Sains Malaysiana 55(10)(2026): 1653-1663
*http://doi.org/10.17576/jsm-2026-5510-05*

# Molecular Insights into Genes and Pathways Involved in the Interaction between *Rafflesia cantleyi* and Its Host *Tetrastigma rafflesiae* Shown by Dual RNA Sequencing

(Pandangan Molekul terhadap Gen dan Laluan yang Terlibat dalam Interaksi antara *Rafflesia cantleyi* dan Perumahnya *Tetrastigma rafflesiae* yang Ditunjukkan melalui Penjujukan Dwi RNA)

ANWARALI-KHAN MURSYIDAH[1], MOHD-NOOR MAT-ISA[1,2], NOR AZLAN NOR MUHAMMAD[3], MOHD FIRDAUS-RAIH[2,3,4] & KIEW-LIAN WAN[1,2,*]

[1]*Department of Biological Sciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia*
[2]*Malaysia Genome and Vaccine Institute, National Institutes of Biotechnology Malaysia, Jalan Bangi, 43000 Kajang, Selangor, Malaysia*
[3]*Institute of Systems Biology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia*
[4]*Department of Applied Physics, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia*

*Received: 20 February 2026/Accepted: 18 September 2026*

*Corresponding author; email: klwan@ukm.edu.my

ABSTRACT

*Rafflesia* is an obligate holoparasitic plant that depends entirely on its host, *Tetrastigma*, for survival. Understanding the molecular interaction between these species is essential for elucidating holoparasitism and supporting conservation strategies. In this study, dual RNA sequencing was performed to characterise the concurrent transcriptomic responses of *Rafflesia cantleyi* and *Tetrastigma rafflesiae* at their interface during floral bud development. Illumina sequencing generated 152.45 million raw reads. Following pre-processing, 24% of the reads were mapped to the *T. rafflesiae* genome, while the remaining reads were inferred to originate from *R. cantleyi*. *De novo* assembly produced 35 255 transcripts for *R. cantleyi* and 15 897 transcripts for *T. rafflesiae*. In *R. cantleyi*, candidate transcripts associated with parasite-host interaction included those encoding ABC transporters, retroelement-associated proteins, aminopeptidase M1, receptor-like kinases, resistance proteins and auxin response factor 19, suggesting involvement in transmembrane transport, genome plasticity, proteostasis, defence-related signalling and hormonal regulation. In *T. rafflesiae,* candidate transcripts associated with parasite-host interaction included those encoding receptor-like kinases, transcription factors and dephospho-CoA kinase, indicating host responses related to perception, signal transduction and metabolic adjustment. Enzyme class profiling and KEGG Orthology and pathway mapping further identified functional pathways associated with transport, signalling, proteostasis, and the plant-pathogen interaction pathway. These findings provide molecular insights into the *Rafflesia-Tetrastigma* interaction at the interface and establish a basis for future functional and conservation-oriented studies.

Keywords: Dual RNA sequencing; holoparasite; host plant; parasite-host interaction; transcriptomics

ABSTRAK

*Rafflesia* ialah tumbuhan holoparasit obligat yang bergantung sepenuhnya kepada perumahnya, *Tetrastigma*, untuk kelangsungan hidup. Pemahaman mengenai interaksi molekul antara kedua-dua spesies ini adalah penting bagi menghuraikan mekanisme holoparasitisme serta menyokong strategi pemuliharaan. Dalam kajian ini, penjujukan dwi RNA telah dijalankan untuk mencirikan tindak balas transkriptomik serentak *Rafflesia cantleyi* dan *Tetrastigma rafflesiae* pada kawasan antara muka semasa perkembangan kudup bunga. Penjujukan menggunakan pelantar Illumina menghasilkan 152.45 juta bacaan mentah. Selepas pra-pemprosesan, sebanyak 24% daripada bacaan dipetakan kepada genom *T. rafflesiae*, manakala selebihnya dianggap merupakan bacaan *R. cantleyi.* Penghimpunan secara *de novo* menghasilkan 35 255 transkrip bagi *R. cantleyi* dan 15 897 transkrip bagi *T. rafflesiae*. Dalam *R. cantleyi*, calon transkrip yang dikaitkan dengan interaksi parasit-perumah termasuk yang mengekod pengangkut ABC, protein berkaitan retroelemen, aminopeptidase M1, kinase mirip reseptor, protein rintangan dan faktor tindak balas auksin 19,

mencadangkan peranan dalam pengangkutan merentas membran, keplastikan genom, proteostasis, pengisyaratan berkaitan pertahanan dan pengawal aturan hormon. Dalam *T. rafflesiae*, calon transkrip yang dikaitkan dengan interaksi parasit-perumah termasuk yang mengekod kinase mirip reseptor, faktor transkripsi dan kinase defosfo-KoA, menunjukkan tindak balas perumah melibatkan persepsi, transduksi isyarat dan pelarasan metabolik. Pemprofilan kelas enzim dan pemetaan Ortologi KEGG serta laluan KEGG selanjutnya mengenal pasti laluan fungsian yang berkaitan dengan pengangkutan, pengisyaratan, proteostasis dan laluan interaksi tumbuhan-patogen. Penemuan ini memberikan pandangan molekul terhadap interaksi *Rafflesia-Tetrastigma* pada kawasan antara muka serta mewujudkan asas untuk kajian fungsian dan berorientasikan pemuliharaan pada masa hadapan.

Kata kunci: Holoparasit; interaksi perumah-parasit; penjujukan dwi RNA; transkriptomik; tumbuhan perumah

## INTRODUCTION

*Rafflesia* is an obligate holoparasitic angiosperm renowned for producing the largest flowers in the world and for its extreme dependence on its host, *Tetrastigma* (Vitaceae), throughout its entire life cycle. Lacking leaves, stems, and functional roots, *Rafflesia* relies exclusively on its host for water, nutrients, and carbon (Nais 2001). This highly reduced morphology and specialised parasitic lifestyle are associated with extensive genomic restructuring, gene loss, and unusual evolutionary trajectories (Mursyidah et al. 2023). As natural populations of *Rafflesia* continue to decline due to habitat destruction and the loss of host plants, understanding the biological basis of host dependency has become increasingly important for conservation and management efforts.

