



Delayed leaf greening involves a major shift in the expression of cytosolic and mitochondrial ribosomes to plastid ribosomes in the highly phosphorus-use-efficient *Hakea prostrata* (Proteaceae)

Toby Bird · Benjamin J. Nestor · Philipp E. Bayer · Guannan Wang · Albina Ilyasova · Clément E. Gille · Bryce E. H. Soraru · Kosala Ranathunge · Anita A. Severn-Ellis · Ricarda Jost · Wolf-Rüdiger Scheible · Maheshi Dassanayake · Jacqueline Batley · David Edwards · Hans Lambers · Patrick M. Finnegan

Received: 21 April 2023 / Accepted: 31 August 2023 / Published online: 27 September 2023
© The Author(s) 2023

Abstract

Background and aims *Hakea prostrata* (Proteaceae) is a highly phosphorus-use-efficient plant native to southwest Australia. It maintains a high photosynthetic rate at low leaf phosphorus (P) and exhibits delayed leaf greening, a convergent adaptation that increases nutrient-use efficiency. This study aimed to

provide broad physiological and gene expression profiles across leaf development, uncovering pathways leading from young leaves as nutrient sinks to mature leaves as low-nutrient, energy-transducing sources.

Methods To explore gene expression underlying delayed greening, we analysed a de novo transcriptome for *H. prostrata* across five stages of leaf development. Photosynthesis and respiration rates, and foliar pigment, P and nitrogen (N) concentrations were determined, including the division of P into five biochemical fractions.

Key results Transcripts encoding functions associated with leaf structure generally decreased in abundance across leaf development, concomitant with decreases in foliar concentrations of 85% for

Responsible Editor: Mian Gu.

Toby Bird and Benjamin J. Nestor are co-first authors of this article.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11104-023-06275-1>.

T. Bird (✉) · B. J. Nestor (✉) · P. E. Bayer · A. Ilyasova · C. E. Gille · B. E. H. Soraru · K. Ranathunge · A. A. Severn-Ellis · R. Jost · J. Batley · D. Edwards · H. Lambers · P. M. Finnegan
School of Biological Sciences, University of Western Australia, Perth, WA 6009, Australia
e-mail: toby.bird@research.uwa.edu.au

B. J. Nestor
e-mail: benjamin.nestor@research.uwa.edu.au

P. E. Bayer
e-mail: philipp.bayer@uwa.edu.au

A. Ilyasova
e-mail: albina.ilyasova6012@gmail.com

C. E. Gille
e-mail: clement.gille@research.uwa.edu.au

B. E. H. Soraru
e-mail: bsoraru@gmail.com

K. Ranathunge
e-mail: kosala.ranathunge@uwa.edu.au

A. A. Severn-Ellis
e-mail: anita.severn-ellis@dpird.wa.gov.au

R. Jost
e-mail: R.Jost@latrobe.edu.au

J. Batley
e-mail: jacqueline.batley@uwa.edu.au

D. Edwards
e-mail: dave.edwards@uwa.edu.au

H. Lambers
e-mail: hans.lambers@uwa.edu.au

anthocyanins, 90% for P and 70% for N. The expression of genes associated with photosynthetic function increased during or after leaf expansion, in parallel with increases in photosynthetic pigments and activity, much later in leaf development than in species that do not have delayed greening. As leaves developed, transcript abundance for cytosolic and mitochondrial ribosomal proteins generally declined, whilst transcripts for chloroplast ribosomal proteins increased.

Conclusions There was a much longer temporal separation of leaf cell growth from chloroplast development in *H. prostrata* than is found in species that lack delayed greening. Transcriptome-guided analysis of leaf development in *H. prostrata* provided insight into delayed greening as a nutrient-saving strategy in severely phosphorus-impooverished landscapes.

Keywords *Hakea prostrata* · Phosphorus · Leaf development · Delayed greening · Photosynthesis · Ribosomes

Introduction

Southwest Australia has severely nutrient-impooverished soils (Turner and Laliberté 2015). These soils are particularly low in soluble inorganic phosphate (Pi), which is the only form of phosphorus (P) available to plants and is essential for plant growth (Lambers 2022). The region is also a

biodiversity hotspot with extremely high levels of plant species richness and endemism (Hopper and Gioia 2004). A prominent plant family on the most P-impooverished soils of southwest Australia is the Proteaceae (Lambers et al. 2015b). Proteaceae are generally considered to be non-mycorrhizal. Instead, several traits that contribute to high P-use efficiency (PUE) have evolved in the Proteaceae that allow them to thrive in soils that are too low in P for mycorrhizal associations to be effective (Lambers et al. 2015a). *Hakea prostrata* is a member of the Proteaceae and is a model Australian species for high nutrient-use-efficiency studies because of its high PUE and widespread distribution in southwest Australia. It is an endemic, prostrate or upright non-lignotuberous shrub that grows 1 to 3 m high and is typically found on sandy soils, laterite, loam, or gravel in woodlands, heath, granite outcrops or coastal dunes (Fig. 1) (Western Australian Herbarium 1998). Like many Proteaceae, *H. prostrata* has a very low P concentration in mature leaves with an average of about 0.3 mg P g⁻¹ dry weight in plants in their natural habitat (Lambers et al. 2012; Sulpice et al. 2014; Wright et al. 2004). This value is among the lowest 20% of mature leaf P concentrations recorded in a dataset encompassing 750 plants worldwide (Wright et al. 2004). The nitrogen (N) concentration in mature leaves of *H. prostrata* is also low at an average of 12.7 mg P g⁻¹ dry weight (Sulpice et al. 2014) compared to a range of 8.7 to 40.9 mg P g⁻¹ dry weight in C₃ plant species (Reich et al. 1997).

P. M. Finnegan
e-mail: patrick.finnegan@uwa.edu.au

B. J. Nestor · P. E. Bayer · D. Edwards · P. M. Finnegan
Centre for Applied Bioinformatics, University of Western Australia, Perth, WA 6009, Australia
e-mail: philipp.bayer@uwa.edu.au · dave.edwards@uwa.edu.au · patrick.finnegan@uwa.edu.au

Present Address:
P. E. Bayer
Minderoo Foundation, Forrest Hall 2, 21 Hackett Dr., Crawley, WA 6009, Australia

G. Wang · M. Dassanayake
Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA
e-mail: gwang23@lsu.edu

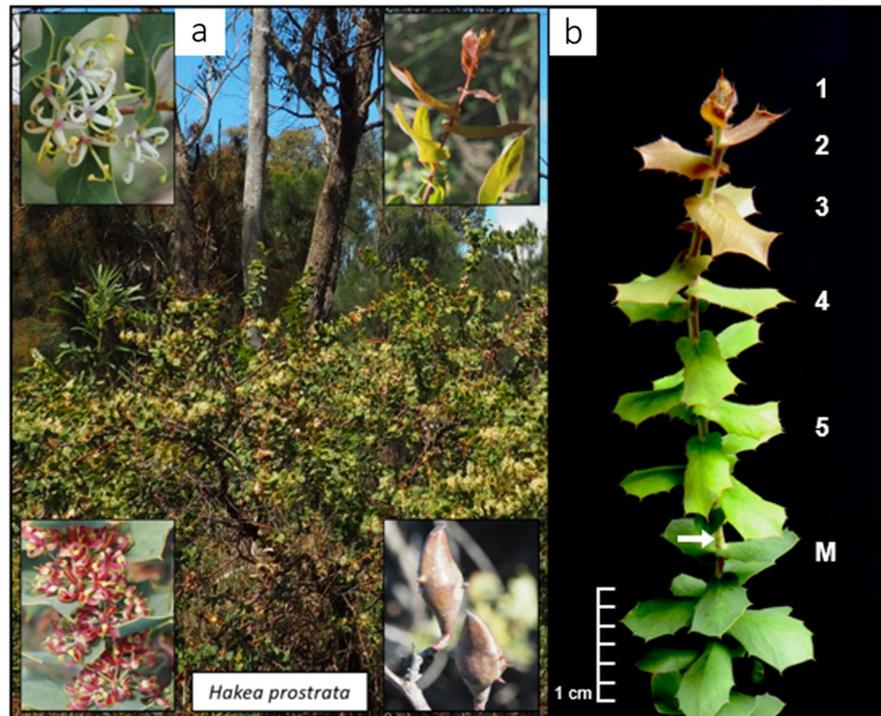
M. Dassanayake
e-mail: maheshid@lsu.edu

Present Address:
A. A. Severn-Ellis
Aquatic Animal Health, Department of Primary Industries and Development, Indian Ocean Marine Research Centre, 86 West Coast Dr., Watermans Bay, WA 6020, Australia

Present Address:
R. Jost
Department of Animal, Plant and Soil Sciences, School of Life Sciences, La Trobe Institute for Agriculture and Food, La Trobe University and Australian Research Council Research HUB for Medicinal Agriculture, La Trobe University, Bundoora, VIC 3086, Australia

W.-R. Scheible
Noble Research Institute, Ardmore, OK, USA
e-mail: tomales@gmx.de

Fig. 1 *Hakea prostrata* growing in its natural habitat in southwest Australia. **a)** Flowers coloured white-cream (upper left) or red (lower left) appear between July and October (Western Australian Herbarium 1998). The species has delayed greening of young leaves (upper right) and produces small fruits (lower right) that bear two seeds. **b)** A sprig of new growth showing delayed greening of the newly emerged leaves. Immature leaf stages 1 to 5 above and Mature (M) leaves below the terminal bud scale scar (arrow) are indicated



Hakea prostrata shares many of the high PUE traits found in other Proteaceae. It forms cluster roots, which are dense clusters of short determinate lateral roots that release carboxylates that mobilise soil-sorbed Pi and greatly aid P acquisition (Shane and Lambers 2005). Furthermore, *H. prostrata* has fast rates of photosynthesis despite its very low leaf P concentration (Lambers et al. 2012, 2015b). These fast photosynthetic rates at low leaf P concentrations are due in part to the preferential allocation of P to photosynthetically-active mesophyll cells, low levels of P-rich rRNA and the replacement of phospholipids with other lipids during leaf development (Hayes et al. 2018; Lambers et al. 2012; Sulpice et al. 2014).

