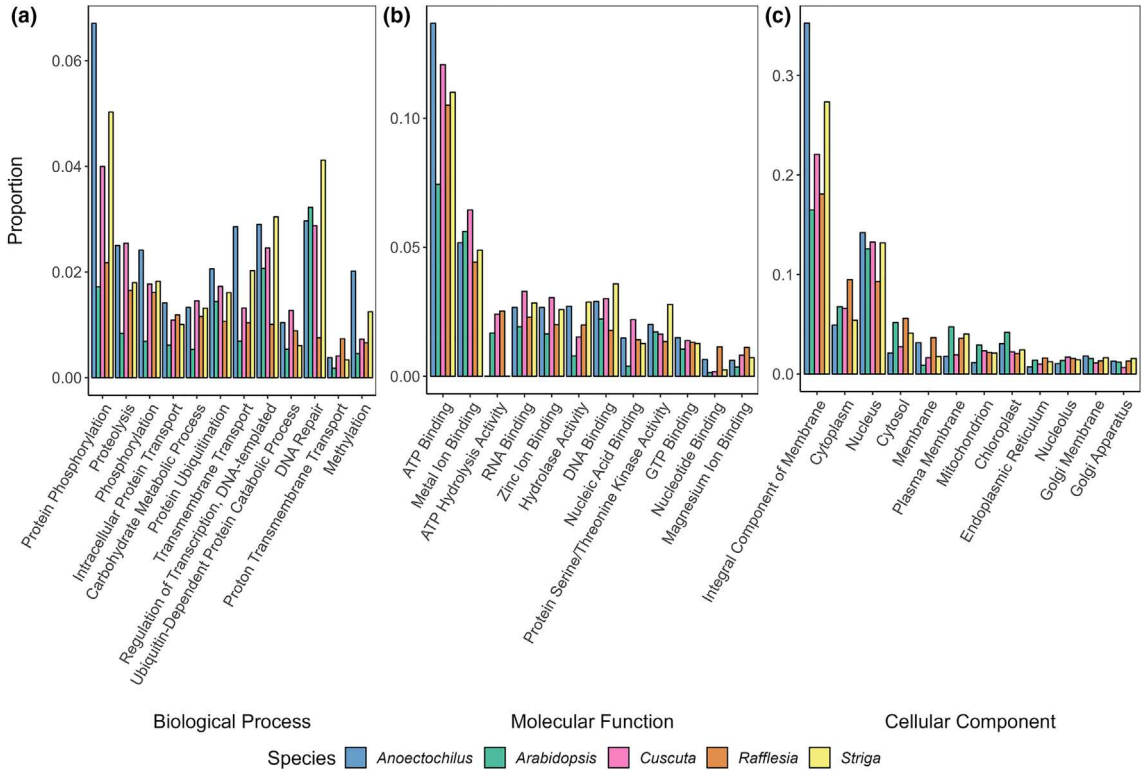


FIGURE 3 (a) Top biological processes in *Rafflesia* seed compared with *Anoectochilus*, *Cuscuta*, *Striga*, and *Arabidopsis*. The top three enriched processes in *Rafflesia* were "protein phosphorylation," "proteolysis," and "phosphorylation," though still relatively lower compared with other taxa. Proportion of *Rafflesia* seed transcripts was the lowest for "protein ubiquitination," "regulation of transcription, DNA-templated," and "DNA repair," though comparatively higher for "proton transmembrane transport" compared with the other taxa. (b) Top molecular function processes in *Rafflesia* seed compared with other taxa. *Rafflesia* seed transcripts were enriched for "ATP binding," "metal ion binding," and "ATP hydrolysis activity," like other taxa. *Rafflesia* was relatively lower for "DNA binding" and "protein serine/threonine kinase activity" yet higher in "nucleotide binding" and "magnesium ion binding" compared with the other taxa. (c) Top cellular component structures in *Rafflesia* seed compared with other taxa. *Rafflesia* transcripts were enriched for "integral component of membrane," "cytoplasm," and "nucleus."

the tricarboxylic acid cycle (Bao et al., 2017). PDHA expression was found to be significantly upregulated in conditioned and strigol-stimulated *Phelipanche* seeds of Orobanchaceae (Bao et al., 2017). Metabolism of the nonproteinogenic amino acid beta-alanine was also relatively increased in the *Rafflesia* seed which may be due to its precursor role in the synthesis of acetyl Co-A (Parthasarathy et al., 2019). It seems that even in the absence of its host, high levels of carbohydrate and lipid metabolism do occur in the *Rafflesia* seed following imbibition.

Synthesis of proteins involved in cell signaling using the seed-stored mRNAs is also increased (Nelson et al., 2017). The KEGG pathway "mRNA surveillance," a quality control mechanism that degrades abnormal mRNAs (Kanehisa et al., 2016), was higher in the *Rafflesia* seed (Figure 4). mRNAs accumulate in seeds during development and are called stored mRNAs/long-lived mRNAs retaining their function as long as the seed is viable and serve as templates for protein synthesis during the germination process. Long-lived mRNAs occur in all plant seeds, even conserved between the dicot *Arabidopsis* and the monocot barley suggesting functional importance in their role to kickstart protein synthesis in early germination (Zhao et al., 2019). Degradation of certain stored seed mRNAs has been shown to promote seed germination (Nelson et al., 2017), resulting in a more "germination-friendly" transcriptome (Carrera-Castano et al., 2020), which may explain the increased "mRNA surveillance" pathway observed in the *Rafflesia* seed.

More than two thirds of *Rafflesia* seed transcripts were annotated as enzymes. Similar proportional trends for various enzyme classes were apparent across taxa (Figure 6), with transferases, hydrolases, and oxidoreductases together making up at least 75% the enzymes. These enzyme classes were also the most abundant types in the *Cuscuta* whole plant transcriptome (Ranjan et al., 2014). The abundance of hydrolase genes in *Rafflesia* seed may be explained by the role of hydrolases in cell wall loosening and seed reserve mobilization during seed germination (Cao et al., 2021). Oxidoreductases catalyze redox reactions that produce reactive oxygen species (ROS), and in parasitic plants, development of the haustoria begins with ROS from the parasite converting host cell wall phenols to benzoquinones stimulating haustorium formation (Cui et al., 2018). It is suspected that oxidoreductase-mediated production of ROS may also be involved in *Rafflesia* seed germination and host infection.

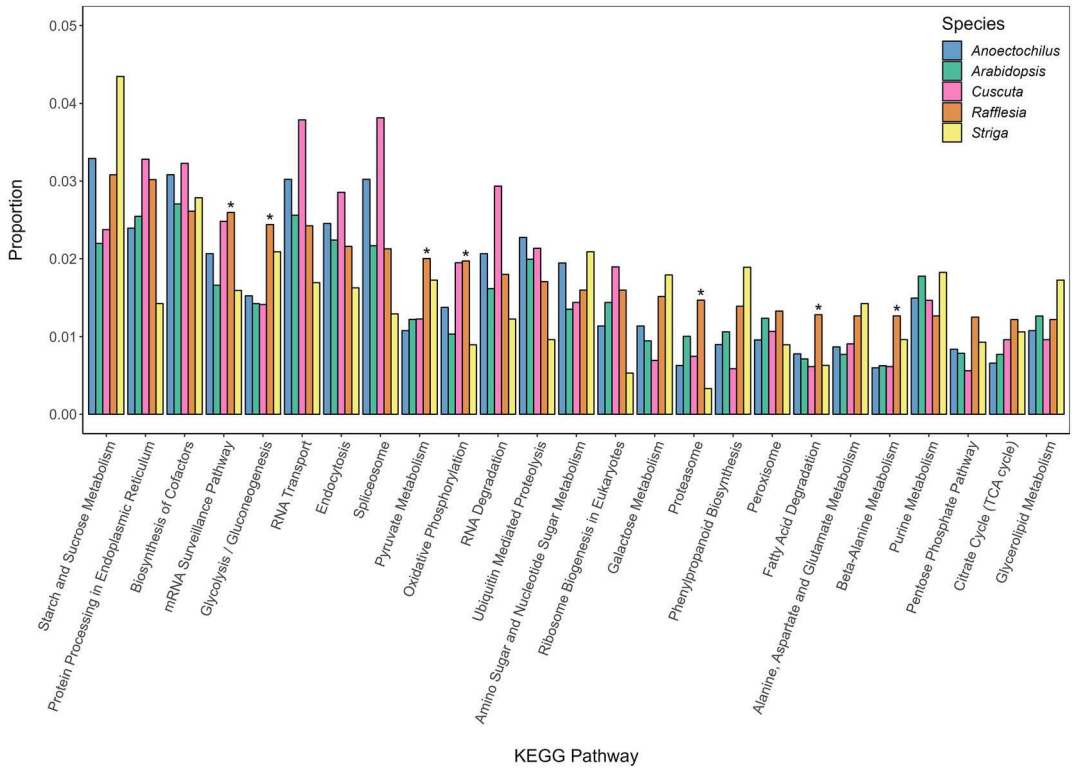


FIGURE 4 Top Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways in *Rafflesia* seed transcriptome compared with *Anoectochilus*, *Cuscuta*, *Striga*, and *Arabidopsis*. *Rafflesia* seed transcripts were relatively higher for the following pathways (indicated with "*" and discussed): mRNA surveillance pathway, glycolysis/gluconeogenesis, pyruvate metabolism, oxidative phosphorylation, proteasome, fatty acid degradation, and beta-alanine metabolism compared with other taxa.