Interactions between parasitic plants and their hosts involve complex molecular crosstalk encompassing nutrient exchange, signal perception, defence activation and counter-defence strategies (Chen et al. 2025; Wu 2018; Yoshida et al. 2016). Host plants often respond to parasitic invasion by activating innate immune pathways similar to those triggered by microbial pathogens, including receptor-mediated perception, kinase-driven signal transduction, transcriptional reprogramming and metabolic adjustment (Kaga et al. 2020; Saucet & Shirasu 2016). Conversely, parasitic plants have evolved mechanisms to suppress, evade, or exploit host defence responses while optimising nutrient acquisition and tissue integration (Johnson & Axtell 2019; Yoshida et al. 2016). Despite growing interest in parasitic plant biology, the molecular mechanisms underpinning these interactions remain poorly understood, particularly for non-model systems such as *Rafflesia*.

Dual RNA sequencing has emerged as a powerful approach for simultaneously profiling transcriptomes of interacting organisms, enabling the investigation of molecular dialogues between hosts and their interacting organisms within the same biological sample (Westermann, Barquist & Vogel 2017; Westermann, Gorski & Vogel 2012). This approach has been successfully applied to plant-pathogen and parasitic plant systems, showing coordinated transcriptional responses and key regulatory pathways underlying interaction outcomes (Fruggiero et al. 2024; Li et al. 2022; Petitot et al. 2016). Applying dual RNA sequencing to the *Rafflesia-Tetrastigma* system therefore provides a unique opportunity to identify genes and biological processes involved in parasitism, host defence, and metabolic integration at the interface.

In this study, we employed dual RNA sequencing to characterise transcriptomic features associated with the interaction between *Rafflesia cantleyi* and *Tetrastigma rafflesiae*. By integrating *de novo* transcriptome assembly and functional annotation, including Gene Ontology (GO) classification, enzyme class profiling, and KEGG pathway mapping, this work aimed to identify candidate genes and functional pathways potentially involved in signalling, defence, transport, and metabolic regulation. The findings contribute to a deeper understanding of holoparasitic plant biology and provide molecular insights relevant to evolutionary studies and conservation strategies for *Rafflesia*.

## MATERIALS AND METHODS

### SAMPLE PREPARATION AND TRANSCRIPTOME SEQUENCING

Fieldwork was carried out in Raub, Pahang, Malaysia with permission from the Pahang State Forestry Department (Permit No.: PHD.R&CH.100-0/28 Jld.6 (57)). Tissue samples were obtained from the interface region between *R. cantleyi* floral buds at the intermediate developmental stage (Amini et al. 2019) and the host, *T. rafflesiae*. The floral buds were sectioned to isolate interface tissues, which were further subdivided into smaller fragments. Samples were surface-sterilised by immersion in 10% (v/v) Clorox solution, followed by three rinses with distilled water. Tissues were immediately cryopreserved in liquid nitrogen during transport and subsequently stored at -80 °C until processing.

Total RNA was extracted using a modified cetyltrimethylammonium bromide (CTAB) protocol (Abdul-Rahman et al. 2017). RNA concentration was measured using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, USA), and RNA integrity was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, USA) prior to library preparation. Transcriptome sequencing was performed on the Illumina HiSeq 4000 PE150 platform.

### DUAL RNA SEQUENCING READ PREPROCESSING AND *de novo* ASSEMBLY

Dual RNA sequencing reads were quality-filtered using SolexaQA (v3.1.5) (Cox, Peterson & Biggs 2010). Adaptor sequences were removed and reads with a Phred quality score below 20 or lengths shorter than 50 bp were discarded. High-quality reads were aligned using Bowtie2 (v2.4.2) (Langmead & Salzberg 2012) against the PLAZA 5.1 plant database (Van Bel et al. 2022) and reads lacking matches to plant sequences were excluded. The remaining reads were mapped to the *T. rafflesiae* genome using Bowtie2. Reads that failed to map to the *T. rafflesiae* genome were considered to be of *R. cantleyi* origin. Clean reads were then assembled *de novo* using Trinity (v2.15.2) (Grabherr et al. 2011).

### GENOME SEQUENCING AND *de novo* ASSEMBLY OF *Tetrastigma rafflesiae*

To facilitate separation of host- and parasite-derived transcripts, a draft genome assembly of *T. rafflesiae* was generated and used as the reference for host read identification. Genomic DNA was extracted from mature leaves of *T. rafflesiae*. Fresh leaf tissues were ground into a fine powder using a pre-chilled mortar and pestle under liquid nitrogen, and genomic DNA was isolated using the CTAB method described by Doyle and Doyle (1990). DNA quality and concentration were evaluated prior to library construction.

Whole-genome sequencing libraries comprising one paired-end library with an insert size of approximately 700 bp and four mate-pair libraries with insert sizes of 2, 4, 6, and 8 kbp were prepared according to the manufacturer's protocols (Illumina Inc., San Diego, CA, USA). Sequencing was performed on the Illumina HiSeq 2000 platform. Raw genomic reads were quality-filtered using SolexaQA++ (v3.1.5) (Cox, Peterson & Biggs 2010). Adapter sequences and low-quality bases were removed, and reads with Phred quality scores below 20 or lengths shorter than 50 bp were discarded. The resulting high-quality reads were *de novo* assembled using Velvet (version 1.2.10) (Zerbino & Birney 2008). The draft genome assembly was used solely as an internal reference for partitioning host-derived reads from the dual RNA sequencing dataset prior to transcriptome assembly. Owing to the large genome size, abundant repetitive sequences and probable polyploid nature of *T. rafflesiae*, the genome assembly remains under refinement and is not intended as the primary outcome of the present study.

### SEQUENCE DATA AVAILABILITY

The raw dual RNA transcriptome and *T. rafflesiae* genome sequencing data generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject PRJNA1494198. The dual RNA sequencing libraries are available under BioSample accessions SAMN61596609-SAMN61596610 with SRA Run accessions SRR39592366-SRR39592367, whereas the draft *T. rafflesiae* genome sequencing datasets are available under BioSample accession SAMN61596611, with SRA Run accessions SRR39592361-SRR39592365.