Delayed leaf greening is another PUE trait exhibited by *H. prostrata* and other Proteaceae (Kuppusamy et al. 2014, 2021; Lambers et al. 2012). The delayed leaf greening trait is a highly convergent adaptation found in many plant families, including shade-tolerant species and plants growing in full sun (Kursar and Coley 1992b; Lambers et al. 2012; Sulpice et al. 2014). Delayed greening of leaves in these species coincides with delayed photosynthetic competence (Kursar and

Coley 1992a; Lambers et al. 2012). Young leaves of plants with delayed greening, including *H. prostrata*, are often pale yellow or reddish-brown due to high concentrations of anthocyanin pigments coupled with a low concentration of chlorophyll (Kuppusamy et al. 2021; Miyazawa et al. 2003; Sulpice et al. 2014). Chlorophyll concentrations and photosynthetic competence do not increase until after full leaf expansion (Kuppusamy et al. 2021; Kursar and Coley 1992a; Lambers et al. 2012). Thus, nutrient use in *H. prostrata*, particularly of P and N, is prioritised first to developing leaf structure in the youngest leaf stages before being re-prioritised to photosynthesis in maturing leaf stages (Kuppusamy et al. 2021; Sulpice et al. 2014). The spreading of nutrient use across leaf development is thought to minimise the risk of losing valuable nutrients in young leaves to herbivory (Kursar and Coley 1992c). In the case of *H. prostrata* and other plants that grow in highly P-impo- verished soils, the trait is also a strategy for the efficient use of scarce P (Kuppusamy et al. 2021; Lambers et al. 2015a). Amongst plant species with delayed leaf greening, including *H. prostrata*, a high flavonoid abundance is also likely to protect

young leaves against herbivory, fungal infections and cellular damage from high light intensities (Close and Beadle 2003; Hughes et al. 2007; Kuppusamy et al. 2021; Kursar and Coley 1992a, c; Woodall and Stewart 1998).

A deeper analysis of delayed leaf greening in *H. prostrata* has shown that in addition to high levels of anthocyanins and low levels of chlorophylls, young leaves have low levels of carotenoids and carbohydrates (Kuppusamy et al. 2021; Sulpice et al. 2014). Furthermore, photosynthetic enzymes such as Rubisco only increase in abundance when the leaf is fully expanded (Sulpice et al. 2014). As leaves develop from young to mature, the total P concentration declines strongly (Kuppusamy et al. 2021; Lambers et al. 2012). The inorganic P concentration remains fairly constant, and nearly all of the decline is from P that is esterified into organic forms (Kuppusamy et al. 2021). In mature leaves, the concentrations of esterified and inorganic P are similar. The predominant form of esterified P in leaves is generally nucleic acids such as rRNA (Veneklaas et al. 2012). In several species of *Banksia* (Proteaceae) and *Hakea*, the population of rRNAs in leaves shifts from being dominated by cytosolic rRNA in young leaves to being either equalised with or slightly dominated by plastid rRNA in mature leaves (Sulpice et al. 2014). This shift in rRNA population is thought to underlie a change in ribosome population to support chloroplast biogenesis later in leaf development after much of the cytosolic protein needed for the production of leaf structure has been synthesised (Lambers et al. 2015b). However, the dominance of cytosolic rRNA in young leaves was not observed in *H. prostrata* (Sulpice et al. 2014), presumably because the shift occurred at an earlier stage of leaf development than was sampled in the study.

Current knowledge on the mechanisms underlying delayed greening is supplied by physiological and biochemical studies (Kuppusamy et al. 2021; Kursar and Coley 1992b; Lambers et al. 2012; Miyazawa et al. 2003; Sulpice et al. 2014; Vogado et al. 2022). In the Proteaceae, the study by Kuppusamy et al. (2014) began to uncover gene networks and expression patterns involved in lipid remodelling during delayed greening. Developing these concepts further, we hypothesised that the expression of *H.*

prostrata genes involved in the development of photosynthetic functions such as chlorophyll biogenesis and carbon fixation would increase in later stages of leaf development as the expression of genes involved in cell division and expansion declined. We also hypothesised that the expression of genes encoding cytosolic ribosomal proteins would decrease over leaf development, while those encoding plastid ribosomal proteins would increase, consistent with the pattern of rRNA abundance found in the majority of Proteaceae examined (Sulpice et al. 2014). To address these hypotheses, we constructed a reference transcriptome for *H. prostrata* and used differential gene expression analysis to examine gene expression patterns across five stages of leaf development that were clearly distinguished by differences in leaf morphology. The resulting analysis and transcriptome resource provides valuable insight into delayed greening as a nutrient-saving strategy in *H. prostrata* and likely other Proteaceae that exhibit delayed greening.

Materials and methods

Sample collection from plants grown in hydroponics

Hakea prostrata plants were grown in a glasshouse at the University of Western Australia, Perth, Australia to provide hydroponically-grown white root, cluster root, and mature leaf samples for the *H. prostrata* reference transcriptome. Six-month-old soil-grown *H. prostrata* plants propagated from cuttings of a single mature plant (Nuts about Natives, Karnup, Western Australia) were transferred to hydroponics supplemented with 1, 10 or 50 μM inorganic phosphate (Pi) with five individuals per Pi treatment, as described by Kuppusamy et al. (2014). Conditions of growth included average minimum and maximum temperatures of 14°C (night) and 22°C (day), with an average maximum light intensity of 820 $\text{mmol m}^{-2} \text{s}^{-1}$ (70% transmission) occurring between 12 noon and 2 pm. The relative humidity varied from 47% (day) to 66% (night). Cluster roots were harvested from plants grown at 1 μM Pi approximately 1, 4, 7 and 12 days after emergence. These four stages ranged in development from just emerging to mature (Shane et al. 2003). All samples were snap-frozen in liquid nitrogen (N_2) and stored at -80°C.

Collection of leaves from wild plants

Leaves were harvested from five naturally-occurring *H. prostrata* plants at the UWA Field Station, Shenton Park, Western Australia (31.9496 S, 115.7973 E) to provide different leaf stage samples for the *H. prostrata* reference transcriptome and replicated RNA libraries for differential expression analysis across leaf development. Leaves of the same stage from three newly-emerged flushes on a single plant were pooled together to form a replicate and snap-frozen in liquid N₂. Five replicates of each leaf stage were harvested from each of the five plants. Immature leaves were designated stage 1 (2% to 10% fully-expanded, soft and reddish), stage 2 (20% to 60% fully-expanded, soft and reddish), stage 3 (60% to 90% fully-expanded, soft and greenish-reddish in colour) and stage 4 (fully-expanded, stiff but somewhat malleable and bright green (Fig. 1b). Mature leaves were the youngest of the fully-expanded hard, green leaves in the flush immediately preceding the flush harvested for immature leaf stages (Fig. 1b). Frozen leaves were ground under liquid N₂ and stored at -80°C until analysis.

Photosynthesis, dark respiration and pigment determination

Net CO₂ assimilation at 1800 μmol quanta m⁻² s⁻¹ and dark respiration in the absence of light after 1 h of dark acclimation were measured on at least four intact, attached leaves from each replicate plant using a portable gas-exchange system (LI-6400 with LI-6400-02B red-blue light source, LI-COR, Lincoln, NE, USA) at ambient CO₂ concentration (400 μmol mol⁻¹) and temperature (20°C to 25°C). Leaves similar to those used for gas exchange were sampled, snap frozen in liquid N₂ and ground under liquid N₂ using a mortar and pestle. To determine chlorophylls *a* and *b* and carotenoid concentrations, approx. 20 mg ground leaf material was extracted by adding 2 mL 100% methanol, mixing and incubating at 4°C overnight in darkness. The concentrations of extracted pigments were determined spectrophotometrically (Multiscan Spectrum Plate Reader, Thermo Scientific, Scoresby, Australia) using the wavelengths and equations described in literature (Wellburn 1994). For determining leaf anthocyanin concentration, 50 mg of the leaf material ground

under liquid N₂ was mixed with 500 μL 1% (v/v) acetic acid and homogenised (Precellys 24 Tissue Disruptor, Bertin Technologies, Montigny-le-Bretonneux, France) for three cycles of 30 s each at 4500 rpm. The supernatant was cleared by centrifugation at 21,000 g for 15 min (twice) and anthocyanin concentrations determined using a modified pH-differential method (Wrolstad et al. 2005). Spectrophotometric absorbance was measured (Multiscan Spectrum Plate Reader) and concentrations were calculated using the molar absorptivity of the cyanidin-3-glucoside ($\epsilon = 26,900 \text{ L mol}^{-1} \text{ cm}^{-1}$) in the sample (Giusti and Wrolstad 2001). Cyanidin-3-glucoside was chosen, since it is the most common anthocyanin in plants (Francis and Markakis 1989).

Determination of leaf N, P and P fractions

Leaf material was harvested into liquid N₂ and stored at -80°C until freeze-dried for 7 days (VirTis Benchtop 'K'; New York, NY, USA). Freeze-dried material was ground to a fine powder using a zirconium ball mill (Geno/Grinder; Spex SamplePrep, Metuchen, NJ, USA). Total leaf N was determined in a 30 mg subsample by combustion (Vario Macro Combustion Analyser; Elementar Analysensysteme GmbH, Langensfeld, Germany). A 150 mg subsample was used to determine total P after the conversion of esterified P to Pi by acid digestion (Matusiewicz and Golik 2004). A final 50 mg subsample was used to determine P present in nucleic acid, phospholipid, metabolic P (small metabolites+Pi) and residual P fractions as described (Yan et al. 2019).

RNA extraction

Frozen leaf powder (100 mg) was mixed with 700 μL of 100 mM Tris-HCl, pH 8, 500 mM LiCl, 10 mM EDTA, 1% (w/v) lithium dodecyl sulfate, 5 mM dithiothreitol, 0.5% (w/v) polyvinylpyrrolidone, and 0.5% (w/v) polyvinylpolypyrrolidone. Following the addition of 350 μL of Fruit-mate (Takara Biotechnology, Kusatsu, Shiga, Japan), samples were vortexed well. After several chloroform:isoamyl alcohol (24:1) extractions, RNA was concentrated by isopropanol precipitation, followed by another precipitation of the resuspended RNA from 2 M LiCl. The RNA pellet was washed with 2 M LiCl followed by two washes with 70% (v/v) ethanol, air dried, and resuspended

in 20 μL of RNase-free water. After the removal of contaminating DNA (RQ1 DNase; Promega Corporation, Madison, WI, USA), the RNA concentration was determined by fluorometry (Qubit 3, Invitrogen, Carlsbad, CA, USA). RNA quality was assessed using spectrophotometry (NanoDrop 1000 Thermo Fisher Scientific Waltham, MA, USA) and capillary electrophoresis (LabChip GX Touch 24, HT RNA Reagents, 24 DNA5K/RNA/CZE Chip HT DNA, PerkinElmer, Waltham, MA, USA).

Transcriptome sequencing for plants grown in hydroponics

The three highest-quality RNA preparations, based on A260/A280 ratio and RNA integrity number (RIN), for each analogous organ or cluster root developmental stage were pooled. This produced three pools for white roots (one pool for each level of Pi during growth), three pools for mature leaves (one pool for each level of Pi during growth), and five pools for cluster roots (one pool for each developmental stage 0, 1, 4, 7 and 12 days post-emergence) to give a total of 11 RNA pools. Each pool was sequenced to give 100 bp paired-end reads (Illumina HiSeq 2000, Illumina, San Diego, CA, USA) by LGC Genomics GmbH (Berlin, Germany), leading to 11 RNAseq libraries each containing between 33.3 M and 63.0 M reads (Table S1).