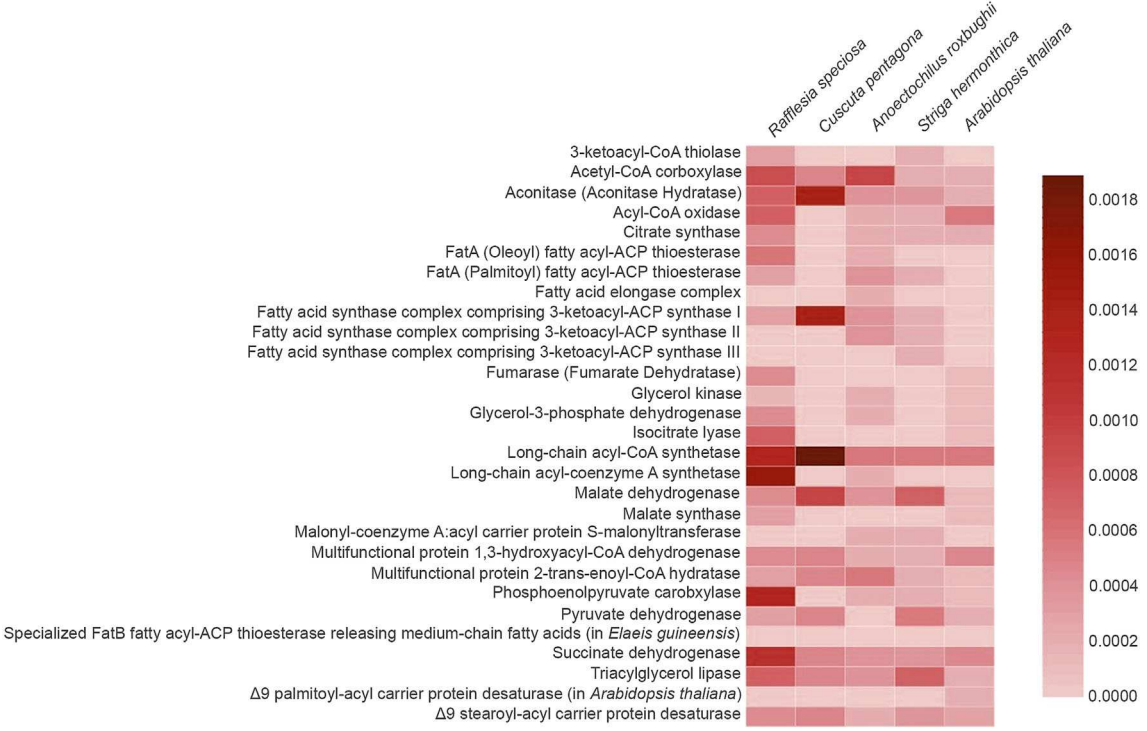


FIGURE 5 Heatmap of oil metabolism genes involved in seed germination. *Rafflesia* seed transcriptome was proportionally higher for fatty acid metabolism (cf. Figure 4) compared with the seed transcriptomes of the other taxa studied (*Anoectochilus*, *Cuscuta*, *Striga*, and *Arabidopsis*).

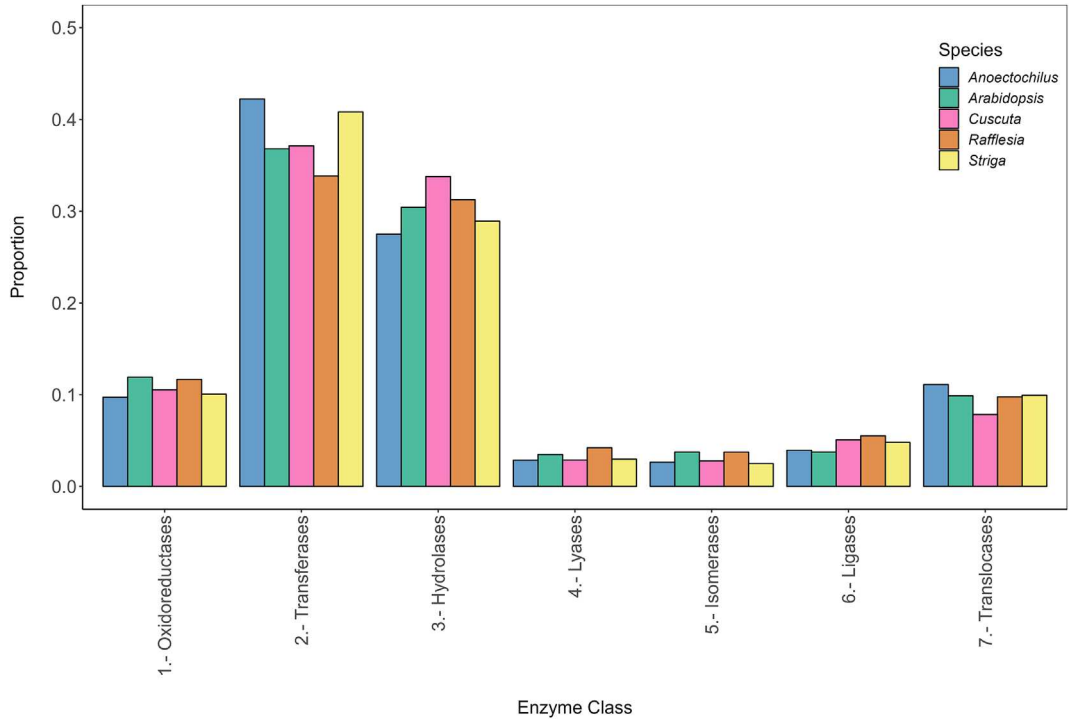
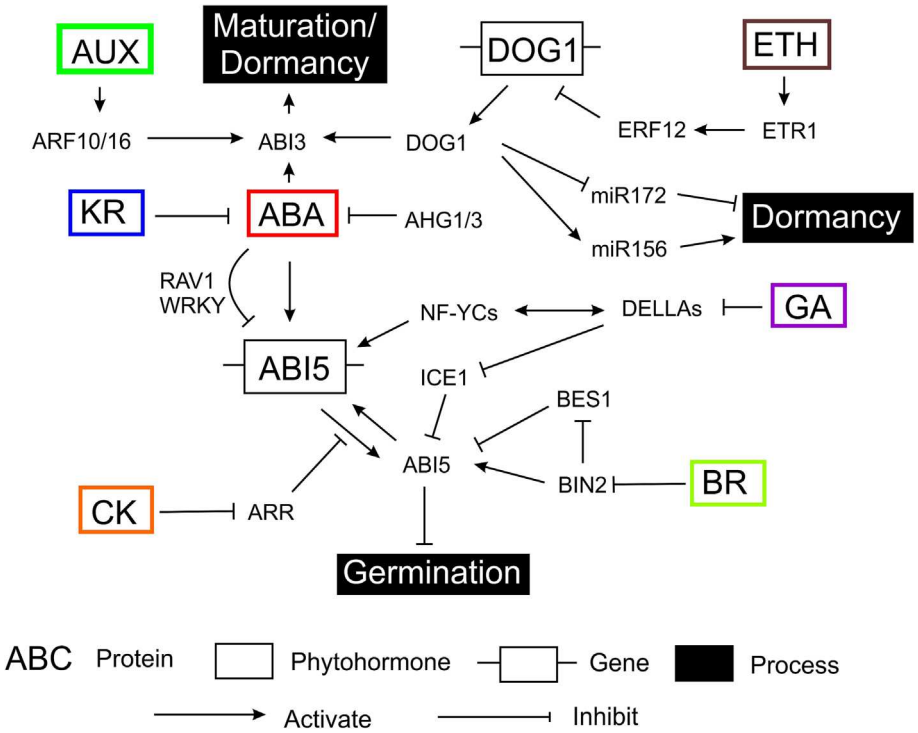


FIGURE 6 Enzyme classes in *Rafflesia* seed compared with *Anoectochilus*, *Cuscuta*, *Striga*, and *Arabidopsis*. Sixty-seven percent (4732/7025) of *Rafflesia* transcripts were annotated as enzymes.

FIGURE 7 Regulatory network of phytohormones involved in germination and dormancy. Figure was adapted from Carrera-Castano et al. (2020). The interaction of auxins (AUX), cytokinins (CKs), gibberellic acids/gibberellins (GA), abscisic acids (ABA), ethylene (ETH), brassinosteroid (BR), and karrikin (KR) controls seed development. Colored outlines are used to distinguish each hormone.



4.3 | Phytohormone signaling during seed development

In general, two hormones, GA and ABA, play antagonistic roles in seed development, with GA promoting seed germination whereas ABA maintaining seed dormancy. The ratio of GA/ABA is also regulated by other hormones such as AUX, KR, ETH, and BR (Figure 7). At e-val < 0.001, homologs of DELLA and DOG1 (delay of germination), which increase ABA sensitivity, were not observed in the *Rafflesia* seed transcriptome, though ABI homologs were recovered,

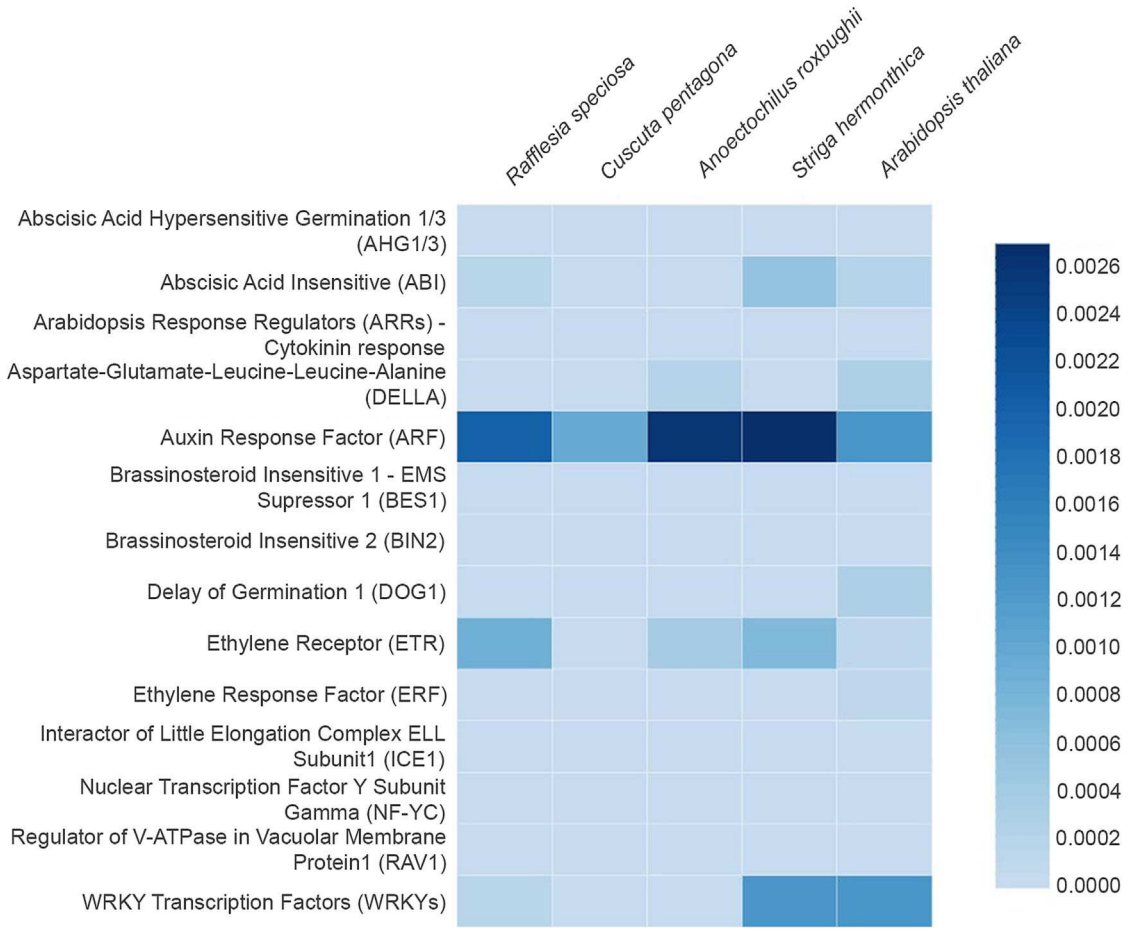


FIGURE 8 Heatmap of phytohormonal response proteins and germination-dormancy molecular regulators, showing the proportion of seed transcripts for each species. Levels of auxin response factor were similarly higher across taxa investigated (*Anoectochilus*, *Cuscuta*, *Rafflesia*, *Striga*, and *Arabidopsis*).