### HOMOLOGY SEARCH AND FUNCTIONAL GO ANNOTATION

The initial phase of gene annotation involved the identification of homologous sequences in public databases. Assembled transcripts were aligned against the NCBI non-redundant (NR) GenBank protein database using DIAMOND BLASTX, which translates nucleotide sequences and searches against protein databases. An E-value threshold of <1e-6 was applied to retain significant matches. GO annotation was performed using Blast2GO (Conesa et al. 2005). Transcript sequences were first associated with homologous protein sequences in the NR database based on the DIAMOND BLASTX results, and GO terms were assigned according to the accession numbers of the best-supported BLAST hits. Annotation was refined by selecting the most specific GO terms for each unigene based on BLAST score calculations. In addition, protein domain information obtained from InterProScan was integrated to support and complement GO assignments. GO annotation enabled the classification of transcripts into three main categories: Molecular Function, Biological Process, and Cellular Component. This functional categorisation facilitated a comprehensive overview of the transcriptomic functional landscape and enabled comparison among transcript subsets. GO classification summaries were visualised using the WEGO web application (Ye et al. 2006). Transcripts potentially associated with the parasite-host interaction were identified through manual curation based on functional annotations and relevant literature.

### ENZYME CLASS PROFILING AND KO-KEGG PATHWAY MAPPING

Enzyme class profiling and KEGG Orthology (KO)-KEGG pathway mapping (Kanehisa et al. 2023, 2016) were performed to summarise and contextualise functional signatures in *R. cantleyi* and *T. rafflesiae* transcriptomes. Enzyme Commission (EC) numbers were retrieved from the homology-based functional annotation outputs, including DIAMOND BLASTX searches against the NCBI NR database, followed by GO assignment using Blast2GO with InterProScan-supported evidence. For each organism, enzyme class terms were extracted from the annotation tables and aggregated by counting the number of transcripts associated with each EC number, with each transcript contributing a single count per enzyme class assignment. Enzyme class frequencies were subsequently normalised to relative proportions (%)

based on the total number of enzyme-class-annotated transcripts per organism. Dominant enzyme classes were ranked to define organism-specific enzymatic signatures for the parasite and host.

For KO-KEGG analysis, KO identifiers were assigned to annotated transcripts based on the best-supported homology evidence, linking high-confidence protein matches to their corresponding KO entries where available. Assignments were cross-validated against functional descriptions and conserved domain evidence to ensure consistency. KO identifiers were then mapped to KEGG pathways and BRITE hierarchies to generate organism-specific pathway profiles. BRITE is a hierarchical functional classification system that groups KEGG objects (commonly KO groups) into higher-level biological categories. Comparative analyses were conducted to identify shared and complementary functional modules associated with the parasite-host interface, with particular emphasis on pathways consistently supported by both enzyme class profiles and candidate gene annotations. These included ABC transport, the plant-pathogen interaction pathway, MAPK signalling, plant hormone signal transduction and proteostasis-related pathways. Finally, KO-KEGG pathway information was integrated with enzyme class profiling results to interpret transcript-level signals as mechanistic, pathway-linked modules underpinning parasite-host interaction at the interface site.

The KO-KEGG analysis performed in this study was intended as a functional pathway-mapping approach rather than a statistical pathway-enrichment analysis. Accordingly, pathway representation was determined from KO assignments of annotated transcripts, and no over-representation test or pathway-level significance testing was performed. The aim was to identify functional pathways represented at the *R. cantleyi* - *T. rafflesiae* interface and to relate these pathways to annotated transcripts potentially involved in parasite-host interaction.

## RESULTS

### DUAL RNA SEQUENCING AND *de novo* TRANSCRIPTOME ASSEMBLY

Dual RNA sequencing of *R. cantleyi* - *T. rafflesiae* interface tissues generated 152.45 million raw paired-end reads. Following quality filtering and adaptor trimming, high-quality reads were mapped to the *T. rafflesiae* genome using Bowtie2. A total of 24% (31.8 million) high-quality reads aligned successfully to the *T. rafflesiae* genome, while the remaining reads were retained and considered to be of *R. cantleyi* origin. *De novo* assembly of the separated read sets yielded 15 897 transcripts for *T. rafflesiae* and 35 255 transcripts for *R. cantleyi*. Summary statistics of transcript assembly are presented in Table 1.

### GENE ONTOLOGY CLASSIFICATION

Overall annotation rates reflected the non-model status of both *R. cantleyi* and its host *T. rafflesiae*. GO annotation profiles showed distinct yet complementary functional signatures at the interface of these two organisms. For *R. cantleyi,* the GO distributions indicate that transcripts are predominantly associated with catalytic, metabolic and transcription-associated functions (Figure 1). The Molecular Function category was dominated by catalytic activity (62.2%), with high representation of transferase activity (26.6%), and hydrolase activity (23.3%). More specific terms included catalytic activity, acting on a protein (18.5%) and RNA binding (11.7%). Overall, the Molecular Function distribution indicates strong representation of enzymatic and nucleic-acid-associated functions. In the Biological Process category, transcripts were predominantly associated with regulatory and transport/metabolic functions. The most abundant term was regulation of DNA-templated transcription (10.7%), followed by transmembrane transport (9.2%). Metabolic processes were also strongly represented, including carbohydrate metabolic process (6.8%) and lipid metabolic process (5.9%). Genome- and RNA-related terms such as chromatin organisation (5.5%), mRNA metabolic process (4.2%), and DNA repair (4.0%) were notable, together with vesicle-mediated transport (4.0%), intracellular protein transport (3.5%) and catabolic process (3.5%). Collectively, the Biological Process profile indicates representation of transcriptional regulation, intracellular trafficking, and metabolic adjustment. In Cellular Component, transcripts were primarily associated with organelles (40.0%) and the nucleus (21.3%), with substantial representation in the plasma membrane (8.2%), and cytosol (6.5%). Additional localisation to mitochondrion (4.1%) and plastid (4.0%) was observed. Lower but consistent proportions were associated with the Golgi apparatus (1.9%) and cytoplasmic vesicle (1.7%), supporting the representation of trafficking-related processes.