Transcriptome sequencing for plants grown in their natural habitat

Only the leaves of four of the five plants were used for RNA preparation, and stage 5 leaves were omitted from the transcript analysis because they were physiologically very similar to stage 4 leaves. Three out of four RNA replicates from each leaf developmental stage (one to mature) that had the highest quality based on A260/A280 ratio and RIN value was selected for individual library preparation and sequencing. Libraries for RNAseq were prepared from 1 μg RNA according to the manufacturer's instructions (TruSeq Stranded Total RNA Kit with Ribo-Zero Plant, Illumina, San Diego, CA, USA). The libraries were quantified by fluorometry (Qubit 3, Invitrogen, Carlsbad, CA, USA), and then quality assessed (LabChip GX Touch 24 with HT DNA HiSens Dual Protocol Reagents using a 24 DNA

Extended Range Chip, PerkinElmer, Waltham, MA, USA). Libraries were sequenced as 150 bp paired-end reads (Illumina HiSeq 3000/4000, Illumina, San Diego, CA, USA) by the Australian Genome Research Facility (Melbourne, VIC, Australia), leading to 15 libraries each containing between 12.4 M and 55.2 M reads (Table S1).

De novo transcriptome assembly and annotation

All RNAseq read libraries from leaf and root samples of hydroponically-grown plants and the leaf developmental series of plants in their natural habitat were combined to construct a reference de novo transcriptome assembly for *H. prostrata*. Illumina RNAseq paired-end reads were analysed using FastQC (v0.11.9, <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) for quality control. Reads with erroneous kmers were corrected using rCorrector (v1.0.4, Song and Florea 2015) and low-quality bases and adapters were trimmed with Trim-galore! (v0.6.6, https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Trimmed reads were assembled in Trinity (v2.12.0, Grabherr et al. 2011). The draft assembly was assessed by mapping RNAseq reads back to the transcriptome using HISAT2 (v2.2.1, Kim et al. 2019). Contigs were searched by Basic Local Alignment Search Tool (BLAST, blastn v2.6.0+, Altschul et al. 1990) with the SILVA database for rRNA (138.1 LSU and SSU NR99, Quast et al. 2012), and *Arabidopsis* chloroplast and mitochondrial genome sequences (NC_000932.1, NC_037304.1) to identify rRNA, chloroplast and mitochondrial DNA sequences. Contigs with e-values less than $1e^{-10}$ were deemed rRNA, chloroplast, or mitochondrial sequences and removed from the transcriptome (Geng et al. 2021).

Putative peptide sequences encoded by the assembled transcripts were predicted using TransDecoder (v5.5.0, Haas et al. 2013). Predicted peptide sequences were then searched for sequence similarity to known protein sequences using BLAST (blastp v2.6.0+) with the Swiss-Prot database (v2022_02, Figueira et al. 2015) and HMMER (hmmscan v3.3.2, Eddy 2011) with the Pfam database (v34.0, El-Gebali et al. 2019), only keeping hits that had e-values less than $1e^{-5}$. Contaminant non-plant sequences were identified and removed by searching against the National Center

for Biotechnology Information (NCBI) Non-Redundant (NR) database (Sayers et al. 2021) with DIAMOND BLAST (v2.0.9, Buchfink et al. 2021). Only sequences where the top hit was an Embryophyta (land plants) sequence were retained. Due to the large number of *H. prostrata* individuals used to generate the transcriptome, representative peptide sequences were extracted by clustering sequences using CDHIT (v4.8.1, Fu et al. 2012) with a 95% identity threshold. The extracted peptide sequences were annotated using the Trinotate (v3.2.2, Grabherr et al. 2011) pipeline including BLAST sequence hits (blastp and blastx), Pfam domains (HMMER hmmscan v2.3), signal peptide predictions (SignalP v4.1, Petersen et al. 2011) and transmembrane domains (TMHMM v2.0c, Krogh et al. 2001). Peptide sequences were also annotated with gene ontology (GO) terms using both blastp and Interproscan results (Blast2GO basic, Conesa and Götz 2008). KEGG orthology (KO) identifiers were obtained for pathway analysis (EggNOG-mapper v2, <http://eggnog5.embl.de>, Huerta-Cepas et al. 2019). A reference transcriptome for expression quantification was generated by extracting transcripts corresponding to the peptide sequences. Transcriptome and peptide sequences were assessed for completeness after each filtering step using BUSCO (v5.1.2, Manni et al. 2021) with the embryophyta_odb10 transcriptome or protein databases for the embryophyta_odb10 lineage. The final set of selected annotated non-redundant sequences will be referred to as transcripts in the reference transcriptome from here on.

Differential expression and co-expression analysis of transcripts

Transcript expression abundance was determined and quality checked before differential expression analysis. The expression of transcripts in the reference transcriptome was quantified using Kallisto (v0.46.2, Bray et al. 2016) through the `align_and_estimate_abundance.pl` script in Trinity (v2.12.0, Grabherr et al. 2011). Transcript abundance among organ samples and among replicates from the leaf developmental series were compared by constructing violin plots of transcript abundance distributions for representative samples with ggplot2 (Wickham 2016) using log normalised transcripts per million (TPM) estimates. Transcript profiles were compared using a principal

component analysis (PCA) graph generated with the PtR script in Trinity (v2.12.0, Grabherr et al. 2011) using counts per million estimates.

Differential expression of transcripts among the leaf developmental stages was tested using DESeq2 (ver. 1.30.1, Love et al. 2014) within R Studio (R studio version 1.4.1106, R version 4.0.4, R Core Team 2022; RStudio Team 2021). Raw transcript counts were normalised using regularised logarithm transformation (rlog) and differential abundance was tested using Wald test with default parameters (Love et al. 2014). Contrasts among leaf developmental stages were established to view any pairwise significance. Gene transcripts with a False Discovery Rate (FDR) adjusted *P*-value (< 0.05) and a fold-difference ≥ 2 (\log_2 FoldChange ≤ -1 or ≥ 1) were considered to represent differentially expressed genes (DEGs). Co-expression profiles for the DEGs were identified and assigned using hierarchical clustering of the mean rlog-transformed transcript counts (Perseus, v1.6.15.0, Tyanova et al. 2016) after normalising the matrix with z-scores. Sub-cellular location of genes encoding ribosomal subunit proteins was predicted (TargetP – v2.0, Armenteros et al. 2019). Fisher's exact test with Benjamini–Hochberg adjusted *P*-values (< 0.05) was used to calculate enrichment for up-regulated and down-regulated ribosomal subunit transcripts in the assigned categorical data Location (cellular) and Group (co-expression) (Tyanova et al. 2016).

Gene ontology term, KEGG pathway enrichment and pigment biosynthesis analyses

The current version of the GO OBO file was downloaded (format version 1.2, release/22–03–2022, Ashburner et al. 2000; Consortium 2021) and used for GO term over-representation analysis of the reference *H. prostrata* transcriptome (BiNGO version 3.0.3, Maere et al. 2005) on an open-source platform (Cytoscape version 3.8.2, Shannon et al. 2003). The default Hypergeometric test and Benjamini and Hochberg *P*-value adjustment method were used. Gene ontology terms for co-expressed DEGs were consolidated into representative clusters (GOMCL, Wang et al. 2020). For primary clustering, the default settings were used including the similarity index of the Overlap coefficient. The `-gsize` parameter was set to 2000 to exclude very

high-level GO terms. Similarity score distributions were retrieved to identify sub-clustering potential. Primary clusters that had a similarity score distribution value of less than 50% ($P > 0.5$) were analysed by sub-clustering. The $-g$ size parameter was reduced to 500 for sub-cluster analysis to reveal finer detail. Over-representation analysis of KEGG orthology (KO) IDs assigned to groups of DEGs in comparison to transcriptome-wide annotations was done for KEGG pathway enrichment (clusterProfiler v4.66 R, Wu et al. 2021). Pathways enriched for DEGs were subsequently filtered to remove non-plant pathways (KEGGREST, Tenenbaum and Maintainer 2022). For transcriptomic analyses of pigment biosynthesis, the GO terms of GO:0009718 (anthocyanin-containing compound

biosynthetic process), GO:0015995 (chlorophyll biosynthetic process) and GO:0016117 (carotenoid biosynthetic process) were used to retrieve annotated sequences. The TPM values for the retrieved functions were then used to create ridge plots to map gene expression density across leaf development (R Core Team 2022; RStudio Team 2021).

Results

Photosynthesis and pigment dynamics across leaf development

Photosynthetic carbon assimilation activity increased progressively in developing leaves of *H. prostrata*,

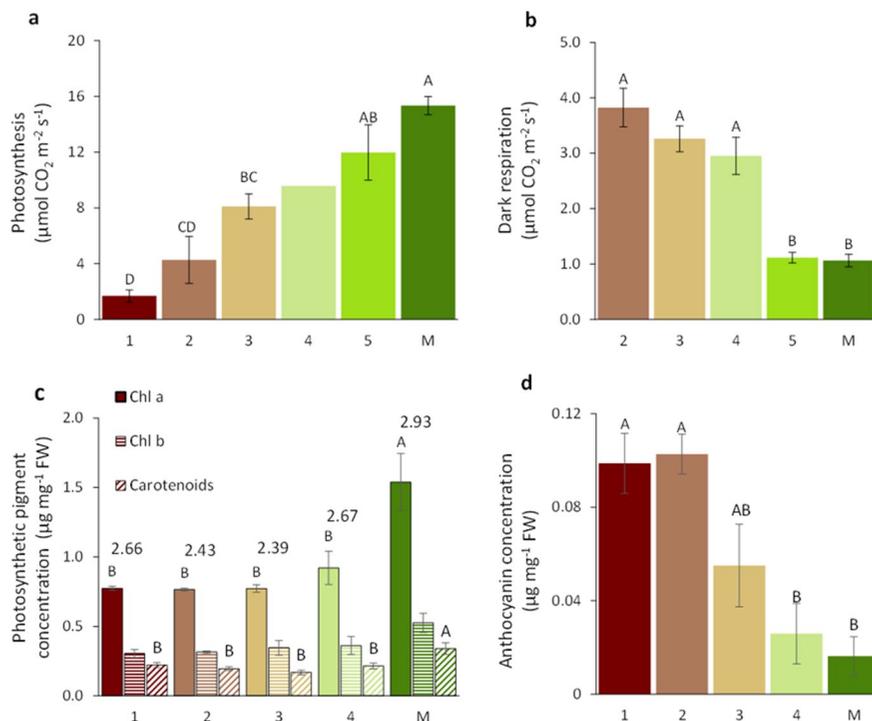


Fig. 2 Leaf photosynthesis and dark respiration activities and pigment concentrations across leaf development in *Hakea prostrata*. **a**) Photosynthesis (gross CO_2 assimilation). **b**) Dark respiration (gross CO_2 release by dark acclimated leaves). **c**) Foliar chlorophyll a (Chl a), chlorophyll b (Chl b) and carotenoid concentrations. The chlorophyll a : chlorophyll b ratio for each stage of leaf development is shown over the bars, and **d**) Foliar anthocyanin concentrations. Leaf development stages (Fig. 1) tested were stages 1 to 5 and mature for photosynthesis measurements, stages 2 to 5 and mature for dark respiration,

and stages 1 to 4 and mature for all pigment determinations. Stage 5 leaves were omitted because they were physiologically similar to mature leaves. Data are means \pm standard error ($n=3$ plants, with at least four measurements per plant) except for **a**) where $n=2$ plants for stage 4 leaves, which were not included in statistical analysis, and **b**) where $n=6$ to 8 plants. Significant differences between stages were tested by ANOVA followed by Tukey HSD significance test. Different upper-case letters indicate significant differences between stages ($P < 0.05$)

with a significant increase between stage 1 leaves and nearly fully-expanded stage 3 leaves (Fig. 2a). During the visible ‘greening’ phase from rust-coloured stage 3 leaves to deep-green mature leaves, there was another twofold increase in photosynthetic carbon assimilation activity (Fig. 2a). In contrast to photosynthetic activity, dark respiration was indistinguishable in leaf stages 2 to 4 (stage 1 could not be measured), but was significantly lower in stage 5 and mature leaves than in younger leaves (Fig. 2b). Photosynthesis and dark respiration activities were approximately balanced in stage 2 leaves, likely favouring dark respiration in stage 1 leaves. Therefore, stage 3 leaves had net carbon assimilation activity, despite the lack of a distinct green colour to the eye.