along with WRKY (Figure 8; Data S1). WRKY is a transcription factor induced by ABA and may function as a negative growth regulator. Levels of auxin response factors were similarly high across all taxa (Figure 8), as expected, given that this gene family is greatly involved in regulating the development of seeds. Additionally, transcripts for ETR1, which acts as a negative regulator of response to ethylene (The Uniprot Consortium, 2021), a germination stimulant by DOG1 repression, were also detected. Detection of ABI, WRKY, and ETR1 transcripts may indicate that the *Rafflesia* seeds sampled, though transcriptionally active in the KEGG pathways described above, have yet to actually germinate sensu stricto, that is, when the embryo radicle protrudes from the seed (Carrera-Castano et al., 2020), which in *Rafflesia* conceivably happens in the presence of its host.

Though transcripts for *ent*-kaurene synthetase A and *ent*-kaurenoic acid oxidase, which are vital in GA synthesis, were recovered in the *Rafflesia* seed (Figure 9), such enzymes are also involved in production of secondary metabolites like terpenoids, which *Rafflesia* is known to produce (Ng et al., 2018). Homologs of genes for synthesis of carotenoids were not detected at all in the developmental stage of *Rafflesia* seed sampled, though relevant transcripts were detected in the other seed transcriptomes studied here (Data S1). Chemical inhibition of carotenoid synthesis using fluridone in *Orobanche* and *Phelipanche* (Orobanchaceae) has been found to promote their seed germination (Bao et al., 2017) which may suggest that fluridone may not work in inducing *Rafflesia* germination because carotenoid synthesis genes were not detected. The strigolactone responsive gene D14 (Wang et al., 2017) was also not found in *Rafflesia* seed which may explain the lack of morphological development in the *Rafflesia* seed when exposed to the synthetic strigol GR24 (Molina et al., 2017), though KAI2, which is important in perception of karrikins—fire-related compounds that can stimulate seed germination (Wang et al., 2017), was observed. In addition, the F-box protein MAX2, another key component in karrikin (and strigolactone) signaling, was also recovered (Wang et al., 2017). However, in previous germination experiments of *R. speciosa* seeds, karrikin did not seem to stimulate germination—no protrusion was seen developing on the seeds (Molina et al., 2017), though it could be possible that tested seeds were agamospermous (Nais, 2001). Thus, hand-pollination has been recommended to produce viable seeds (Bänziger, 2004).

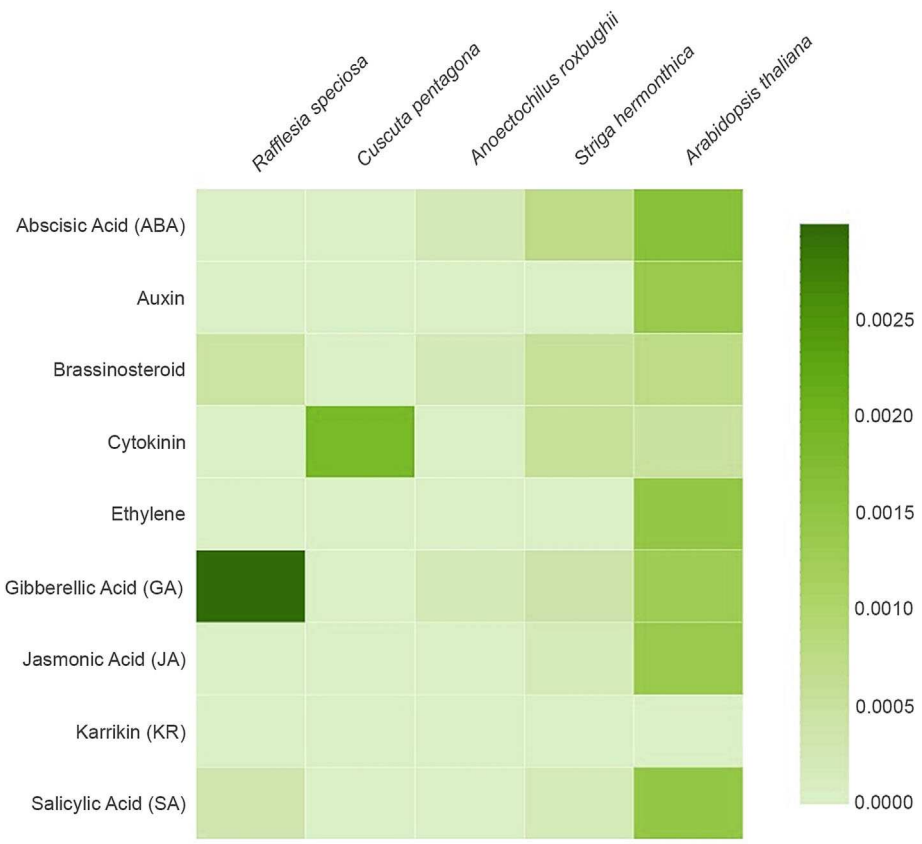


FIGURE 9 Gene ontology count heatmap of phytohormone biosynthesis genes in *Anoectochilus*, *Cuscuta*, *Striga*, and *Arabidopsis*. The *Rafflesia* seed had transcriptionally active genes for gibberellic acids/gibberellins, brassinosteroid, and salicylic acid, with higher proportion of gibberellic acids/gibberellins synthesis genes compared with the other taxa.

TABLE 3 Transcripts that are involved in phenylpropanoid synthesis in *Rafflesia* seed transcriptome.

Identified transcripts	Gene name	*Rafflesia* contig	Expect value	Species of identified homolog
Regulatory factors				
TRANSPARENT TESTA GLABRA1	*TTG1*	41185	9.73E−116	*Manihot esculenta*
GLABRA2-EXPRESSION MODULATOR	*GEM*	86527	2.30E−119	*M. esculenta*
GLABRA2-EXPRESSION MODULATOR	*GEM*	86534	6.34E−119	*M. esculenta*
GLABRA2-EXPRESSION MODULATOR	*GEM*	35727	1.31E−116	*Ricinus communis*
HOMEODOMAIN GLABRA2	*HDG2*	19125	1.03E−159	*Populus alba*
Phenylpropanoid pathway enzymes				
Phenylalanine ammonia-lyase	*PAL*	73687	0.00E+00	*R. communis*
Phenylalanine ammonia-lyase	*PAL*	73688	0.00E+00	*R. communis*
4-Coumarate-CoA ligase	*4CL3*	86494	0.00E+00	*Jatropha curcas*
Chalcone synthase	*CHS*	53706	0.00E+00	*Actinidia chinensis*
Flavonoid 3′-monooxygenase	*F3′H, TT7*	86571	0.00E+00	*P. alba*
Dihydroflavonol 4-reductase	*DFR, TT3*	30307	5.14E−135	*M. esculenta*
Leucoanthocyanidin reductase	*LAR*	60680	2.96E−167	*Quercus suber*
Cinnamoyl-CoA reductase	*CCR*	76321	5.01E−108	*Populus euphratica*
Cinnamoyl-CoA reductase	*CCR*	46001	3.55E−116	*Salix brachista*
Anthocyanidin 3-O-glucosyltransferase	*UGT*	10343	1.94E−168	*Herrania umbratica*
Anthocyanidin 3-O-glucosyltransferase	*UGT*	10348	2.49E−170	*H. umbratica*
Anthocyanin 5-aromatic acyltransferase	*AACT1, ACT*	37598	1.21E−134	*R. communis*
Anthocyanin 5-aromatic acyltransferase	*AACT1, ACT*	37594	8.33E−137	*R. communis*
Flavonoid transporter				
TRANSPARENT TESTA9	*TT9*	74015	7.14E−114	*R. communis*
TRANSPARENT TESTA9	*TT9*	74018	3.24E−126	*R. communis*