For *T. rafflesiae*, transcripts were predominantly represented by terms related to ATP binding, nucleic-acid binding, kinase activity, and membrane localisation (Figure 2). For Molecular Function, ATP binding (16.6%) was the most abundant term, followed by zinc ion binding (9.6%) and RNA binding (5.6%). Notably, protein serine/threonine kinase activity (3.3%) was present, alongside mRNA binding (3.0%), DNA-binding transcription factor activity (1.5%), and structural constituent of ribosome (1.3%). Within Biological Process, the most represented term was signal transduction (1.7%), followed by response to stress (0.2%). Additional terms related to regulation and development were present at lower proportions. In Cellular Component, the dominant term was membrane (24.1%), followed by nucleus (14.8%) and cytoplasm (10.1%). Plasma membrane (7.5%) and cytosol (5.2%) were also strongly represented. Organellar terms included

TABLE 1. Transcriptome assembly statistics

| | *R. cantleyi* | *T. rafflesiae* |
|---|---|---|
| Total transcripts | 35 255 | 15 897 |
| Median transcript length | 404 | 302 |
| Mean transcript length (bp) | 535.36 | 368.54 |
| Total assembled bases | 18 873 959 | 5 858 753 |
| GC content (%) | 44.93 | 43.29 |
| N50 | 646 | 377 |

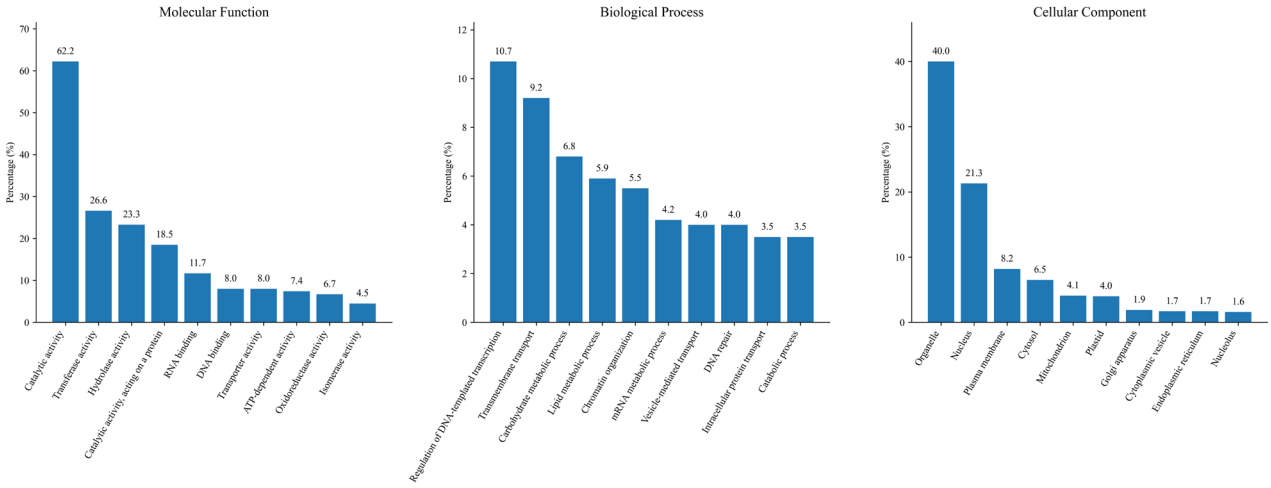



FIGURE 1. GO annotation of transcripts in *Rafflesia cantleyi*. The y-axis indicates the percentage of transcripts assigned to each GO term, while the x-axis shows the corresponding GO categories

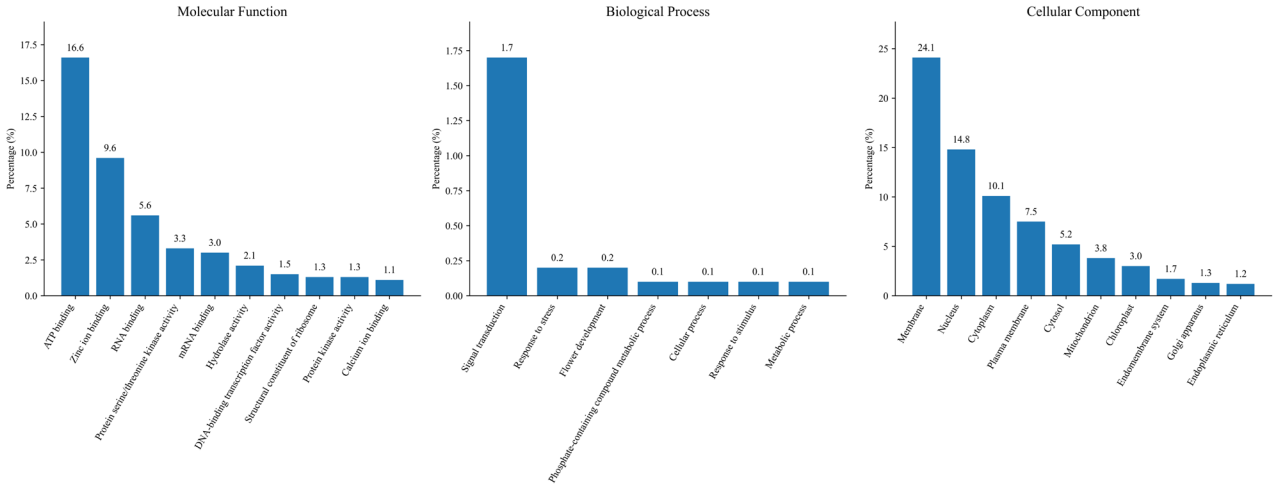



FIGURE 2. GO annotation of transcripts in *Tetrastigma rafflesiae*. The y-axis indicates the percentage of transcripts assigned to each GO term, while the x-axis shows the corresponding GO categories

mitochondrion (3.8%) and chloroplast (3.0%), together with lower-level representation of endomembrane-associated compartments.

### TRANSCRIPTS ASSOCIATED WITH THE PARASITE-HOST INTERACTION

Several transcripts potentially associated with the parasite-host interaction were identified in *R. cantleyi* (Table 2), including those encoding ABC transporters, retroelement-associated proteins, aminopeptidase M1, receptor-like kinases (RLKs), disease resistance proteins and auxin response factor 19 (ARF19). These transcripts are associated with transmembrane transport, proteolysis, signal perception, stress-related responses, and hormonal regulation. Notably, the presence of retroelement-related transcripts is consistent with the representation of mobile genetic element-related functions, while the identification of transporters and kinases indicates the presence of transcripts associated with transport and signalling at the parasite-host interface.