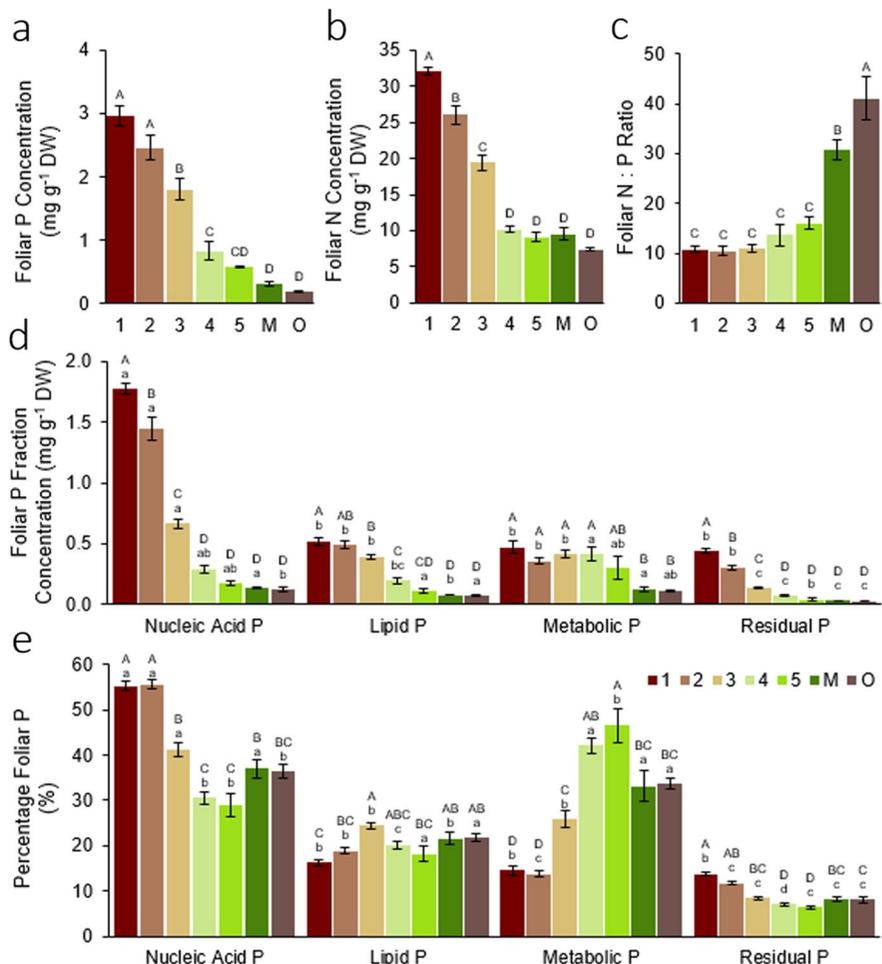
Foliar concentrations of chlorophyll a were higher than the concentrations of chlorophyll b and carotenoids (Fig. 2c) at each stage of *H. prostrata* leaf

development. The concentration of each pigment type was indistinguishable from stages 1 to 4 (Fig. 2c). However, as the leaves progressed from stage 4 to maturity there was an almost twofold increase in chlorophyll a concentration and a slight increase in carotenoid concentration (Fig. 2c). Chlorophyll b concentration was indistinguishable across all leaf developmental stages. In contrast to photosynthetic pigments, foliar anthocyanin concentrations decreased across leaf development, being significantly higher in stage 1 and 2 leaves than in stage 4 and mature leaves (Fig. 2d).

Phosphorus and nitrogen dynamics over leaf lifetime in *Hakea prostrata*

As leaves aged from the very youngest immature stage 1 to the oldest leaves on the plant, the total P

Fig. 3 Changes in total phosphorus (P), total nitrogen (N), and P fraction concentrations over leaf development in *Hakea prostrata*. **a)** foliar P concentration, **b)** foliar N concentration, **c)** foliar N : P ratio, **d)** concentration of different foliar P fractions, and **e)** percentage of total foliar P in various fractions. Colour key on panel indicates immature leaf stages 1 through 5, mature leaves (M), and old leaves (O). Data are means ± standard error (n=3 to 5). Significant differences between stages and fractions were tested by ANOVA followed by Tukey HSD significance test. Different upper-case letters indicate significant differences between stages (P < 0.05). In **d** and **e**, different lower-case letters indicate significant differences between P fractions within the same leaf stage (P < 0.05)



and N concentrations both declined (Fig. 3). The P concentration declined about 90% from stage 1 to mature and old leaves. Each of the major P-containing biochemical fractions also had a declining trend over the assessed leaf lifetime. Nucleic acids were by far the main P fraction in the youngest three stages, followed by phospholipids. In stage 1 leaves, there was nearly 3.5-fold more P in nucleic acids than in phospholipids (Fig. 3d). However, by stage 4, the metabolic pool became the largest P pool as the other P pools declined in abundance. Unlike the other P pools, the metabolic P pool did not decline strongly until late in leaf development at the transition from stage 5 to maturity (Fig. 3d). The concentration of leaf N decreased by about 70% from stage 1 to stage 4 but was then indistinguishable from stage 4 to old leaves (Fig. 3b). The N : P ratio in young leaves was about 12 and did not change significantly across leaf development. However, the ratio increased approximately twofold to 30 between stage 5 and mature leaves, and reached 40 in old leaves (Fig. 3c). These results highlight a priority to recycle P over N from the mature and old leaves.

De novo transcriptome assembly and annotation for *Hakea prostrata*

A transcriptome for *H. prostrata* was assembled and annotated to analyse changes in gene expression over leaf development. Stage 5 leaves were omitted from the transcriptome and from further analyses as they were physiologically similar to mature leaves. Raw reads were corrected and trimmed, reducing the number of reads to between 11.4 M and 59.8 M reads per library (Table S2). After filtering, 97.9% of reads had quality scores above 20 and 93.1% had quality scores above 30, indicating high quality for the RNAseq data. Assembly of the trimmed reads de novo produced a draft transcriptome that contained 678,635 transcript sequences and covered a total length of 602 Mb (Table S3). The mapping rates of reads to the transcriptome for samples from hydroponically-grown *H. prostrata* ranged from 81.5% to 88.9%, while the mapping rates for leaf samples from plants in their natural habitat ranged from 98.8% to 99.8%. Removal of ribosomal RNA, chloroplast, and mitochondrial sequences left 672,206 assembled sequence contigs with a total length of 596 Mb. A total of

274,228 putative peptide sequences were predicted in this draft transcriptome. Removal of non-Embryophyta and redundant sequences left 53,710 peptide sequences that corresponded to 51,697 unique transcripts in the transcriptome. These 51,697 transcripts were then extracted as a reference transcriptome for *H. prostrata*. The 51,697 transcripts gave a complete BUSCO score of 97.2%, indicating high quality for the reference transcriptome. Each filtering step was validated by calculating BUSCO scores and assembly statistics (Table S3).

The overall transcript distribution of individual leaf, root and cluster root samples was compared to determine the validity of measuring transcript expression using this reference *H. prostrata* transcriptome. The overall expression profile of transcripts in TPM did not differ substantially among the several representative samples used to create the transcriptome (Fig. S1). Transcripts with high and low expression were similarly distributed and the proportions did not differ greatly among the samples. Thus, the *H. prostrata* reference transcriptome was a representative reference for all samples used in its construction, including those from the leaf developmental series. Furthermore, a PCA graph demonstrated that the individual transcript expression patterns in leaves were distinct when compared to those of roots and cluster roots (Fig. S2). The PCA also demonstrated that each leaf developmental stage was distinct from every other stage and that there was a progression in expression patterns over leaf development from stage 1 to mature leaves. Moreover, the three replicates of each leaf developmental stage were closely associated in the PCA. Thus, the *H. prostrata* reference transcriptome and the samples representing leaf developmental stages were suitable for further analysis of differential gene expression patterns across leaf development.

Differential gene expression across leaf development

Pairwise comparisons among all leaf stages of the average expression levels for the 51,697 transcripts in the *H. prostrata* reference transcriptome identified 18,374 DEGs (FDR < 0.05, \log_2 FoldChange ≤ -1 or ≥ 1) (Table S4). Among these, 16,861 DEGs were differentially expressed between stage 1 leaves and another leaf developmental stage (Fig. 4). At some point in leaf development, a total of 10,338 DEGs

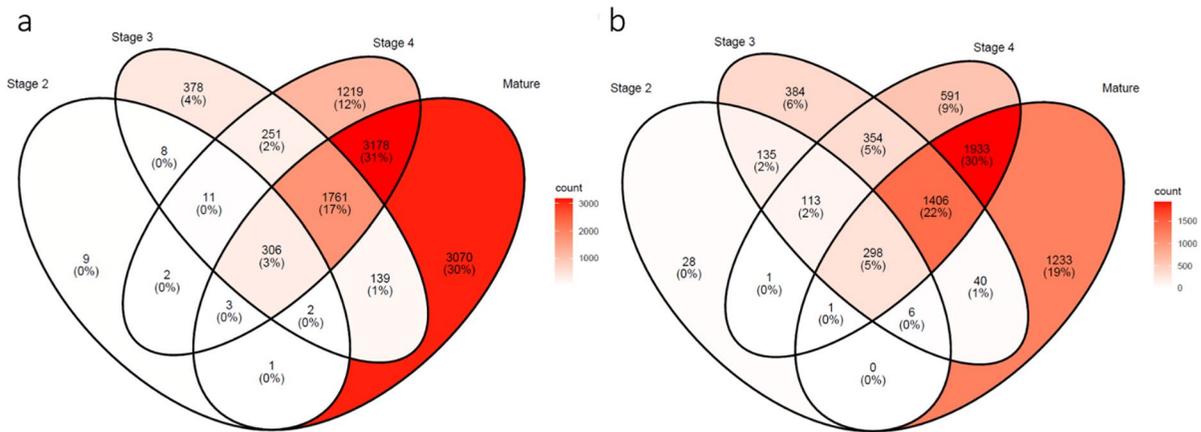


Fig. 4 Overlaps in the number of differentially expressed genes (DEGs) that had significantly **a)** lower or **b)** higher relative expression compared with the stage 1 leaves of *Hakea prostrata* ($FDR < 0.05$, \log_2 Fold Change ≤ -1 or ≥ 1). The per-

centage (%) of genes within each diagram that are within each overlap group is shown. Significance was determined using the default settings in DESeq2 based on Wald's test ($n = 3$ plants for each leaf stage)

had lower relative expression in comparison to stage 1 leaves (Fig. 4a), whilst 6,523 DEGs had higher relative expression after stage 1 (Fig. 4b). Within the 16,861 DEGs, 358 DEGs had both higher and lower expression levels relative to stage 1 at some later point in leaf development. These 358 DEGs either increased or decreased in expression compared to stage 1, and then changed significantly in the opposite direction at some later time in leaf development. Unsurprisingly, the total number of DEGs compared to stage 1 leaves increased as leaves developed.