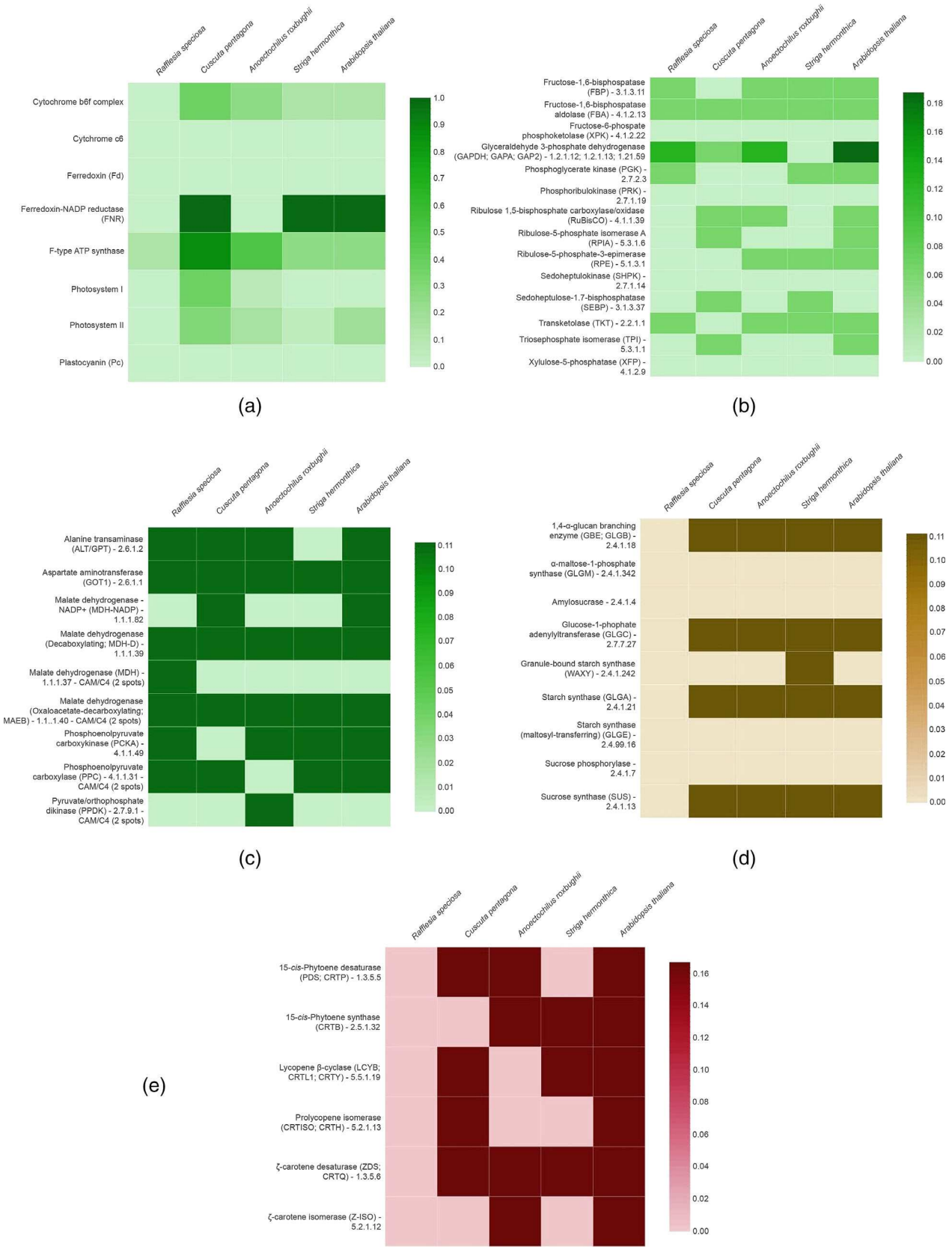


FIGURE 10 Heatmap of plastid-related genes across the five taxa (*Rafflesia*, *Anoectochilus*, *Cuscuta*, *Striga*, and *Arabidopsis*). Genes for light-dependent reaction (a), C3 carbon fixation (b), C4-CAM carbon fixation (c), starch synthesis process (d), and carotene synthesis (e). All enzymes are noted with their respective Kyoto Encyclopedia of Genes and Genomes enzyme IDs (except for the light dependent because some are composed of multiple proteins).

TABLE 4 *Rafflesia* seed transcripts by GO biological process with top blastx hits to *Vitis* (vs. members of Malpighiales) which may indicate putatively horizontally transferred transcripts.

GO biological process	Transcript count
DNA integration	36
Proteolysis	21
RNA-templated DNA biosynthetic process	13
Nucleic acid phosphodiester bond hydrolysis	10
Glutathione metabolic process	7
Protein phosphorylation	5
Peptide catabolic process	4
Others	96

Abbreviation: GO, gene ontology.

4.4 | Proanthocyanidin synthesis in *Rafflesia* seed coat

In mature seeds, proanthocyanidins, produced in the phenylpropanoid pathway, accumulate in the innermost layer of the seed coat, oxidizing to a brown color and are known to protect the seed from environmental stresses and improve the longevity of seeds (Debeaujon et al., 2003). Transcripts involved in this pathway were detected in the *Rafflesia* seed including TTG1. In *A. thaliana*, TTG1 activates the transcription of downstream transcription factors such as GL2 and HOMEODOMAIN GLABRA2, which then regulate the expression of metabolic genes involved in phenylpropanoid synthesis (Li et al., 2020). It is also possible for the *Rafflesia* TTG1 to function in trichome development or other cell fate determination processes because TTG1 is also involved in trichome development and root epidermal pattern formation in *A. thaliana* (Wang et al., 2010). In the unrelated holoparasite, *Balanophora abbreviata* (Balanophoraceae), trichomes originating from endosperm cells, facilitated host attachment (Arekal & Shivamurthy, 1976). Whether the TTG1 gene is involved in the development of a similar root-like structure in the *Rafflesia* seed as it germinates waits to be seen. Other metabolic genes involved in phenylpropanoid synthesis pathway, such as cinnamoyl-CoA reductase—a key enzyme in lignin biosynthesis (Kawasaki et al., 2006), were also identified in the *Rafflesia* seed transcriptome (Table 3).

The presence of the key transcriptional regulator, TTG1, anthocyanidin, and lignin metabolic genes in the *Rafflesia* seed transcriptome suggests that proanthocyanidins accumulate via phenylpropanoid synthesis pathway in the lignified seed coat of *Rafflesia*, which explains the brown color of *Rafflesia* seeds. *Rafflesia*'s lignified seed coat conceivably protects the seed during its long dormancy (Wicaksono et al., 2021) until a suitable host is found. Proanthocyanidins have also been demonstrated to inhibit seed germination in *Arabidopsis* (Jia et al., 2012). As mentioned above, though the *Rafflesia* seeds were transcriptionally active in energy metabolism, detection of the proanthocyanidin-related transcripts may suggest that when the seeds were sampled, germination was still inhibited, pending host stimulation.

4.5 | Plastid processes

Dry seeds of *Arabidopsis* already possess a substantial portion of their plastid-encoded mRNA as stored/long-lived mRNA with de novo mRNA synthesis commencing during imbibition (Demarsy et al., 2012). However, none of the genes in the KEGG light-dependent reactions of photosynthesis (Kanehisa et al., 2016) were present in the *Rafflesia* seed, yet *psbA* was present in the other four taxa studied including *Striga* and *Cuscuta* as well as the mycorrhizae-dependent seed *Anoectochilus* suggesting the importance of the presence of this gene even in dry seeds (Figure 10). *psbA* encodes for the D1 protein of Photosystem II that splits water in the presence of light energy and converts this to ATP to fuel subsequent carbon fixation. It seems that even in the plant parasites, *Striga* (Tuquet et al., 1990) and *Cuscuta* (Dinelli et al., 1993), light-mediated reactions occur to some extent until they are established in their hosts, and this is especially true in *Cuscuta*, which is initially photosynthetic. The endophytic nature of *Rafflesia* may explain *psbA* loss.

Transcripts for *rpl32* (35223–35225; Data S1), a ribosomal plastid protein typically encoded in the chloroplast genome, were also observed in the *Rafflesia* seed transcriptome and, if shown to be indeed localized in the plastome, would contradict the purported unprecedented loss of a plastome in plants (Cai et al., 2021; Molina et al., 2014). The *Rafflesia* seed *rpl32* copy was also phylogenetically allied with Euphorbiaceae (results not shown), thus a native copy and not HT. However, it is possible that this gene has moved from the plastome and is now nucleus encoded, as was discovered in *Rafflesia's* distantly related Malpighiales photosynthetic relative, *Populus* (Ueda et al., 2007).

The *Rafflesia* seed transcriptome had multiple transcripts present in all three carbon fixation pathways comparable with the seed transcriptomes of the other taxa examined (Figure 10). Despite the absence of the light-dependent reactions of photosynthesis, carbon fixation does occur in the *Rafflesia* seed. However, ribulose-1,-5-bisphosphate carboxylase-oxygenase gene (EC 4.1.1.39) was not detected in the *Rafflesia* seed transcriptome, unlike the other four taxa examined. However, *rbcL* subunit binding proteins alpha and beta, also known as chaperonin 60α and 60β, respectively, were detected in *Rafflesia* seed and had intact open reading frames, suggesting they are functional. They also strongly phylogenetically cluster with Euphorbiaceae taxa (results not shown), thus are not HT. Both chaperone proteins are nuclear encoded and essential in *rbcL* folding and binding the subunits of ribulose-1,5-bisphosphate carboxylase-oxygenase, but in *Rafflesia*, the absence of *rbcL* may indicate other functions for the chaperone proteins perhaps related to embryogenesis. In *Arabidopsis*, alpha mutation arrested embryo development (AT2G28000.1, TAIR, 2022), whereas beta mutants, on the other hand, show accelerated cell death to heat shock stress (AT1G55490.1, TAIR, 2022).