In *T. rafflesiae*, interaction-associated transcripts included Gag-protease-integrase-RT-RNase H polyprotein, multiple classes of receptor-like kinases (including leucine-rich repeat RLKs, Nod factor perception proteins, MIK2 and cysteine-rich RLKs), NAC, and NF-Y transcription factors, and dephospho-CoA kinase. These genes are linked to signal perception, transcriptional regulation, transport processes and metabolic functions, with the diversity of RLKs indicating representation of multiple receptor-mediated signalling functions at the interface site.

### KEGG PATHWAY ANALYSIS

KEGG pathway mapping identified multiple pathways represented in both organisms. In *R. cantleyi*, represented pathways included ABC transporters, the plant-pathogen interaction pathway, MAPK signalling, plant hormone signal transduction, proteostasis-related pathways, and genetic information processing, including retroelement-related processes. In *T. rafflesiae*, KEGG pathways were predominantly associated with signal transduction, phosphate metabolism, the plant-pathogen interaction pathway and metabolic regulation. Shared pathway representation suggests overlapping functional processes during the interaction, whereas organism-specific pathways indicate functional differences between the parasite and host (Table 3).

### DOMINANT ENZYME CLASS SIGNATURES

For *R. cantleyi*, the most frequently represented enzyme class was RNA-directed DNA polymerase (Enzyme Commission (EC) 2.7.7.49), consistent with the representation of reverse-transcription- and retroelement-related functions. Aspartic endopeptidase activity (EC 3.4.23) and phosphorylation-related enzyme class categories were also prominent. In *T. rafflesiae*, alkaline phosphatase (EC 3.1.3.1) was the most abundant enzyme class, with serine/threonine protein kinase activity (EC 2.7.11.1) also strongly represented, consistent with the representation of kinase-associated signalling functions (Figure 3).

### KO-KEGG MAPPING

KO-KEGG mapping provided an integrated framework linking candidate genes and interaction-relevant pathways. In *R. cantleyi*, the ABC transporter transcript mapped to ko02010 (ABC transporters), supporting its potential role in metabolite exchange and efflux at the interface. Retroelement-associated transcripts mapped to viral-like genetic information processing pathways within the KEGG BRITE hierarchy, consistent with the representation of mobile-element-related functions. Aminopeptidase M1 aligned with proteostasis and protein-processing pathways (ko04141, ko04120, ko03050), while the S-domain receptor-like kinase and TMV N resistance protein mapped to immune-related perception and signalling pathways (ko04626, ko04016). ARF19 mapped to ko04075 (plant hormone signal transduction), linking auxin signalling to transcriptional regulation in parasite tissues. Together, these mappings integrate transport, genome plasticity, proteostasis, signalling and hormone regulation as functional features potentially associated with *R. cantleyi* adaptation to host association.

In *T. rafflesiae*, KO-KEGG mapping highlighted host responses encompassing perception, signalling, transcriptional regulation and metabolic support. Multiple RLKs, including lectin receptor-like kinases, LRR-RLKs, CRK29, NFP, and MIK2, mapped to ko04626 (the plant-pathogen interaction pathway) and ko04016 (MAPK signalling), supporting the representation of receptor-mediated perception and kinase-associated signalling. NAC and NF-Y transcription factors were associated with hormone and defence signalling pathways (ko04075, ko04016), while dephospho-CoA kinase mapped to ko00770 (pantothenate and CoA biosynthesis), indicating a potential role in metabolic processes associated with sustained defence responses. Detection of a Gag-protease-integrase-RT-RNase H polyprotein further indicated the presence of transcripts associated with retroelement activity. Collectively, these results indicated a host response involving surface perception, signal amplification, transcriptional regulation, transport and metabolic adjustment during parasitic invasion.

## DISCUSSION

This study provides an integrated transcriptomic perspective on the molecular interaction between *R. cantleyi* and its host *T. rafflesiae*. By applying a dual RNA sequencing strategy, transcripts from both organisms were analysed simultaneously, enabling a

TABLE 2. Candidate transcripts associated with the *Rafflesia cantleyi* - *Tetrastigma rafflesiae* interaction

| Organism | GO ID | GO Name | No. Transcripts | Transcript ID | Gene |
|---|---|---|---|---|---|
| *R. cantleyi* | GO:0006508 | Proteolysis | 719 | TRINITY_DN29126_c0_g1_i1 | ABC transporter protein |
| | GO:0015074 | DNA integration | 69 | TRINITY_DN1228_c0_g1_i1 | Ty3-I transposon Gag-Pol polyprotein |
| | GO:0015074 | DNA integration | 69 | TRINITY_DN1228_c0_g1_i18 | Retrovirus-related Pol polyprotein (reverse transcriptase/integrase) |
| | GO:0015074 | DNA integration | 69 | TRINITY_DN1228_c0_g1_i25 | Retrovirus-related Pol polyprotein (reverse transcriptase/integrase) |
| | GO:0006278 | RNA-templated DNA biosynthetic process | 33 | TRINITY_DN6297_c0_g1_i3 | Reverse transcriptase |
| | GO:0034976 | Response to endoplasmic reticulum stress | 32 | TRINITY_DN1599_c0_g1_i1 | Protein disulfide isomerase-like 2-3 |
| | GO:0006508 | Proteolysis | 719 | TRINITY_DN379_c0_g1_i1 | Aminopeptidase M1 (proteolysis/peptide processing) |
| | GO:0009742 | Brassinosteroid mediated signaling pathway | 39 | TRINITY_DN9935_c0_g1_i1 | S receptor-like serine/ threonine protein kinase |
| | GO:0008289 | Lipid binding | 117 | TRINITY_DN13344_c0_g1_i1 | TMV N resistance protein |
| | GO:0006355 | Regulation of DNA-templated transcription | 1,379 | TRINITY_DN14613_c0_g1_i1 | Auxin response factor 19 (ARF19) |
| *T. rafflesiae* | GO:0015074 | DNA integration | 93 | TRINITY_DN215_c0_g1_i14 | Gag-protease-integrase-RT-RNase H polyprotein |
| | GO:0006338 | Chromatin remodeling | 484 | TRINITY_DN2679_c0_g1_i2 | Lectin G-type S receptor-like serine/ threonine protein kinase (SD2-5) |
| | GO:0004672 | Protein kinase activity | 199 | TRINITY_DN8862_c0_g1_i1 | Cysteine-rich receptor-like kinase 29 |
| | GO:0001682 | tRNA 5'-leader removal | 7 | TRINITY_DN9786_c0_g1_i1 | ABC transporter family C member 9 (ABCC9) |
| | GO:0004672 | Protein kinase activity | 199 | TRINITY_DN12129_c0_g1_i1 | NFP receptor-like serine/threonine kinase |
| | GO:0006952 | Defense response | 149 | TRINITY_DN9005_c0_g1_i1 | LRR receptor-like serine/threonine protein kinase |
| | GO:0004672 | Protein kinase activity | 199 | TRINITY_DN717_c0_g1_i2 | MDIS1-interacting receptor-like kinase 2 (MIK2) |
| | GO:0006355 | Regulation of DNA-templated transcription | 548 | TRINITY_DN5812_c0_g1_i1 | NAC transcription factor 035 |
| | GO:0006357 | Regulation of transcription by RNA polymerase II | 297 | TRINITY_DN628_c0_g1_i1 | Nuclear transcription factor Y subunit (NF-Y subunit) |
| | GO:0015937 | Coenzyme A biosynthetic process | 10 | TRINITY_DN2705_c0_g1_i1 | Dephospho-CoA kinase |