Functional associations enriched in co-expression profiles during leaf development in *Hakea prostrata*

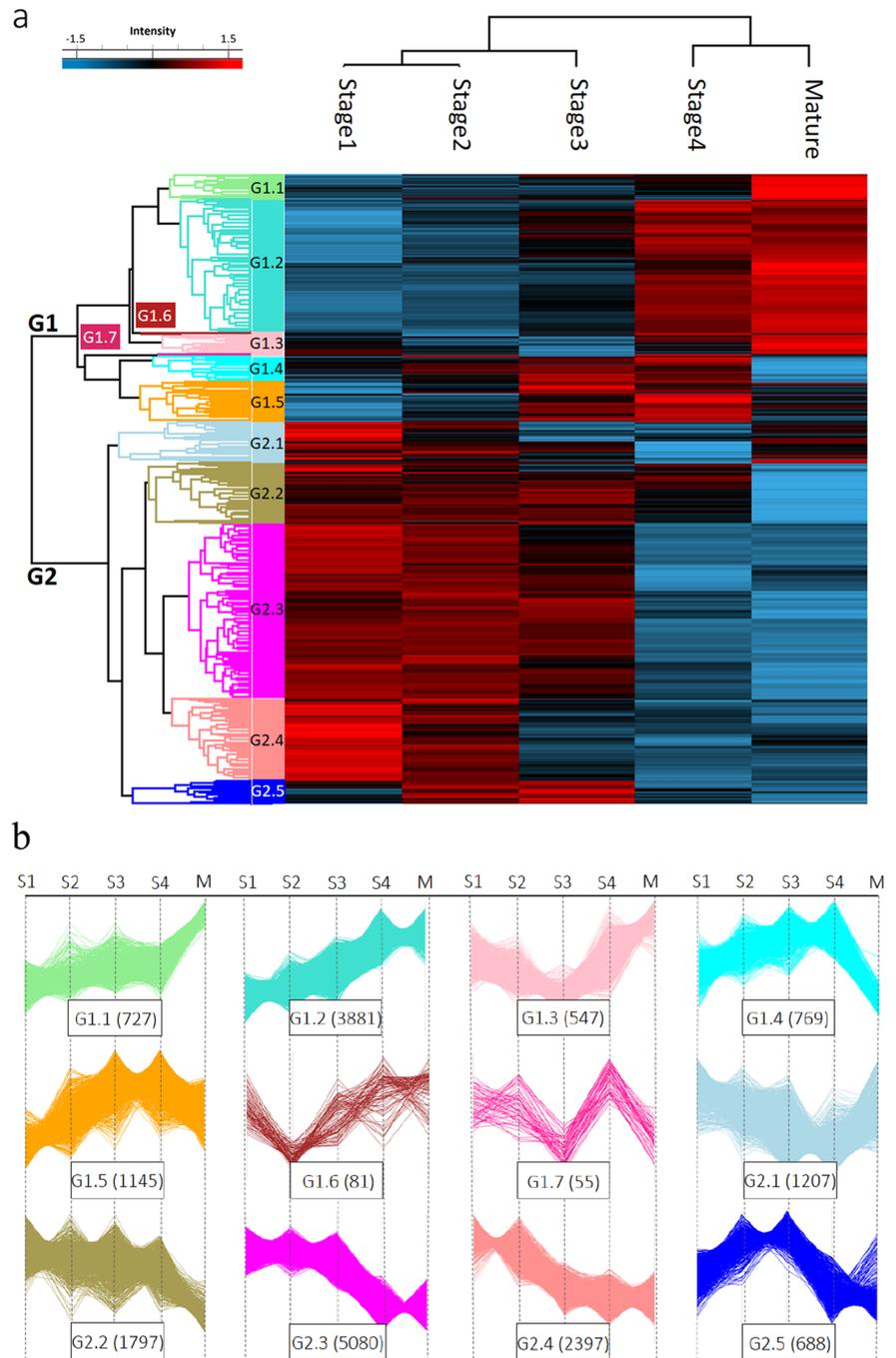
Hierarchical clustering of the 18,374 DEGs showed two distinct expression groups, designated G1 and G2 (Fig. 5a). Group G1 contained 7,205 DEGs that showed a general increase in relative expression with leaf development, while group G2 contained 11,169 DEGs that showed a general decrease in relative expression. The DEGs of group G1 were divided into seven co-expression sub-groups, G1.1 to G1.7 (Fig. 5a, b). Sub-groups G1.1 to G1.5 had more than 500 DEGs each (Fig. 5b). Sub-groups G1.6 and G1.7 had fewer than 100 DEGs each and were excluded from further analysis. The DEGs of group G2 were divided into co-expression sub-groups G2.1 to G2.5. Co-expression profiles G1.3 and G2.5 (Fig. 5b) likely included the 358 DEGs that had both higher and

lower expression at some point in leaf development compared to stage 1 leaves.

The DEGs in group G1, which generally increased in expression across leaf development, were enriched in GO functions associated with plastids and carbon metabolism, both consistent with the increase in photosynthetic competence (Fig. 2a). Other enriched functions were ion and long-chain fatty-acid transport, and protein remodelling (Fig. 6a, c). The DEGs in group G2, which generally decreased in expression across leaf development, were enriched in GO terms for functions associated with cell division, DNA replication, regulation of gene expression, and lipid and carbohydrate metabolism (Fig. 6b, d).

Gene ontology term enrichment analysis exploring the 10 main co-expression profiles offered a more detailed identification of the functional processes that changed during leaf development. We focussed first on DEGs that increased in expression as leaves developed. The DEGs within co-expression groups G1.4 and G1.5 reached peak expression earliest, at stage 3, generally maintained expression in stage 4, before decreasing to varying degrees in mature leaves (Fig. 5b). These two groups of DEGs were both particularly enriched for functions involved in establishing cell structure, particularly cell wall biosynthetic processes, DNA metabolism and general metabolic functions (Table S5e, f). Group G1.5 also included a number of functions associated with photosynthesis. The genes in

Fig. 5 Co-expression patterns among differentially expressed genes (DEGs) during leaf development in *Hakea prostrata*. **a)** Hierarchical clustering for 18,374 DEGs from low (blue) to high (red) expression across the leaf stages shown at the top. Intensity is based on z-score values for rlog transformed transcript counts. Groups G1, containing sub-groups G1.1 to G1.7, and G2, containing sub-groups G2.1 to G2.5, are shown on the left of the panel. **b)** Co-expression profiles for each sub-group in G1 and G2. The number of DEGs in each expression pattern is given in parentheses



group G1.2 increased in expression throughout leaf development, reached peak expression at stage 4, which was maintained in mature leaves (Fig. 5b). These DEGs were enriched in GO terms related to photosynthesis, including functions associated with plastid biosynthesis and carbohydrate metabolism,

along with functions around photorespiration, such as serine metabolism and peroxisomes (Table S5c). The DEGs in group G1.1 increased strongly in expression late in leaf development, between stage 4 and mature leaves (Fig. 5b.). These genes were enriched in GO terms associated with carbohydrate

binding, kinases, organic nitrogen catabolism (consistent with N recycling, Fig. 3b), iron/sulfur cluster binding and ion transport (Table S5b).

The two largest groups of DEGs that declined in transcript abundance during leaf development were G2.3 and G2.4 (Fig. 5). The DEGs in group G2.4 declined in expression earlier in leaf development, between stages 2 and 3. The genes in group G2.4 were mainly associated with the transfer of genetic information, including GO terms such as chromatin assembly, transcriptional regulation, RNA processing, nucleolus, ribosome biogenesis, and ribosomal RNA and transfer RNA metabolism (Table S5j). Other functional associations enriched in this group were cell cycle, developmental processes and hormone signalling. The DEGs in group G2.3 maintained peak expression until approaching full leaf expansion at stage 3 before declining by stage 4. The DEGs in group G2.3 included genes associated with carbohydrates, ribose-phosphate, lipid and fatty acid metabolisms, cell-wall modification, Golgi function, cytoskeleton and microtubule-based processes (Table S5i). The decline in expression of genes associated with mitochondrion may relate to the lower dark respiration activity in stage 5 and mature leaves compared to younger leaves (Fig. 2b). There was also a decline in expression of genes involved in the biogenesis of anthocyanins and the photosynthetic pigments (Fig. S3) that was consistent with the decline in anthocyanin concentration and peak of photosynthetic pigment concentration during leaf development (Fig. 2).