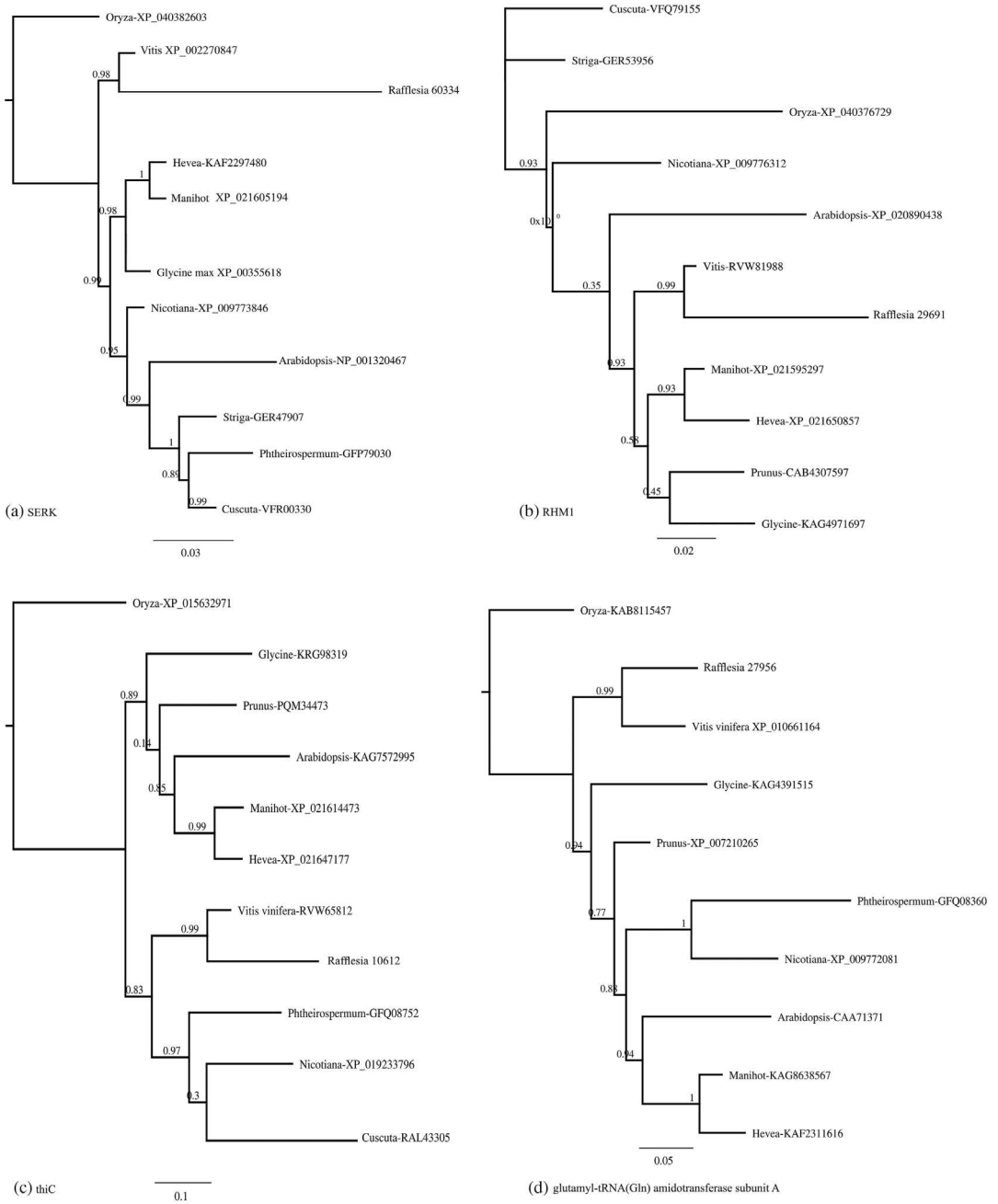


FIGURE 11 Gene trees of *Rafflesia* seed transcripts showing evidence of horizontal gene transfer from the host family (Vitaceae). (a) SERK; (b) RHM1 (rhamnose biosynthesis 1); (c) thiC (phosphomethylpyrimidine synthase gene); and (d) glutamyl-tRNA (Gln) amidotransferase subunit A. BlastX (NCBI) was performed on the *Rafflesia* seed transcripts against the taxa shown in the phylogenies including *Vitis* (representative of the host family, Vitaceae) and Euphorbiaceae (the closest photosynthetic relative of *Rafflesia*, represented by *Manihot* and *Hevea*). Multiple sequence alignment was performed using MAFFT v 7.490, followed by phylogeny reconstruction using PhyML 3.3.20180621 (with LG substitution model and SH-branch like supports), as implemented in Geneious Prime 2022.2.1 (Biomatters, Ltd).

4.6 | Horizontal transfer

Over 2% of the transcripts had *Vitis* as the top hit (vs. a member of Malpighiales) and presumably HT. Table 4 shows that many transcripts are involved in "DNA integration" including transposon integration perhaps mediated by "nucleic acid phosphodiester bond hydrolysis," another enriched GO. Multiple *Rafflesia* seed HT transcripts were annotated as "transposon polyproteins" (Data S1), and an abundance of these was also observed among *Sapria* HT orthogroups (Cai et al., 2021). "RNA-dependent DNA biosynthetic process" was enriched due to the presence of multiple retrotransposon transcripts (Data S1). Horizontal gene transfer is disproportionately

increased in parasitic plants with these host-derived genes being co-opted by the parasite in haustorial function, playing a role in its adaptive evolution (Yang et al., 2016).

Interestingly, in *Rafflesia* seed, among putatively HT genes, one notable transcript (60334) was annotated as SERK1 and phylogenetically allied to *Vitis* (Figure 11). SERK family members interact with R (resistance) proteins to activate the hypersensitive immune response (Su et al., 2020). The rice homolog OsSERK1 has been shown to be involved in defense response against the rice blast fungus (Hu et al., 2005). Cowpea *Vigna unguiculata* possesses RSG-301, which interacts with SERK upon infection by *Striga*. However, resistant *Striga* strains subvert RSG-301 by secreting the effector Suppressor of Host Resistance 4z that has homology to a region of SERK. This inhibits VuPOB1 necessary for the hypersensitive response. Thus, Suppressor of Host Resistance 4z in *Striga* acts as a decoy to circumvent host immune responses (Su et al., 2020). Could it be that *Rafflesia* seed 60334, which is 96% similar to *Vitis* SERK1, is acting similarly to enable *Rafflesia* to undermine *Tetrastigma* host defenses?

Another *Rafflesia* seed transcript 29691 annotated as RHM1 (rhamnose biosynthesis 1), which converts UDP-D-glucose to UDP-L-rhamnose (The Uniprot Consortium, 2021), was also phylogenetically allied to *Vitis* (Figure 11), again suggesting HT. We find this interesting because rhamnose is a precursor of primary cell wall components, and it is tempting to speculate that this could be a mechanism of *Rafflesia* to mimic host cell wall components to avoid host immune response and/or allow *Rafflesia* to colonize host tissue, as rhamnose synthase genes were demonstrated to be crucial for host attachment and penetration and fungal pathogenicity of *Verticillium dahliae* (Santhanam et al., 2017).

In Cai et al. (2021), phosphomethylpyrimidine synthase gene, thiC, involved in biosynthesis of pyrimidine was shown to be HT. Analyzing this gene in *Rafflesia* phylogenetically also showed support for this (transcript 10612; Figure 11), as well as *Rafflesia* seed transcript 27956 annotated as "glutamyl-tRNA (Gln) amidotransferase subunit A" (Figure 11) which was also shown to be HT in *Rafflesia* in Xi et al. (2012). Cai et al. (2021) suggested that host-derived HT genes in Rafflesiaceae may substitute for its genes lost during parasitic evolution.

4.7 | Convergent evolution

Among flowering plants, parasitism has independently evolved at least 12 times, with 1% of all angiosperms parasitic (Nickrent, 2020). There seem to be conserved molecular mechanisms underlying this heterotrophic life history. Across the 12 lineages, parasitic plants similarly secrete cell-wall degrading enzymes and proteolytic enzymes when attached to host plant tissues. Such enzymes have been detected in the haustoria of Orobanchaceae parasites and even in the unrelated stem parasite, *Cuscuta reflexa* (Convolvulaceae; Ranjan et al., 2014). Yang et al. (2015) termed these "core parasitism genes" in Orobanchaceae, including genes for cell-wall modifying enzymes (cellulase, etc.), pectinolytic enzymes (pectin methylesterase and pectate lyase), and for phytohormone synthesis—genes for auxin, gibberellin, brassinesteroid, ethylene, and strigolactone biosynthesis. These genes were also upregulated in *Cuscuta* (Ranjan et al., 2014). Here, in the *Rafflesia* seed transcriptome, we found annotated terms enriched for host-degrading enzymes similar to the "core parasitism genes" of Orobanchaceae (Yang et al., 2015) including cellulases, pectate lyases, and pectin methylesterase, hinting at an evolutionary convergence among disparate lineages of parasitic plants. The homolog of cuscutain gene, a cysteine protease upregulated in the haustoria of the stem parasite *C. reflexa* important for host infection (Bleischwitz et al., 2010), was also detected in the root parasite *P. aegyptiaca* (Rehker et al., 2012) and in the *Rafflesia* seed transcriptome in this study (Data S1; *Rafflesia* contig 4146). Whether this transcript is involved in *Tetrastigma* host infection remains to be seen.

Laccase enzyme transcripts were also recovered in both *Striga* and *Rafflesia* seed transcriptomes (Data S1). Interestingly, laccases are enzymes that can degrade host lignin, producing haustorium-inducing factors that promoted haustorium formation in *Striga* and the confamilial *Phtheirospermum* (Cui et al., 2018). It will be interesting to know if laccase genes in *Rafflesia* function similarly. Metabolic profiling of *Tetrastigma loheri*, host of *Rafflesia lagascae* (Molina et al., 2021), showed the presence of sinapaldehyde, sinapyl alcohol, syringic acid, and acetosyringone, which are molecules that induced haustoria in both *S. hermonthica* and *Phtheirospermum japonicum* (Cui et al., 2018). Testing these compounds in germination experiments of *Rafflesia* seeds may be worthwhile in the future.

During the conditioning of seeds of the parasitic plant *Phelipanche*, the gene gibberellin 20-oxidase (GA20OX1) was upregulated resulting in increased GA3 levels. However, in the *Rafflesia* seed, this GA20OX1 was not detected, though present in the seed transcriptome of *Striga*, which was obtained from imbibed seeds (Yang et al., 2015). Gibberellin 2-oxidase (GA2OX2), which in *Arabidopsis* deactivates GA and promotes seed dormancy, was detected in the *Rafflesia* seed instead. Transcripts for methionine adenosyltransferase, aminocyclopropane-carboxylate synthase, and aminocyclopropanecarboxylate oxidase, which are three key enzymes of ethylene biosynthesis, were also detected in *Rafflesia* seed, and these corresponding genes were upregulated significantly in *Phelipanche* after GR24/strigol treatment. Taken together, the *Rafflesia* seed sampled here may have been imbibed, kickstarting the transcription of the early seed processes detected so far but still need to undergo host stimulation. The increased fatty acid breakdown seen in the *Rafflesia* seed transcriptome may be associated with lipid mobilization and increased gluconeogenesis in the *Rafflesia* seed to supply sugars to the embryo, as it awaits host stimulation.