TABLE 3. KEGG pathway mapping of transcripts from *Rafflesia cantleyi* and *Tetrastigma rafflesiae*

| Organism | Pathway ID | Pathway name | No. KOs | No. transcripts |
|---|---|---|---|---|
| *R. cantleyi* | ko02010 | ABC transporters | 8 | 114 |
| | ko03250 | Viral life cycle - HIV-1 (BRITE: Viral-like genetic information processing) | 16 | 50 |
| | ko04141 | Protein processing in endoplasmic reticulum | 75 | 311 |
| | ko04120 | Ubiquitin mediated proteolysis | 58 | 278 |
| | ko03050 | Proteasome | 35 | 96 |
| | ko04626 | Plant-pathogen interaction | 30 | 191 |
| | ko04016 | MAPK signaling pathway - plant | 34 | 156 |
| | ko04075 | Plant hormone signal transduction | 52 | 370 |
| *T. rafflesiae* | ko04626 | Plant-pathogen interaction | 6 | 8 |
| | ko04016 | MAPK signaling pathway - plant | 9 | 10 |
| | ko04075 | Plant hormone signal transduction | 9 | 18 |
| | ko00770 | Pantothenate and CoA biosynthesis | 1 | 1 |
| | ko03250 | Viral life cycle - HIV-1 (BRITE: Viral-like genetic information processing) | 5 | 7 |

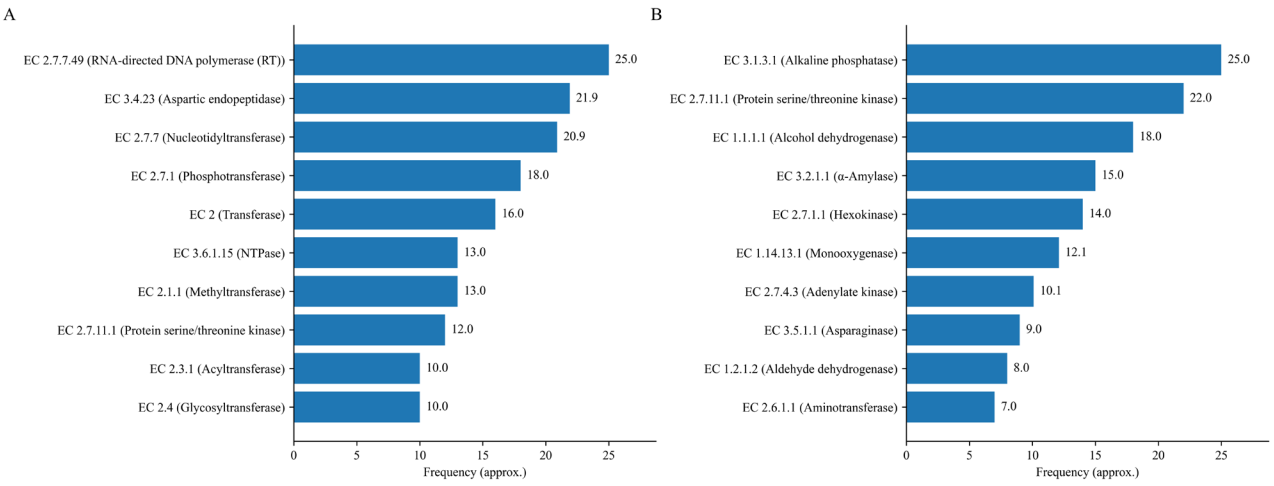



FIGURE 3. Enzyme class profiling of *Rafflesia cantleyi* and *Tetrastigma rafflesiae*. Enzyme Commission (EC) codes for *Rafflesia cantleyi* (A) and *Tetrastigma rafflesiae* (B) are presented on the y-axis, with their corresponding frequencies of occurrence shown on the x-axis

comprehensive characterisation of molecular activities at the parasite-host interface. This approach captures both parasite-associated processes to resource acquisition and adaptation, as well as host regulatory and defence responses, offering a more holistic view of holoparasitic interaction than single-organism analysis. Unlike previous transcriptomic studies in parasitic plant systems that focused on either the host or parasite separately or on bulk tissue-level expression without specific resolution of the interface, this study captures simultaneous bidirectional transcriptional activity at the physical parasite-host interface. This provides an opportunity to infer potential molecular exchange and localised regulatory responses that may otherwise be obscured in single-organism or whole-tissue datasets.

The KO-KEGG framework was employed as a unifying analytical framework to link transcript annotations to biological pathways, thereby providing a consistent basis for interpreting molecular functions within the interface region. In *R. cantleyi*, candidate genes - including an ABC transporter, retroelement- and RT-integrase-associated modules (Ty3-I-like Gag-Pol

polyproteins and reverse transcriptase), aminopeptidase M1, receptor-like kinases (RLK), TMV N and ARF19 - collectively point to functional themes of transmembrane trafficking, genomic plasticity, proteolysis and peptide processing, defence-related signalling and hormonal regulation associated with interface tissue development. Importantly, these gene categories should not be interpreted independently but rather as functional modules that may contribute to localised interface specialisation. The representation of transport, signalling and proteolysis-related genes suggests that the interface region may represent a metabolically active zone associated with molecular exchange, rather than simply a passive attachment site.