Leaf development in *Hakea prostrata* involves up-regulation of energy transduction and down-regulation of cellular development functions

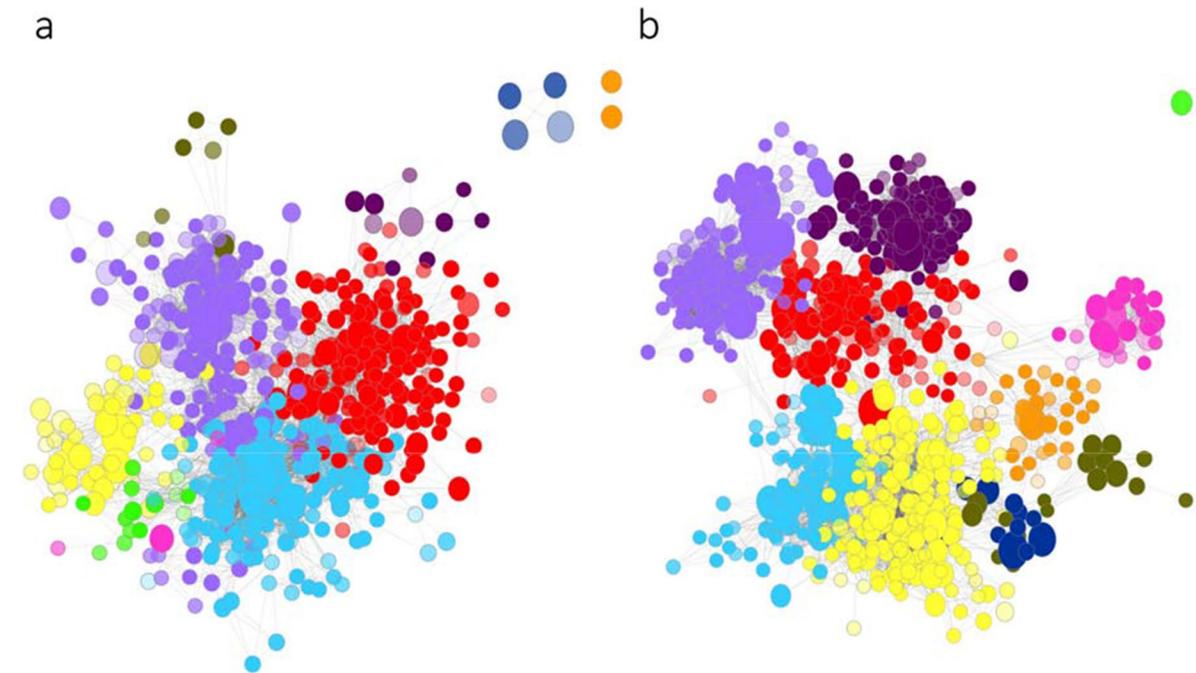
Within the DEGs of group G1, whose transcript abundance generally increased across leaf development, there was an overall enrichment in pathways associated with photosynthesis. These pathways encompassed pigment metabolism, including carotenoid biosynthesis, light harvesting, photosynthetic carbon metabolism, including carbon fixation, the pentose phosphate pathway and glyoxylate metabolism of photorespiration, as well as starch and sucrose metabolism (Table 1). These increases in transcript abundance were consistent with the increase in photosynthetic activity across leaf development (Fig. 2a). There was also an over-representation of functions

associated with N and sulfur metabolism, biosynthesis of phenylpropanoids and cofactors, as well as plant hormone signal transduction (Table 1). In contrast, the pathways enriched in DEGs from group G2, which generally decreased in transcript abundance as leaves developed, were largely associated with cell proliferation. These pathways were associated with DNA replication and repair, biosynthesis and metabolism of nucleotides for DNA and RNA, RNA degradation and ribosomes (Table 1). There appeared to be antagonistic enrichment of DEGs associated with the pathways for starch and sucrose metabolism in both group G1 and group G2 (Table 1). There were 32 unique identifiers that were linked to starch and sucrose metabolism, with isoforms for 24 of these identifiers shared between both increasing and decreasing DEGs. For example, genes encoding enzymes necessary for converting cellulose to cello-dextrin and then to cellobiose were among both the G1 and G2 profiles. Overall, enrichment analysis of functions assigned to KEGG pathways strengthened the perspective gained from the GO term enrichment analysis on co-expression groups (Fig. 6) (Table 1).

Ribosome dynamics during leaf development in *Hakea prostrata*

In relation to the delayed greening syndrome exhibited by *H. prostrata*, the ribosome pathway was the most significantly enriched KEGG pathway for DEGs in group G2, which generally exhibited a decrease in transcript abundance during leaf development. DEGs belonging to groups G2.4 and G2.1 were also enriched for GO terms associated with ribosomes (Fig. 7, Table S5g, j). We explored the dynamics of genes encoding ribosomal subunits across *H. prostrata* leaf development, focussing on potential differences among ribosomes located in the cytosol, mitochondria and chloroplasts.

There were 346 transcripts annotated as encoding ribosomal subunit proteins within the *H. prostrata* reference transcriptome (Table S6). Of these, the majority encoded cytosolic subunits, while chloroplast and mitochondrial subunits were encoded by 14% and 10% of transcripts, respectively. There were 292 ribosomal subunit transcripts expressed in leaves. Interestingly, this included all of the mitochondrial subunit transcripts and all but one of the chloroplast subunit



Cluster	GO # ^A	Gene # ^B	Largest GO term (x/n, p-value) ^C	GO term with smallest p-value (x/n, p-value) ^C	Most connected GO term (x/n, p-value) ^C
G1.C1	169	1592	response to chemical (303/1677, 1.2E-3)	plastid (622/1675, 0E-100)	plastid (622/1675, 0E-100)
G1.C2	169	1425	cellular catabolic process (333/1968, 2.9E-2)	iron-sulfur cluster binding (110/261, 1.7E-24)	organic acid metabolic process (341/1679, 9.9E-9)
G1.C3	165	1071	carbohydrate metabolic process (392/1855, 2E-12)	cellular carbohydrate metabolic process (221/768, 1.2E-21)	carbohydrate metabolic process (392/1855, 2E-12)
G1.C4	51	535	ion transport (293/1559, 1E-4)	active transmembrane transporter activity (228/1158, 5.5E-5)	ion transport (293/1559, 1E-4)
G1.C5	10	185	endopeptidase activity (140/766, 3.3E-2)	ATP-dependent peptidase activity (32/62, 7.6E-10)	endopeptidase activity (140/766, 3.3E-2)
G1.C6	4	169	protein modification by small protein conjugation (167/937, 4E-2)	ubiquitin-like protein transferase activity (126/667, 1.9E-2)	protein modification by small protein conjugation (167/937, 4E-2)
G1.C7	4	145	lyase activity (145/751, 4.7E-3)	lyase activity (145/751, 4.7E-3)	fructose-bisphosphate aldolase activity (7/15, 2.9E-2)
G1.C8	7	117	isomerase activity (116/560, 1.3E-3)	cis-trans isomerase activity (38/126, 1.8E-4)	isomerase activity (116/560, 1.3E-3)
G1.C9	2	99	oxidoreductase activity, acting on CH-OH group... (99/490, 7.3E-3)	oxidoreductase activity, acting on CH-OH group... (99/490, 7.3E-3)	oxidoreductase activity, acting on CH-OH group... (99/490, 7.3E-3)
G1.C10	15	77	microbody (62/247, 2.8E-4)	microbody (62/247, 2.8E-4)	long-chain fatty acid t transport (8/15, 6.8E-3)

Cluster	GO # ^A	Gene # ^B	Largest GO term (x/n, p-value) ^C	GO term with smallest p-value (x/n, p-value) ^C	Most connected GO term (x/n, p-value) ^C
G2.C1	109	2061	carbohydrate metabolic process (551/1855, 3.3E-10)	structural molecule activity (267/604, 1.1E-28)	cellular component biogenesis (499/1696, 1.4E-8)
G2.C2	80	1721	ATP-dependent activity (497/1773, 8.5E-6)	DNA replication (107/237, 4.3E-12)	chromosome (200/539, 7.6E-12)
G2.C3	134	1615	lipid metabolic process (442/1688, 9.7E-3)	carbohydrate derivative metabolic process (420/1260, 1.3E-15)	carbohydrate derivative metabolic process (420/1260, 1.3E-15)
G2.C4	197	1583	cellular component assembly (281/952, 3.9E-5)	microtubule-based process (207/367, 8.6E-41)	cell cycle (286/664, 1.7E-28)
G2.C5	100	1580	nitrogen compound transport (449/1762, 4.8E-2)	Golgi apparatus (443/1133, 2.1E-32)	cellular localization (446/1514, 9.3E-8)
G2.C6	11	524	transcription regulator activity (408/1454, 6.3E-5)	DNA-binding transcription factor activity (356/1238, 2.9E-5)	sequence-specific DNA binding (241/856, 3.1E-3)
G2.C7	27	438	response to chemical (435/1677, 1.9E-2)	response to hormone (220/724, 5.6E-5)	response to chemical (435/1677, 1.9E-2)
G2.C8	17	404	response to external stimulus (222/779, 2.7E-3)	negative regulation of macromolecule metabolic process (180/603, 9.1E-4)	negative regulation of metabolic process (186/626, 9.1E-4)
G2.C9	33	343	developmental process (335/1187, 2.4E-4)	cellular developmental process (112/265, 2E-10)	maintenance of root meristem identity (6/8, 2.6E-2)
G2.C10	1	188	mRNA binding (188/674, 1.6E-2)	mRNA binding (188/674, 1.6E-2)	mRNA binding (188/674, 1.6E-2)

◀Fig. 6 Gene Ontology (GO) term enrichment for co-expression groups of differentially expressed genes (DEGs) created by GOMCL for **a**) and **c**) group G1 (generally increased in expression during leaf development) and **b**) and **d**) group G2 (generally decreased in expression during leaf development) from Fig. 5. The GO term clusters were ranked by the number of DEGs within them. **a**) Group G1 network and **b**) Group G2 network. Each colour represents a unique cluster. A darker shaded node (circle) represents a lower P -value (<0.05). Node size represents the number of genes in the test set that were assigned to that GO term. ^ANumber of GO terms within the cluster. ^BNumber of genes in each cluster. ^C x = number of genes in the test set, n = number of genes associated with a GO term in the reference set

transcripts, but only 80% of the cytosolic subunit transcripts in the reference transcriptome (Table S6). Of the 292 leaf ribosomal subunit transcripts, 75% were differentially expressed during leaf development. This included about 77% of both cytosolic and chloroplast subunit transcripts found in the leaf, but only 62% of the mitochondrial transcripts (Table S6). Hierarchical clustering separated the ribosomal subunit DEGs into 52 up-regulated genes and 168 down-regulated genes across leaf development (Fig. 8).

To determine trends for subunits targeted to different cellular locations, as well as their respective expression patterns, an enrichment analysis was performed based on the two clades in Fig. 8 (Table S7). Chloroplast ribosomal subunit genes were significantly enriched in the set that was up-regulated during leaf development ($P < 0.05$) (Table S7). The transcripts for all 36 of the up-regulated chloroplast subunit genes belonged to group G1.2, which had a large increase in expression between stages 3 and 4. Unsurprisingly, this co-expression pattern was also highly enriched with DEGs associated with the GO term plastid, consistent with the increase in photosynthetic competence across leaf development (Fig. 2a). In contrast, differentially-expressed cytosolic ribosomal subunit genes were significantly enriched among the down-regulated group of ribosomal DEGs ($P < 0.05$) (Table S7). The majority of these genes belonged to co-expression group G2.4, which contained almost twice the number of ribosomal subunit genes as group G2.3 (Table S7). Nineteen of the 21 DEGs encoding mitochondrial ribosomal subunit proteins decreased in expression across leaf development (adjusted $P = 0.064$, Table S7), consistent with the decrease in dark respiration (Fig. 2b).