In the previously published *Cuscuta* seed transcriptome (Ranjan et al., 2014, Supplement 13), transcripts involved in "fructose metabolic process," "mannose metabolic process," "glucose 6-phosphate," "glucokinase activity," "response to heat," and "hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds, in linear amides" were enriched and were also detected in *Rafflesia* seed. Moreover, *Rafflesia* and *Striga* were similar proportionally for certain pathways such as "galactose metabolism" and

"phenylpropanoid metabolism," in contrast to *Arabidopsis*. In Orobanchaceae, planteose, a galactose-derived molecule, was found as a storage carbohydrate required for early-stage germination (Wakabayashi et al., 2015). Perhaps, a similar carbohydrate is required by the *Rafflesia* seed during germination which may be causing the uptick in "galactose metabolism." The similar proportion in "phenylpropanoid metabolism" for *Striga* and *Rafflesia* seeds may be due to the synthesis of the lignin components in their seed coats and other proanthocyanidins (Table 3). This woody seed coat perhaps allows *Striga* seeds to survive for prolonged periods of time (>10 years) in the seed bank prior to germination (Atera et al., 2011), which may also be the case for *Rafflesia* seeds.

Comparing *Rafflesia* seed with the mycoheterotrophic seeds of *Anoectochilus*, we were unable to find in *Rafflesia* the homolog of the lipid biosynthetic enzyme, RAM2, which promotes arbuscular mycorrhizal symbiosis (Bravo et al., 2017). However, RAM2 was found in mycorrhizal *Anoectochilus*, as expected, annotated as "glycerol-3-phosphate 2-O-acyltransferase." The absence of RAM2 in *Rafflesia* seed may indicate that *Rafflesia* is not dependent on arbuscular mycorrhiza or that this gene is not yet expressed in the seed developmental stage we sampled.

5 | CONCLUSIONS AND IMPLICATIONS FOR *RAFFLESIA* CONSERVATION

In this study, we produced the first *Rafflesia* seed transcriptome and compared it with seed transcriptomes of a variety of plant species across the phylogenetic tree. We were particularly interested in explaining the difficulty of germinating *Rafflesia* seeds in the laboratory as well as elucidating metabolic pathways that could be involved in seed development of *Rafflesia*. Overall, metabolic pathways of *Rafflesia* are similar to the other plant species we compared it to; however, there are notable findings. Our data suggest that following imbibition, the *Rafflesia* seed is already transcriptionally active, showing increased energy metabolism, preparing the seed for germination upon host stimulation. Our transcriptome results provide further evidence that the plastome is absent in *Rafflesia* yet rife in multiple transposable elements, many apparently HT from the host, similar to its confamilial relative *Sapria*. Furthermore, we found evidence of genes in the *Rafflesia* seed putatively HT from the host. *Rafflesia* could have co-opted these HT genes to potentially inhibit host immune response during infection.

The *Rafflesia* seed transcriptome is also enriched for many classes of the "core parasitism genes" found in Orobanchaceae. The lack of carotenoid synthesis genes in *Rafflesia* seed suggests that fluridone may not induce germination as it does in parasitic Orobanchaceae. Similarly, lack of strigolactone responsive genes may explain the absence of morphological development when *Rafflesia* seeds were exposed to synthetic strigol. However, detection of karrikin response transcripts in *Rafflesia* seed suggests that karrikin may stimulate *Rafflesia* seed germination especially if they were collected from fruits produced by hand-pollination. It is also worthwhile inoculating *Rafflesia* seeds in sinapaldehyde, sinapyl alcohol, syringic acid, and acetosyringone, which are molecules that induced haustoria in Orobanchaceae parasites and are also present in *Tetrastigma* hosts. Transcripts for host-degrading laccase enzymes implicated in haustorium formation in Orobanchaceae were also detected in *Rafflesia*, which may suggest that *Rafflesia* seed inoculation in *Tetrastigma* may benefit from laccase application. It is also possible that *Rafflesia* seeds, like *Striga* seeds, can survive for prolonged periods, given its thick woody seed coat and accumulation of phenylpropanoids. Additionally, we discovered the homolog of the cuscutain gene, important for host infection by *Cuscuta*, in the *Rafflesia* seed. The present study thus provides strong evidence for a theory of generalized convergent evolution among disparate lineages of parasitic plants. We were unable to find homologs of genes involved in mycorrhizal symbiosis in orchids, suggesting that mycoheterotrophy may not be a feature of *Rafflesia* nor required for its germination/infection.

With this research, we take one step closer to understanding the seed biology of *Rafflesia* and recalcitrant parasitic plants in general. Developing the toolbox and skillsets necessary for propagating endangered parasitic plants is required to successfully develop conservation and management strategies for these species and their ecosystems. This research provides several tactical guidelines for continuing ongoing efforts to propagate *Rafflesia*. For example, future germination experiments can deprioritize searching for elicitors that are typical of mycoheterotrophy. However, we should prioritize experiments applying laccase and karrikin on conditioned *Rafflesia* seeds produced from manual pollination. If successful, this approach can also be applied to other difficult to propagate plant species with unique life history traits.

Our research group is dedicated to increase the public's awareness so as to engender engagement in plant and habitat conservation. People are awed every time an *Amorphophallus titanum* blooms at a public garden because of its remarkable size, appearance, and unpleasant smell. It is a charismatic species that can help people connect with and support the natural world. We believe that *Rafflesia* has significant untapped engagement potential. We hope this seed transcriptome study brings us closer to a reality in which botanic gardens all over the world grow *Rafflesia* for conservation, while increasing public awareness and appreciation of an evolutionary marvel—the world's largest flower.

AUTHOR CONTRIBUTIONS

Jeanmaire Molina conceived the project, collected samples, and conducted the experiments. Ronniel D. Pedales, Danilo Tandang, and William McLaughlin collected samples with Jeanmaire Molina. Zoé Joly-Lopez performed experiments. Adhityo Wicaksono, Todd P. Michael, Su-Hwan Kwak, Ronniel D. Pedales, Semar Petrus, Allen Mamerto, Brian Tomek, Sumaya Ahmed, Venkatasivasankar Maddu, Kristina Yakubova, and Jeanmaire Molina analyzed the data. Jeanmaire Molina wrote the manuscript with help from Adhityo Wicaksono, Todd P. Michael, Su-Hwan Kwak, Zoé Joly-Lopez, Joseph W. Morin, So-Yon Park, Hyun-Oh Lee, Kyle Wallick, James Adams, Ari Novy, Susan Pell, and Michael D. Purugganan.

ACKNOWLEDGMENTS

We are most grateful to our collaborators in the Philippines, to our field guides and assistants, to the Miagao municipality headed by Richard Garin and Macario Napulan, and to their very helpful staff: Jane Branton, Mark Moruga, Isidro Mosura, and so forth. We thank Julie Barcelona, Marites Muyong and family, Auring Nopat and family, Tigmalapad Brgy. Cap. Janette Momay, Kagawad Raymond Nugara, DENR's Nermalie Lita, Josefina de Leon, Anson Tagtag, Giselle Garcia, Nato Andraje, Hector Garrido, Llane Orale, Edgardo Ferrer, Livino Duran, Raul Lorilla, Faith of R6 Penro, Pola Bumanglag, Teresita Paderna, Jojie Gereza, Hiro Masuda, Maricar Escuela, former RD Jim Sampulna, former BMB director Mundita Lim, BPI's Lea Gella Blancaflor, PNM's Ana Labrador, Tito Evangelista, JohnRey Callado, Mike Galindon, Jhaydee Pascual, LIU's Lynn Cabralis, Marni Goldner, Angela Stark, Rick Nader, Randy Burd, Michael Marino, Barbara Neal, Jared Littman, USBG's Stephen Jones, and Mick Erickson, the very helpful USDA staff, and the Bio faculty of University of the Philippines-Diliman with special mention to RDP's mentor, Ian Fontanilla, UP scientists Gisela Concepcion, Arturo Lluisma, and to all who have contributed one way or another. Sofi Mursidawati and Hans Bänziger have been a constant source of inspiration to JM and AW in *Rafflesia* conservation. We also thank James Westwood and the reviewers for their helpful comments. JM's Indiegogo funders in 2014 have seeded this project (pun intended). Some of this work was also funded by National Science Foundation PBI award #2204938 to JM. This paper is dedicated to the memory of eminent Filipino botanist, Leonard Co.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in NCBI SRA under BioProject PRJNA879908.

ORCID

Jeanmaire Molina https://orcid.org/0000-0001-5250-2684
Adhityo Wicaksono https://orcid.org/0000-0003-3298-3577
Ronniel D. Pedales https://orcid.org/0000-0002-5335-6313
Zoé Joly-Lopez https://orcid.org/0000-0002-7926-322X
Danilo Tandang https://orcid.org/0000-0003-2708-661X
So-Yon Park https://orcid.org/0000-0002-2126-3382
Ari Novy https://orcid.org/0000-0002-0873-9948
Susan Pell https://orcid.org/0000-0002-7214-3225
Michael D. Purugganan https://orcid.org/0000-0002-9197-4112

REFERENCES

Arekal, G. D., & Shivamurthy, G. R. (1976). Seed germination in *Balanophora abbreviata*. *Phytomorphology*, *26*(2), 135–138.

Atera, E. A., Itoh, K., & Onyango, J. C. (2011). Evaluation of ecologies and severity of Striga weed on rice in sub-Saharan Africa. *Agriculture and Biology Journal of North America*, *2*(5), 752–760. https://doi.org/10.5251/abjna.2011.2.5.752.760

Bänziger, H. (2004). Studies on hitherto unknown fruits and seeds of some Rafflesiaceae, and a method to manually pollinate their flowers for research and conservation. *Linzer Biologische Beitraege*, *36*, 1175–1198.