This interpretation is further supported by the dominant enzyme class profile in *R. cantleyi*, in which RNA-directed DNA polymerase (EC 2.7.7.49) was the most abundant enzyme category, consistent with the representation of reverse-transcription- and retrotransposon-associated functions. Such representation is consistent with previous reports that retrotransposons can be induced under stress conditions and contribute to genome remodelling and adaptive plasticity (Eickbush & Jamburuthugoda 2008; Lisch 2013). In the context of parasitism, retroelement-associated activity may contribute to genomic plasticity and potentially facilitate adaptation to host-derived environmental pressures, including immune responses and nutrient flux variability. This raises the possibility that genome restructuring may have contributed to the long-term parasitic specialisation of *Rafflesia* species. In Rafflesiaceae, extensive horizontal gene transfer from *Tetrastigma* to *Rafflesia* has also been documented, further supporting the view that genome dynamism has contributed to the evolutionary diversification of parasitic traits (Xi et al. 2013, 2012). In addition, the prominence of proteolytic enzymes such as aspartic endopeptidase (EC 3.4.23) suggests that proteolysis may contribute to parasite development and the parasite-host interaction, potentially through signalling peptide maturation, tissue modification, or modulation of host-derived defence components (van der Hoorn 2008).

In *T. rafflesiae*, the dominant enzyme class profile indicates representation of biochemical functions potentially associated with responses to parasitic invasion. Alkaline phosphatase (EC 3.1.3.1), the most frequent enzyme detected, is associated with phosphate metabolism, energy homeostasis, and cellular signalling and regulation during biotic stress (Nussaume & Kanno 2024). These findings suggest that *T. rafflesiae* may actively respond to the presence of the parasite through perception and signal transduction rather than functioning solely as a nutrient source. Such kinase-mediated signalling may represent components of the molecular dialogue between host and parasite, potentially contributing to downstream transcriptional reprogramming and defence modulation. This was followed by serine/threonine protein kinase activity (EC 2.7.11.1), a central component of signal transduction pathways commonly activated during host defence responses (Bigeard, Colcombet & Hirt 2015). Together, these profiles are consistent with host tissues exhibiting signalling and metabolic functions associated with the parasite-host interaction, rather than simply passive resource depletion.

A key strength of this study lies in positioning KO-KEGG mapping as an integrative analytical framework rather than a purely descriptive catalogue of pathways. Instead of simply listing pathways, this approach provides a framework for interpreting the *Rafflesia-Tetrastigma* interaction as a network of potentially related processes involving transport, signalling, transcriptional regulation and physiological adaptation at the interface site. In *R. cantleyi*, this network is characterised by ABC transport (ko02010), potentially supporting metabolite transport and flux, retroelement-associated modules within the KEGG BRITE viral-like genetic information processing hierarchy, consistent with the representation of retroelement-related functions, proteostasis mediated by protein processing and targeted proteolysis (ko04141, ko04120, ko03050), perception and defence-related signalling (ko04626, ko04016), and hormone-mediated regulation via auxin signalling (ko04075), which may contribute to transcriptional regulation and interface tissue formation.

Retroelement-related transcripts are interpreted mechanistically, with the presence of Gag-Pol, reverse transcriptase and integrase domains representing characteristic features of LTR retrotransposons and retrovirus-like elements, explaining their placement within KEGG viral-like information processing categories (Kanehisa et al. 2021; Sabot & Schulman 2006; Vicient 2025). In parallel, the involvement of proteostasis pathways, including the unfolded protein response in the endoplasmic reticulum, is biologically plausible, as elevated production of secretory proteins such as receptors and defence-related factors can impose ER stress, necessitating homeostatic regulation during sustained biotic interactions (Bao & Howell 2017; Howell 2013; Liu & Howell 2010).

The present study was designed primarily as an exploratory functional characterisation of the *R. cantleyi* - *T. rafflesiae* interface, with represented genes and pathways identified based on functional annotation and literature-based manual curation rather than differential expression or statistical enrichment analyses. The principal aim was to identify transcripts derived from the parasite and host at their physical interface and to determine the biological functions and pathways represented by these transcripts. Similarly, the KO-KEGG analysis was used to map annotated transcripts to functional pathways rather than to test for statistical over-representation against a predefined background gene set. Consequently, the pathway profiles reported

here should be interpreted as evidence of functional representation and as a basis for generating mechanistic hypotheses, rather than as statistically significant pathway enrichment. This exploratory approach is particularly useful for the poorly characterised *Rafflesia-Tetrastigma* system because it establishes candidate genes and molecular processes that can subsequently be tested using targeted expression analyses, additional biological comparisons and functional validation. Collectively, these findings support the view that the *Rafflesia-Tetrastigma* interface represents a specialised functional zone characterised by interacting metabolic, signalling and regulatory processes from both organisms.

## CONCLUSIONS

This study provides a dual RNA sequencing-based characterisation of the molecular interaction between *Rafflesia* and its host *Tetrastigma*. The results show a diverse set of transcripts and functional pathways associated with signalling, proteolysis, transport and metabolic regulation. The detection of conserved defence-related pathways, including RLK–MAPK signalling, is consistent with host response to parasitic interaction, while *Rafflesia* expresses genes potentially associated with host interaction, nutrient acquisition and adaptive integration at the interface. Collectively, these findings advance our understanding of the molecular basis of holoparasitism and parasite-host crosstalk and provide a framework for future studies focused on functional validation, evolutionary dynamics and conservation of this remarkable parasitic plant system.

## ACKNOWLEDGEMENTS

This research was supported by the Ministry of Higher Education Malaysia under the Fundamental Research Grant Scheme (FRGS/1/2024/WAS10/UKM/01/1).

## REFERENCES

Abdul-Rahman, A., Suleman, N.I., Zakaria, W.A., Goh, H.H., Mohd Noor, N.M. & Aizat, W.M. 2017. RNA extractions of mangosteen (*Garcinia mangostana* L.) pericarps for sequencing. *Sains Malaysiana* 46(8): 1231-1240.

Amini, S., Rosli, K., Abu-Bakar, M.F., Alias, H., Mat-Isa, M.N., Juhari, M.A.A., Haji-Adam, J., Goh, H.H. & Wan, K.L. 2019. Transcriptome landscape of *Rafflesia cantleyi* floral buds reveals insights into the roles of transcription factors and phytohormones in flower development. *PLoS ONE* 14(12): e0226338.