Discussion

Transcriptome assembly and quality

The *H. prostrata* reference transcriptome compared favourably in quality and size to other recent plant transcriptome assemblies. The 51,697 representative transcripts from white roots, cluster roots, and leaves was roughly twice the approximately 27,000 protein-coding genes found in *A. thaliana* (Jiao and Schneeberger 2020). The larger set of *H. prostrata* transcripts was most likely due to heterozygosity in this outcrossing diploid species. The number of transcripts was similar to the 59,063 transcripts generated for a leaf transcriptome from *Banksia hookeriana* (Proteaceae) (Lim et al. 2017) and the 67,004 transcripts generated for a young leaf and flower transcriptome from *Grevillea juniperina* (Proteaceae) (Damerval et al. 2019). Our *H. prostrata* reference transcriptome had a higher score of 97% complete BUSCOs (57.4% single-copy, 39.8% duplicated) than the *G. juniperina* transcriptome (85% complete: 75% single-copy, 10% duplicated). This difference was likely due to the inclusion of root and mature leaf RNAseq libraries in the *H. prostrata* transcriptome. The BUSCO score for *H. prostrata* also compared well to the BUSCO score for the recently published transcriptome of wild peanut (*Arachis glabrata*) roots, stems, leaves, flowers and pods (93% complete: 29% single-copy, 64% duplicated) (Zhao et al. 2021). Overall, these results indicated that the *H. prostrata* reference transcriptome was a suitable resource for examining changes in functional enrichment during leaf development and would be useful for the identification of genes expressed in leaves, roots and cluster roots in *H. prostrata*.

The development of photosynthesis machinery is concomitant with delayed leaf greening

Whilst a general increase in expression of genes encoding photosynthetic machinery appears necessary during leaf development in both monocots and dicots (Andriankaja et al. 2012; Li et al. 2010; Wang et al. 2014), the refined co-expression patterns for *H. prostrata* gave insights into the delay in reaching photosynthetic maturity that typified delayed-greening in this species (Fig. 9). Leaves of delayed-greening species take longer to reach peak

Table 1 Enriched KEGG pathways ($P < 0.05$) among differentially expressed genes (DEGs) showing a general increase (group G1) or decrease (group G2) in relative transcript abundance during leaf development. Pathways were identified using

clusterProfiler and enrichment tests performed using the Over-Representation Analysis algorithm. The denominator for the test gene ratios was 984 for G1 and 1500 for G2. The denominator for the background (entire transcriptome) gene ratio was 2537

KEGG pathway ID	Pathway description	Genes in test set	Genes in transcriptome	Benjamini–Hochberg adjusted P -value
KEGG pathways enriched in group G1				
ko00195	Photosynthesis	30	33	1.10E-07
ko00500	Starch and sucrose metabolism	28	32	1.60E-06
ko00940	Phenylpropanoid biosynthesis	18	18	2.90E-06
ko00906	Carotenoid biosynthesis	16	17	1.30E-04
ko00630	Glyoxylate and dicarboxylate metabolism	23	29	3.60E-04
ko00196	Photosynthesis—antenna proteins	12	112	3.80E-04
ko01200	Carbon metabolism	56	93	5.00E-04
ko04146	Peroxisome	28	39	7.00E-04
ko04075	Plant hormone signal transduction	29	43	2.80E-03
ko01240	Biosynthesis of cofactors	68	125	4.50E-03
ko00910	Nitrogen metabolism	11	13	1.90E-02
ko00040	Pentose and glucuronate interconversions	12	15	2.40E-02
ko00710	Carbon fixation in photosynthetic organisms	17	24	2.40E-02
ko00030	Pentose phosphate pathway	13	17	2.80E-02
ko00920	Sulfur metabolism	11	14	4.00E-02
KEGG pathways enriched in group G2				
ko03010	Ribosome	98	114	2.40E-08
ko01250	Biosynthesis of nucleotide sugars	34	37	5.00E-04
ko00520	Amino sugar and nucleotide sugar metabolism	37	41	5.00E-04
ko03018	RNA degradation	39	48	3.00E-02
ko03030	DNA replication	28	33	3.60E-02
ko03430	Mismatch repair	19	21	4.00E-02
ko00500	Starch and sucrose metabolism	27	32	4.00E-02

chlorophyll concentration than leaves of most crop species (Kursar and Coley 1992b). The results from *H. prostrata* indicate that this is also likely to be the case for many other chloroplast functions. Moreover, in a recent comparison of developing leaves in tropical plants, Vogado et al. (2022) showed that young leaves of delayed-greening species had significantly lower photosynthetic rates than those of species exhibiting green flushes of young leaves, as was the case in *H. prostrata*. Amongst ‘green-flushing’ species, one study in *A. thaliana* seedlings showed that photosynthetic machinery was functional just 14 h following de-etiolation (Pipitone et al. 2021). In *Triticum aestivum* (wheat), a strong onset of chloroplastic biogenesis and photosynthetic

activity occurred in leaf cells at approximately 1.5 days of age (Loudya et al. 2021). These analyses demonstrated two distinct phases of cell development, with early plastid proliferation preceding chloroplast development in both *A. thaliana* and *T. aestivum* (Loudya et al. 2021; Pipitone et al. 2021). This cell development pattern may also occur in *H. prostrata*, but over a much longer time frame than in those green-flushing species (Fig. 9).

In *H. prostrata*, the largest co-expression group of DEGs that showed increased expression during leaf development (group G1.2) was up-regulated throughout the development of immature leaves. The increase in expression for this group was particularly strong between stages 2 and 4. This

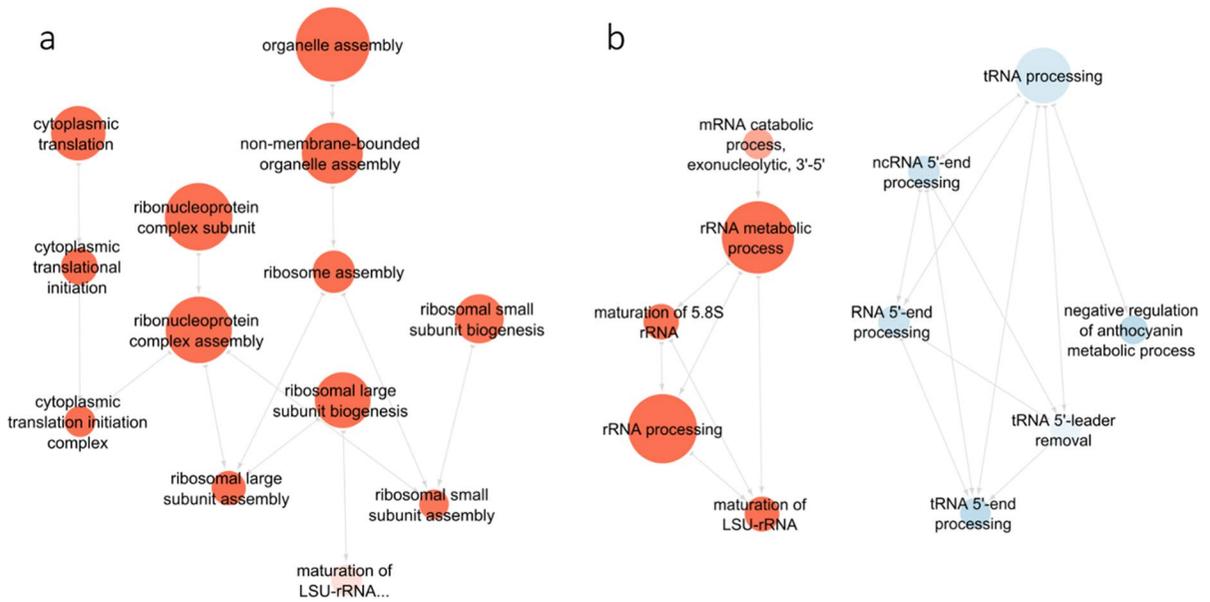


Fig. 7 Network of enriched Gene Ontology (GO) terms for biological processes (BP) related to ribosomes. **a)** Group G2.4, cluster 2, sub-cluster 1 (Table S5j). **b)** Group G2.4, cluster 3, sub-clusters 1 (red) and 2 (blue) (Table S5j). Each colour rep-

resents a unique cluster, with the darker shade of a node (circle) representing a lower *P*-value and thereby higher significance ($P < 0.05$). Node size represents the number of genes in the test set that were assigned to that GO term

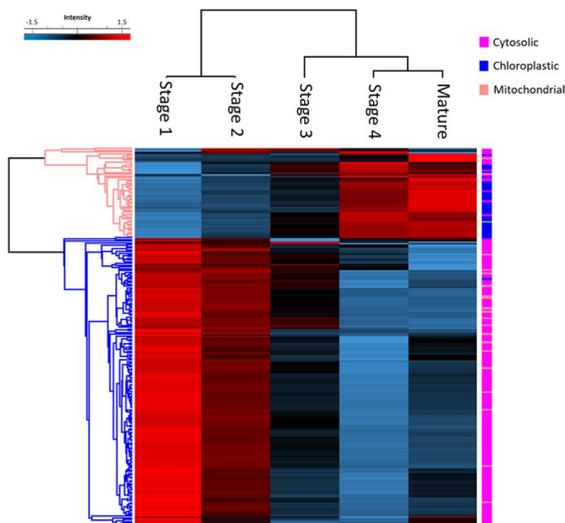
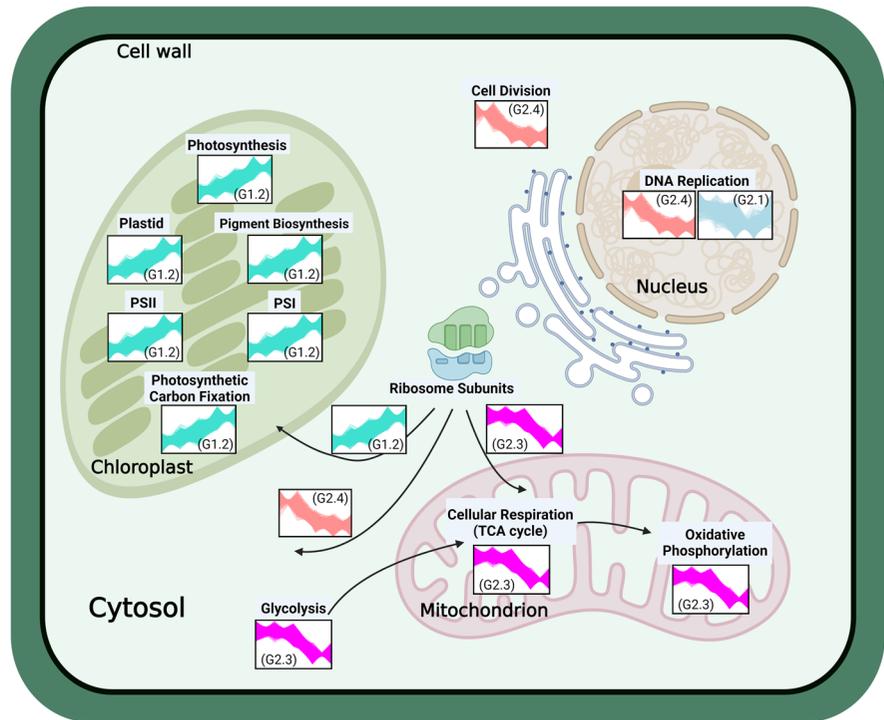


Fig. 8 Expression of differentially-expressed ribosomal subunit genes across leaf development ($P < 0.05$, \log_2 Fold Change ≤ -1 or ≥ 1). In each leaf stage (across top), relative expression counts were log transformed and presented on a scale from lower (blue) to higher (red) values. Intensity is based on z-score values for transcript counts. Hierarchical clustering into two main clades (blue, pink) is shown on the left. Cellular compartment of each subunit is shown on the right

expression pattern, therefore, described genes that showed up-regulation during the visible 'greening' phase. The up-regulation of these genes overlapped with the decrease in leaf anthocyanin concentration after stage 2 and ended before the increase in both chlorophyll a and carotenoid concentrations between stage 4 and mature leaves. Given the changes in leaf pigments, coupled with the increase in photosynthetic capacity throughout leaf development, it was not surprising, then, that the GO term plastid was both the most enriched and most connected term in cluster 1 of this expression profile. This large cluster contained over 25% of DEGs that were in this co-expression group and comprised GO terms for increased plastid and photosynthetic functioning. These functions included carbohydrate biosynthesis associated with carbon metabolism, as well as chloroplast genetic information processing. Therefore, the transcript profiles aligned strongly with the finding that young *H. prostrata* leaves have negligible rates of photosynthesis and approximately 50% of the chlorophyll a concentration of mature leaves. As would be expected for a C_3 plant, functions associated with photorespiration were also enriched among group G1.2. These functions would be necessary to

Fig. 9 Transcript expression patterns for key cellular processes involved in delayed leaf greening in *Hakea prostrata*. The transcript abundance pattern for each process from Fig. 5 is shown, with the group designation in parentheses. Created with BioRender.com



complement an increase in Calvin cycle activity in this C_3 plant (Wingler et al. 2000).