Bao, Y. Z., Yao, Z. Q., Cao, X. L., Peng, J. F., Xu, Y., Chen, M. X., & Zhao, S. F. (2017). Transcriptome analysis of *Phelipanche aegyptiaca* seed germination mechanisms stimulated by fluridone, TIS108, and GR24. *PLoS ONE*, *12*(11), e0187539. https://doi.org/10.1371/journal.pone.0187539

Barcelona, J. F., Pelser, P. B., Balete, D. S., & Co, L. L. (2009). Taxonomy, ecology, and conservation status of Philippine *Rafflesia* (Rafflesiaceae). *Blumea*, *54*(1–3), 77–93. https://doi.org/10.3767/000651909x474122

Bascos, E. M. A., Fernando, E. S., Duya, M. V., & Rodriguez, L. J. V. (2021). Beginnings of a plant parasite: Early development of *Rafflesia consueloae* inside its Tetrastigma host. *Planta*, *254*(3), 61. https://doi.org/10.1007/s00425-021-03710-4

Baskin, J. M., & Baskin, C. C. (2022). Germination and seed/embryo size in holoparasitic flowering plants with "dust seeds" and an undifferentiated embryo. *Botanical Review*, *88*(1), 1–49. https://doi.org/10.1007/s12229-020-09242-y

Bleischwitz, M., Albert, M., Fuchsbauer, H.-L., & Kaldenhoff, R. (2010). Significance of Cuscutain, a cysteine protease from *Cuscuta reflexa*, in host-parasite interactions. *BMC Plant Biology*, *10*(1), 227. https://doi.org/10.1186/1471-2229-10-227

Bouman, F., & Meijer, W. (1994). Comparative structure of ovules and seeds in Rafflesiaceae. *Plant Systematics and Evolution*, *193*(1), 187–212. https://doi.org/10.1007/BF00983550

Bravo, A., Brands, M., Wewer, V., Dörmann, P., & Harrison, M. J. (2017). Arbuscular mycorrhiza-specific enzymes FatM and RAM2 fine-tune lipid biosynthesis to promote development of arbuscular mycorrhiza. *New Phytologist*, *214*(4), 1631–1645. https://doi.org/10.1111/nph.14533

Cai, L., Arnold, B. J., Xi, Z., Khost, D. E., Patel, N., Hartmann, C. B., Manickam, S., Sasirat, S., Nikolov, L. A., Mathews, S., Sackton, T. B., & Davis, C. C. (2021). Deeply altered genome architecture in the endoparasitic flowering plant *Sapria himalayana* Griff. (Rafflesiaceae). *Current Biology*, *31*(5), 1002–1011.e9. https://doi.org/10.1016/j.cub.2020.12.045

Cao, X. L., Li, M. L., Li, J., Song, Y. X., Zhang, X. N., Yang, D. L., Li, M., & Wei, J. (2021). Co-expression of hydrolase genes improves seed germination of *Sinopodophyllum hexandrum*. *Industrial Crops and Products*, *164*, 113414. https://doi.org/10.1016/j.indcrop.2021.113414

Carrera-Castano, G., Calleja-Cabrera, J., Pernas, M., Gomez, L., & Onate-Sanchez, L. (2020). An updated overview on the regulation of seed germination. *Plants (Basel)*, *9*(6), 703. https://doi.org/10.3390/plants9060703

Cui, S., Wada, S., Tobimatsu, Y., Takeda, Y., Saucet, S. B., Takano, T., Umezawa, T., Shirasu, K., & Yoshida, S. (2018). Host lignin composition affects haustorium induction in the parasitic plants *Phtheirospermum japonicum* and *Striga hermonthica*. *New Phytologist*, *218*(2), 710–723. https://doi.org/10.1111/nph.15033

Davis, C. C., Latvis, M., Nickrent, D. L., Wurdack, K. J., & Baum, D. A. (2007). Floral gigantism in Rafflesiaceae. *Science*, *315*(5820), 1812. https://doi.org/10.1126/science.1135260

Debeaujon, I., Nesi, N., Perez, P., Devic, M., Grandjean, O., Caboche, M., & Lepiniec, L. (2003). Proanthocyanidin-accumulating cells in Arabidopsis testa: Regulation of differentiation and role in seed development. *Plant Cell*, *15*(11), 2514–2531. https://doi.org/10.1105/tpc.014043

Demarsy, E., Buhr, F., Lambert, E., & Lerbs-Mache, S. (2012). Characterization of the plastid-specific germination and seedling establishment transcriptional programme. *Journal of Experimental Botany*, *63*(2), 925–939. https://doi.org/10.1093/jxb/err322

Dinelli, G., Bonetti, A., & Tibiletti, E. (1993). Photosynthetic and accessory pigments in *Cuscuta campestris* Yuncker and some host species. *Weed Research*, *33*(3), 253–260. https://doi.org/10.1111/j.1365-3180.1993.tb01939.x

Dorone, Y., Boeynaems, S., Jin, B., Bossi, F., Flores, E., Lazarus, E., Michiels, E., De Decker, M., Baatsen, P., Holehouse, A. S., & Sukenik, S. (2020). Hydration-dependent phase separation of a prion-like protein regulates seed germination during water stress. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3680875

Götz, S., Garcia-Gomez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., Robles, M., Talón, M., Dopazo, J., & Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, *36*(10), 3420–3435. https://doi.org/10.1093/nar/gkn176

Haughn, G., & Chaudhury, A. (2005). Genetic analysis of seed coat development in *Arabidopsis*. *Trends in Plant Science*, *10*(10), 472–477. https://doi.org/10.1016/j.tplants.2005.08.005

Hu, H., Xiong, L., & Yang, Y. (2005). Rice *SERK1* gene positively regulates somatic embryogenesis of cultured cell and host defense response against fungal infection. *Planta*, *222*(1), 107–117. https://doi.org/10.1007/s00425-005-1534-4

Jia, L., Wu, Q., Ye, N., Liu, R., Shi, L., Xu, W., Zhi, H., Rahman, A. R. B., Xia, Y., & Zhang, J. (2012). Proanthocyanidins inhibit seed germination by maintaining a high level of abscisic acid in *Arabidopsis thaliana*. *Journal of Integrative Plant Biology*, *54*(9), 663–673. https://doi.org/10.1111/j.1744-7909.2012.01142.x

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research*, *44*, D457–D462. https://doi.org/10.1093/nar/gkv1070

Kawasaki, T., Koita, H., Nakatsubo, T., Hasegawa, K., Wakabayashi, K., Takahashi, H., Umemura, K., Umezawa, T., & Shimamoto, K. (2006). Cinnamoyl-CoA reductase, a key enzyme in lignin biosynthesis, is an effector of small GTPase Rac in defense signaling in rice. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(1), 230–235. https://doi.org/10.1073/pnas.0509875103

Khalaf, E. M., & Raizada, M. N. (2016). Taxonomic and functional diversity of cultured seed associated microbes of the cucurbit family. *BMC Microbiology*, *16*(1), 131. https://doi.org/10.1186/s12866-016-0743-2

Li, S. F., Allen, P. J., Napoli, R. S., Browne, R. G., Pham, H., & Parish, R. W. (2020). MYB-bHLH-TTG1 regulates *Arabidopsis* seed coat biosynthesis pathways directly and indirectly via multiple tiers of transcription factors. *Plant and Cell Physiology*, *61*(5), 1005–1018. https://doi.org/10.1093/pcp/pcaa027

Li, Y. Y., Meng, Z. X., Zhang, Y., Guo, S. X., & Lee, Y. I. (2019). Embryology of *Anoectochilus roxburghii*: Seed and embryo development. *Botanical Studies*, *60*(1), 6. https://doi.org/10.1186/s40529-019-0254-1

Liu, S. S., Chen, J., Li, S. C., Zeng, X., Meng, Z. X., & Guo, S. X. (2015). Comparative transcriptome analysis of genes involved in GA-GID1-DELLA regulatory module in symbiotic and asymbiotic seed germination of *Anoectochilus roxburghii* (Wall.) Lindl. (Orchidaceae). *International Journal of Molecular Sciences*, *16*(12), 30190–30203. https://doi.org/10.3390/ijms161226224

Miray, R., Kazaz, S., To, A., & Baud, S. (2021). Molecular control of oil metabolism in the endosperm of seeds. *International Journal of Molecular Sciences*, *22*(4), 1621. https://doi.org/10.3390/ijms22041621

Molina, J., Hazzouri, K. M., Nickrent, D., Geisler, M., Meyer, R. S., Pentony, M. M., Flowers, J. M., Pelser, P., Barcelona, J., Inovejas, S. A., & Uy, I. (2014). Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). *Molecular Biology and Evolution*, *31*(4), 793–803. https://doi.org/10.1093/molbev/msu051

Molina, J., McLaughlin, W., Wallick, K., Pedales, R., Marius, V. M., Tandang, D. N., Damatac, A., Stuhr, N., Pell, S. K., Lim, T. M., & Novy, A. (2017). Ex situ propagation of Philippine *Rafflesia* in the United States: Challenges and prospects. *Sibbaldia*, *15*, 77–96. https://doi.org/10.24823/Sibbaldia.2017.224

Molina, J., Nikolic, D., Jeevarathanam, J. R., Abzalimov, R., Park, E.-J., Pedales, R., Mojica, E. R. E., Tandang, D., McLaughlin, W., Wallick, K., Adams, J., Novy, A., Pell, S. K., van Breemen, R., & Pezzuto, J. M. (2021). Living with a giant, flowering parasite: Metabolic differences between *Tetrastigma loheri* Gagnep.(Vitaceae) shoots uninfected and infected with *Rafflesia* (Rafflesiaceae) and potential applications for propagation. *Planta*, *255*(1), 4–12. https://doi.org/10.1007/s00425-021-03787-x

Mursidawati, S., Ngatari, N., Irawati, I., Cardinal, S., & Kusumawati, R. (2015). Ex situ conservation of *Rafflesia patma* blume (Rafflesiaceae): An endangered emblematic parasitic species from Indonesia. *Sibbaldia*, *13*, 99–110. https://doi.org/10.24823/Sibbaldia.2015.77

Nais, J. (2001). *Rafflesia of the world*. Sabah Parks.