Bao, Y. & Howell, S.H. 2017. The unfolded protein response supports plant development and defense as well as responses to abiotic stress. *Frontiers in Plant Science* 8: 344.

Bigeard, J., Colcombet, J. & Hirt, H. 2015. Signaling mechanisms in pattern-triggered immunity (PTI). *Molecular Plant* 8(4): 521-539.

Chen, Y., Wu, H., Cai, J., Guo, S., Gan, X., Liu, X. & Yang, J. 2025. Molecular translocation between parasitic plants and their hosts. *Frontiers in Plant Science* 16: 1716304.

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

Cox, M.P., Peterson, D.A. & Biggs, P.J. 2010. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics* 11: 485.

Doyle, J.J. & Doyle, J.L. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

Eickbush, T.H. & Jamburuthugoda, V.K. 2008. The diversity of retrotransposons and the properties of their reverse transcriptases. *Virus Research* 134(1-2): 221-234.

Fruggiero, C., Aufiero, G., D'Angelo, D., Pasolli, E. & D'Agostino, N. 2024. Refining dual RNA-seq mapping: Sequential and combined approaches in host-parasitic plant dynamics. *Frontiers in Plant Science* 15: 1483717.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. & Regev, A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29(7): 644-652.

Howell, S.H. 2013. Endoplasmic reticulum stress responses in plants. *Annual Review of Plant Biology* 64: 477-499.

Johnson, N.R. & Axtell, M.J. 2019. Small RNA warfare: Exploring origins and function of trans-species microRNAs from the parasitic plant *Cuscuta*. *Current Opinion in Plant Biology* 50: 76-81.

Kaga, Y., Yokoyama, R., Sano, R., Ohtani, M., Demura, T., Kuroha, T., Shinohara, N. & Nishitani, K. 2020. Interspecific signaling between the parasitic plant and the host plants regulate xylem vessel cell differentiation in haustoria of *Cuscuta campestris*. *Frontiers in Plant Science* 11: 193.

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. 2023. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research* 51(D1): D587-D592.

Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M. & Tanabe, M. 2021. KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Research* 49(D1): D545-D551.

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(D1): D457-D462.

Langmead, B. & Salzberg, S.L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9(4): 357-359.

Li, H., Hu, R., Fan, Z., Chen, Q., Jiang, Y., Huang, W. & Tao, X. 2022. Dual RNA sequencing reveals the genome-wide expression profiles during the compatible and incompatible interactions between *Solanum tuberosum* and *Phytophthora infestans*. *Frontiers in Plant Science* 13: 817199.

Liu, J.X. & Howell, S.H. 2010. bZIP28 and NF-Y transcription factors are activated by ER stress and assemble into a transcriptional complex to regulate stress response genes in Arabidopsis. *The Plant Cell* 22(3): 782-796.

Lisch, D. 2013. How important are transposons for plant evolution? *Nature Reviews Genetics* 14(1): 49-61.

Mursyidah, A.-K., Hafizzudin-Fedeli, M., Nor Muhammad, N.A., Latiff, A., Firdaus-Raih, M. & Wan, K-L. 2023. Dissecting the biology of *Rafflesia* species: Current progress and future directions made possible with high-throughput sequencing data. *Plant and Cell Physiology* 64(4): 368-377.

Nais, J. 2001. *Rafflesia of the World*. Malaysia: Natural History Publications (Borneo).

Nussaume, L. & Kanno, S. 2024. Reviewing impacts of biotic and abiotic stresses on the regulation of phosphate homeostasis in plants. *Journal of Plant Research* 137(3): 297-306.

Petitot, A-S., Dereeper, A., Agbessi, M., Da Silva, C., Guy, J., Ardisson, M. & Fernandez, D. 2016. Dual RNA-seq reveals *Meloidogyne graminicola* transcriptome and candidate effectors during the interaction with rice plants. *Molecular Plant Pathology* 17(6): 860-874.

Sabot, F. & Schulman, A.H. 2006. Parasitism and the retrotransposon life cycle in plants: A hitchhiker's guide to the genome. *Heredity* 97(6): 381-388.

Saucet, S.B. & Shirasu, K. 2016. Molecular parasitic plant-host interactions. *PLoS Pathogens* 12(12): e1005978.

Van Bel, M., Silvestri, F., Weitz, E.M., Kreft, L., Botzki, A., Coppens, F. & Vandepoele, K. 2022. PLAZA 5.0: Extending the scope and power of comparative and functional genomics in plants. *Nucleic Acids Research* 50(D1): D1468-D1474.

van der Hoorn, R.A. 2008. Plant proteases: From phenotypes to molecular mechanisms. *Annual Review of Plant Biology* 59: 191-223.

Vicient, C.M. 2025. Retand LTR-retrotransposons in plants: A long way from *pol* to 3'LTR. *Mobile DNA* 16(1): 15.

Westermann, A.J., Barquist, L. & Vogel, J. 2017. Resolving host-pathogen interactions by dual RNA-seq. *PLoS Pathogens* 13(2): e1006033.

Westermann, A.J., Gorski, S.A. & Vogel, J. 2012. Dual RNA-seq of pathogen and host. *Nature Reviews Microbiology* 10(9): 618-630.

Wu, J. 2018. miRNAs as a secret weapon in the battlefield of haustoria, the interface between parasites and host plants. *Molecular Plant* 11(3): 354-356.

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.

Xi, Z., Bradley, R.K., Wurdack, K.J., Wong, K.M., Sugumaran, M., Bomblies, K., Rest, J.S. & Davis, C.C. 2012. Horizontal transfer of expressed genes in a parasitic flowering plant. *BMC Genomics* 13: 227.

Ye, J., Fang, L., Zheng, H., Zhang, Y., Chen, J., Zhang, Z., Wang, J., Li, S., Li, R., Bolund, L. & Wang, J. 2006. WEGO: A web tool for plotting GO annotations. *Nucleic Acids Research* 34: W293-W297.

Yoshida, S., Cui, S., Ichihashi, Y. & Shirasu, K. 2016. The haustorium, a specialized invasive organ in parasitic plants. *Annual Review of Plant Biology* 67: 643-667.

Zerbino, D.R. & Birney, E. 2008. Velvet: Algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Research* 18(5): 821-829.