Nelson, S. K., Ariizumi, T., & Steber, C. M. (2017). Biology in the dry seed: Transcriptome changes associated with dry seed dormancy and dormancy loss in the *Arabidopsis* GA-insensitive sleepy1-2 mutant. *Frontiers in Plant Science*, *8*, 2158. https://doi.org/10.3389/fpls.2017.02158

Ng, S. M., Lee, X. W., Mat-Isa, M. N., Aizat-Juhari, M. A., Adam, J. H., Mohamed, R., Wan, K. L., & Firdaus-Raih, M. (2018). Comparative analysis of nucleus-encoded plastid-targeting proteins in *Rafflesia cantleyi* against photosynthetic and non-photosynthetic representatives reveals orthologous systems with potentially divergent functions. *Scientific Reports*, *8*(1), 17258. https://doi.org/10.1038/s41598-018-35173-1

Nickrent, D. L. (2020). Parasitic angiosperms: How often and how many? *Taxon*, *69*(1), 5–27. https://doi.org/10.1002/tax.12195

Olszewski, M., Dilliott, M., Garcia-Ruiz, I., Bendarvandi, B., & Costea, M. (2020). *Cuscuta* seeds: Diversity and evolution, value for systematics/identification and exploration of allometric relationships. *PLoS ONE*, *15*(6), e0234627. https://doi.org/10.1371/journal.pone.0234627

Parthasarathy, A., Savka, M. A., & Hudson, A. O. (2019). The synthesis and role of beta-alanine in plants. *Frontiers in Plant Science*, *10*, 921. https://doi.org/10.3389/fpls.2019.00921

Pelser, P. B., Nickrent, D. L., & Barcelona, J. F. (2016). Untangling a vine and its parasite: Host specificity of Philippine *Rafflesia* (Rafflesiaceae). *Taxon*, *65*(4), 739–758. https://doi.org/10.12705/654.4

Pelser, P. B., Nickrent, D. L., Callado, J. R. C., & Barcelona, J. F. (2013). Mt. Banahaw reveals: The resurrection and neotypification of the name *Rafflesia lagascae* (Rafflesiaceae) and clues to the dispersal of *Rafflesia* seeds. *Phytotaxa*, *131*(1), 35–40. https://doi.org/10.11646/phytotaxa.131.1.6

Pelser, P. B., Nickrent, D. L., van Ee, B. W., & Barcelona, J. F. (2019). A phylogenetic and biogeographic study of *Rafflesia* (Rafflesiaceae) in the Philippines: Limited dispersal and high island endemism. *Molecular Phylogenetics and Evolution*, *139*, 106555. https://doi.org/10.1016/j.ympev.2019.106555

Pennisi, E. (2014). When is a plant no longer a plant? https://www.science.org/content/article/scienceshot-when-plant-no-longer-plant [Accessed 11 February 2023].

Ranjan, A., Ichihashi, Y., Farhi, M., Zumstein, K., Townsley, B., David-Schwartz, R., & Sinha, N. R. (2014). De novo assembly and characterization of the transcriptome of the parasitic weed dodder identifies genes associated with plant parasitism. *Plant Physiology*, *166*(3), 1186–1199. https://doi.org/10.1104/pp.113.234864

Rehker, J., Lachnit, M., & Kaldenhoff, R. (2012). Molecular convergence of the parasitic plant species *Cuscuta reflexa* and *Phelipanche*

aegyptiaca. *Planta*, *236*(2), 557–566. https://doi.org/10.1007/s00425-012-1626-x

Runo, S., & Kuria, E. K. (2018). Habits of a highly successful cereal killer, Striga. *PLoS Pathogens*, *14*(1), e1006731. https://doi.org/10.1371/journal.ppat.1006731

Santhanam, P., Boshoven, J. C., Salas, O., Bowler, K., Islam, M. T., Saber, M. K., van den Berg, G. C., Bar-Peled, M., & Thomma, B. P. (2017). Rhamnose synthase activity is required for pathogenicity of the vascular wilt fungus *Verticillium dahliae*. *Molecular Plant Pathology*, *18*(3), 347–362. https://doi.org/10.1111/mpp.12401

Seed Identification Guide. (2018). www.idseed.org [Accessed June 1, 2022].

Song, Y., & Zhu, J. (2019). The roles of metabolic pathways in maintaining primary dormancy of *Pinus koraiensis* seeds. *BMC Plant Biology*, *19*(1), 550. https://doi.org/10.1186/s12870-019-2167-2

Su, C., Liu, H., Wafula, E. K., Honaas, L., de Pamphilis, C. W., & Timko, M. P. (2020). SHR4z, a novel decoy effector from the haustorium of the parasitic weed *Striga gesnerioides*, suppresses host plant immunity. *New Phytologist*, *226*(3), 891–908. https://doi.org/10.1111/nph.16351

Teixeira-Costa, L. (2021). A living bridge between two enemies: Haustorium structure and evolution across parasitic flowering plants. *Brazilian Journal of Botany*, *44*, 165–178. https://doi.org/10.1007/s40415-021-00704-0

The Arabidopsis Information Resource (TAIR). (2022). www.arabidopsis.org [Accessed multiple times Jan 2021-Sept 2022].

The UniProt Consortium. (2021). UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Research*, *49*, D480–D489. https://doi.org/10.1093/nar/gkaa1100

Thorogood, C., Witono, J. R., Mursidawati, S., & Fleischmann, A. (2021). Parasitic plant cultivation: Examples, lessons learned and future directions. *Sibbaldia*, *21*, 109–136. https://doi.org/10.24823/Sibbaldia.2022.1892

Tuquet, C., Farineau, N., & Salé, G. (1990). Biochemical composition and photosynthetic activity of chloroplasts from *Striga hermonthica* and *Striga aspera*, root parasites of field-grown cereals. *Physiolia Plantarum*, *78*(4), 574–582. https://doi.org/10.1111/j.1399-3054.1990.tb05244.x

Ueda, M., Fujimoto, M., Arimura, S., Murata, J., Tsutsumi, N., & Kadowaki, K. (2007). Loss of the rpl32 gene from the chloroplast genome and subsequent acquisition of a preexisting transit peptide within the nuclear gene in *Populus*. *Gene*, *402*(1–2), 51–56. https://doi.org/10.1016/j.gene.2007.07.019

Wakabayashi, T., Joseph, B., Yasumoto, S., Akashi, T., Aoki, T., Harada, K., Muranaka, S., Bamba, T., Fukusaki, E., Takeuchi, Y., Yoneyama, K., Muranaka, T., Sugimoto, Y., & Okazawa, A. (2015). Planteose as a storage carbohydrate required for early stage of germination of *Orobanche minor* and its metabolism as a possible target for selective control. *Journal of Experimental Botany*, *66*(11), 3085–3097. https://doi.org/10.1093/jxb/erv116

Wang, B., Wang, Y., & Li, J. (2017). Strigolactones. In J. Li, C. Li, & S. M. Smith (Eds.), *Hormone metabolism and signaling in plants* (pp. 327–359). Elsevier.

Wang, S., Barron, C., Schiefelbein, J., & Chen, J.-G. (2010). Distinct relationships between GLABRA2 and single-repeat R3 MYB transcription factors in the regulation of trichome and root hair patterning in *Arabidopsis*. *New Phytologist*, *185*, 387–400. https://doi.org/10.1111/j.1469-8137.2009.03067.x

Wicaksono, A., Mursidawati, S., & Molina, J. (2021). A plant within a plant: Insights on the development of the *Rafflesia* endophyte within its host. *Botanical Review*, *87*(2), 233–242. https://doi.org/10.1007/s12229-020-09236-w

Wicaksono, A., Mursidawati, S., Sukamto, L. A., & Teixeira da Silva, J. A. (2016). *Rafflesia* spp.: Propagation and conservation. *Planta*, *244*(2), 289–296. https://doi.org/10.1007/s00425-016-2512-8

Xi, Z., Bradley, R. K., Wurdack, K. J., Wong, K., Sugumaran, M., Bomblies, K., Rest, J. S., & Davis, C. C. (2012). Horizontal transfer of expressed genes in a parasitic flowering plant. *BMC Genomics*, *13*(1), 227. https://doi.org/10.1186/1471-2164-13-227

Xi, Z., Wang, Y., Bradley, R. K., Sugumaran, M., Marx, C. J., Rest, J. S., & Davis, C. C. (2013). Massive mitochondrial gene transfer in a parasitic flowering plant clade. *PLoS Genetics*, *9*(2), e1003265. https://doi.org/10.1371/journal.pgen.1003265

Yang, Z., Zhang, Y., Wafula, E. K., Honaas, L. A., Ralph, P. E., Jones, S., Clarke, C. R., Liu, S., Su, C., Zhang, H., Altman, N. S., Schuster, S. C., Timko, M. P., Yoder, J. I., Westwood, J. H., & dePamphilis, C. W. (2016). Horizontal gene transfer is more frequent with increased heterotrophy and contributes to parasite adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(45), E7010–E7019. https://doi.org/10.1073/pnas.1608765113

Yang, Z. Z., Wafula, E. K., Honaas, L. A., Zhang, H. T., das, M., Fernandez-Aparicio, M., Huang, K., Bandaranayake, P. C., Wu, B., der, J. P., Clarke, C. R., Ralph, P. E., Landherr, L., Altman, N. S., Timko, M. P., Yoder, J. I., Westwood, J. H., & dePamphilis, C. W. (2015). Comparative transcriptome analyses reveal core parasitism genes and suggest gene duplication and repurposing as sources of structural novelty. *Molecular Biology and Evolution*, *32*(3), 767–790. https://doi.org/10.1093/molbev/msu343

Yao, Z., Tian, F., Cao, X., Xu, Y., Chen, M., Xiang, B., & Zhao, S. (2016). Global transcriptomic analysis reveals the mechanism of *Phelipanche aegyptiaca* seed germination. *International Journal of Molecular Sciences*, *17*(7), 1139. https://doi.org/10.3390/ijms17071139

Ye, S., Shao, Q., & Zhang, A. (2017). Anoectochilus roxburghii: A review of its phytochemistry, pharmacology, and clinical applications. *Journal of Ethnopharmacology*, *209*, 184–202. https://doi.org/10.1016/j.jep.2017.07.032

Zhao, L., Wang, S., Fu, Y. B., & Wang, H. (2019). *Arabidopsis* seed stored mRNAs are degraded constantly over aging time, as revealed by new quantification methods. *Frontiers in Plant Science*, *10*, 1764. https://doi.org/10.3389/fpls.2019.01764

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Molina, J., Wicaksono, A., Michael, T. P., Kwak, S.-H., Pedales, R. D., Joly-Lopez, Z., Petrus, S., Mamerto, A., Tomek, B., Ahmed, S., Maddu, V., Yakubova, K., Tandang, D., Morin, J. W., Park, S.-Y., Lee, H.-O., McLaughlin, W., Wallick, K., Adams, J., … Purugganan, M. D. (2023). The seed transcriptome of *Rafflesia* reveals horizontal gene transfer and convergent evolution: Implications for conserving the world's largest flower. *Plants, People, Planet*, 1–19. https://doi.org/10.1002/ppp3.10370