Linking transcriptional reprogramming to phases of leaf development

The leaf developmental stages that we sampled provided a good transcriptomic background for detailed analysis of the transition from emergence to maturity for *H. prostrata* leaves. The stage 1 leaves were collected approximately 1 to 3 days after the lamina had emerged from the primordia, suggesting that the expanding leaf was still undergoing primary morphogenesis with the majority of cells undergoing proliferation through division (Bar and Ori 2014; Nelissen et al. 2016). In *A. thaliana*, the transition from the cell division phase to the expansion phase occurs when leaves are about 20% fully expanded (Beemster et al. 2005). This level of expansion coincides with the smallest stage 2 leaves we collected for *H. prostrata*, suggesting that stage 2 leaves were already beyond the transition from the cell division and into the expansion phase. We assume that cells in the near-fully-expanded leaves of stage 3 would be predominantly at the end of the expansion phase, while those

in fully-expanded stage 4 leaves would be beyond the expansion phase and in a maturation phase. The cells in the mature leaves from the previous flush of growth, then, would be mature. It is important to keep in mind that the leaf stages examined represented a mixture of cells at different stages of development, with those near the tip of the leaf being more advanced than those nearer the base (Donnelly et al. 1999). Thus, the patterns of gene expression observed represent an average of all events occurring within the entire leaf, even when most cells were in one phase.

Three co-expression patterns showed only down-regulation across leaf development. Therefore, the DEGs with these patterns might serve primarily in determining the developmental structure of the leaf. These three DEG groups had a clear progression in their down-regulation, starting after stage 2 (group G2.4), after stage 3 (group G2.3) or after stage 4 (group G2.2). The cohort of DEGs that declined in expression earliest in leaf development (group G2.4) likely included genes no longer needed once the cell division phase ended (Fig. 9). Supporting this idea was the finding that GO terms involved in the cell cycle were some of the most enriched in this group of DEGs. The continual decline in expression suggested

that transition from cell division to expansion may still be occurring beyond stage 2. Thus, in *H. prostrata*, as in *A. thaliana*, the transition to a stage of secondary morphogenesis in leaf cells may occur at a cell-cycle arrest front, which remains fixed for a period of time, before making a fairly rapid progression towards the base of the leaf (Andriankaja et al. 2012). Other GO terms enriched among the G2.4 co-expression profile largely surrounded genetic information processing, protein biogenesis and hormone signalling. The inclusion of hormone signalling was consistent with auxin signalling being involved in the cell cycle transition phase (Kalve et al. 2014). Cell expansion may be linked to chloroplast retrograde signalling within *A. thaliana* leaves (Andriankaja et al. 2012). *Hakea prostrata* therefore provides an opportunity to dissect this link more precisely, given the extended time over which the transition occurs.

Functions enriched amongst the DEGs that declined in expression after stage 3 (group G2.3) likely highlighted cellular development processes that were no longer needed once *H. prostrata* leaves reached full expansion. These processes included organelle development, particularly mitochondria and Golgi apparatus, as well as cytoskeletal, cell wall and all associated membranes. The finding that these genes remained close to peak expression until stage 3 supported the view that the transition between expansion and maturation of cells in *H. prostrata* leaves is at stage 3. The inclusion of mitochondrial development in this group of genes was consistent with the late decline in dark respiration that was observed during leaf development. The DEGs that showed a major decline in expression between fully-expanded, bright green leaves and mature leaves (group G2.2) encompassed GO terms enriched for functions involved in cell-wall biogenesis. The production of cellulose during this transition would contribute greatly to the increased rigidity of stage 4 leaves compared with stage 3 leaves, and the continued need for these functions as stage 4 leaves increased in rigidity in this sclerophyllous species. The GO terms with the largest number of DEGs within the co-expression pattern G2.2 were involved in cytoskeletal protein binding. This cluster included GO terms associated with actin filament-based processes and actin-cytoskeleton organisation, which suggested that the construction of the cytoskeleton, which supports the movement

of chloroplasts and other organelles, was completed before full leaf maturity (Wada and Kong 2018).

Ribosome dynamics across *Hakea prostrata* leaf development

The strong decrease in nucleic acid P concentration from stages 1 to 4 of leaf development suggested that there was a strong decrease in rRNA concentration during this phase since most nucleic acid P is found in rRNA (Veneklaas et al. 2012). The likely reduction in rRNA abundance, coupled with the ribosome pathway being the most significantly enriched KEGG pathway amongst DEGs decreasing in expression during leaf development, was associated with a clear shift in expression of genes encoding specific organellar ribosomal subunit proteins in developing leaves of *H. prostrata*. Generally, expression of genes encoding both cytosolic and mitochondrial ribosomal subunits declined during development, whilst expression of genes encoding chloroplast ribosomal subunits increased in expression (Fig. 9). This shift between cytosolic and plastidic ribosome components was captured by Sulpice et al. (2014) at the rRNA level for all six species of Proteaceae that they examined, except *H. prostrata*. This exception was likely due to the *H. prostrata* leaves in that study being sampled later in leaf development compared to the other five species, when rRNA levels had already shifted. Thus, it is likely that the species of *Hakea* and *Bankisia* examined by Sulpice et al. (2014) also undergo a similar shift in the expression of genes involved in the production of ribosomal subunits during leaf development. The decreased expression of genes encoding ribosomal subunits and the decrease in rRNA abundance was closely paralleled by a decrease in leaf N concentration in *H. prostrata*, supporting the previous suggestion that part of the decrease in leaf N concentration was due to breakdown of cytosolic ribosomal subunit proteins and recycling of the released N (Sulpice et al. 2014).

Ribosomal proteins are essential components for cell division and expansion in leaf development. This has been particularly well demonstrated in numerous mutant studies of ribosomal proteins in *A. thaliana* (Schippers and Mueller-Roeber 2010). These experiments included organellar proteins, with a growing number of specific

mitochondrial and chloroplast ribosomal proteins found to be involved in multiple developmental processes (Robles and Quesada 2017, 2022; Romani et al. 2012). The expression patterns during leaf development exhibited by genes encoding *H. prostrata* ribosomal subunits differed in important points from the patterns seen in *A. thaliana*, which may underlie some of the differences in the cell division and expansion programs seen between these two plants. Transcripts encoding chloroplast ribosomal proteins in *H. prostrata* were enriched among DEGs that increased in expression at each sampled stage through leaf development. The majority of these DEGs belonged to group G1.2 and had peak expression in fully-expanded stage 4 and mature leaves of *H. prostrata*. In contrast, *A. thaliana* genes encoding chloroplast ribosomal proteins generally decreased in expression as leaves entered the expansion phase, before increasing somewhat in expression again, but not to their maximal levels, after the leaves matured (Schippers and Mueller-Roeber 2010).

Genes encoding cytosolic ribosomal subunit proteins generally decreased in expression across all four immature stages of *H. prostrata* leaf development. Moreover, transcript abundance for most genes encoding cytosolic subunit proteins in mature *H. prostrata* leaves remained well below their initial expression levels in young leaves. *Arabidopsis thaliana* genes encoding cytosolic ribosomal proteins also show a fairly rapid decrease in expression during the transition from cell division to cell expansion (Beemster et al. 2005; Schippers and Mueller-Roeber 2010). In contrast to *H. prostrata*, though, most genes encoding cytosolic ribosomal proteins in *A. thaliana* returned to peak expression in mature leaves (Schippers and Mueller-Roeber 2010). Therefore, the decline in abundance of cytosolic ribosomal subunit transcripts may occur at a somewhat later stage of leaf cell development in *H. prostrata* than in *A. thaliana* and there was less reliance on enhanced translation in older long-lived leaves of *H. prostrata* than in older short-lived leaves of *A. thaliana*.

The majority of DEGs encoding mitochondrial ribosomal subunits in *H. prostrata* had a similar decrease in expression across leaf development to that of the cytosolic subunits, generally paralleling the

decrease in dark respiration activity. Genes encoding mitochondrial ribosomal subunits in *A. thaliana* do not have a uniform trend but are variable in expression during leaf development (Schippers and Mueller-Roeber 2010). Thus, there was a co-ordinated down-regulation of mitochondrial ribosomal subunit genes during leaf development in *H. prostrata* that is largely lacking in *A. thaliana*.

Overall, the expression dynamics for *H. prostrata* genes encoding ribosomal subunits had a clear trend moving out of the growth phase and into the photosynthetically-active phase of leaf development. These results support conclusions of Sulpice et al. (2014) that delayed leaf greening provides a P-saving mechanism in highly P-efficient *H. prostrata*. This mechanism was enabled by the progressive transition of protein synthesis machinery from cytosolic and mitochondrial ribosomes toward chloroplast ribosomes, integrated with other processes that transitioned in an ordered manner through leaf development in *H. prostrata*.

Acknowledgements High performance computing was provided by the Pawsey Supercomputing Research Centre with funding from the Australian Government and the Government of Western Australia.

Author contribution PMF, HL and RJ conceived the project and designed experiments, PEB, GW, MD, JB, DE and WRS provided advice, HL, PMF, WRS, JB, DE provided financial support, TB, BN, AI, RJ, AASE, KS, CG and BEHS performed experiments, TB, BN, PMF, RJ and HL analysed data, TB, BN and PMF wrote the manuscript, and all authors contributed to editing the manuscript.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions This research was funded by grants from the Australian Research Council to H.L. and P.M.F. (DP200101013, LP0776252) and to H.L., P.M.F. and M.D. (DP140100148), University of Western Australia (UWA) Research Training Fund Fee Offset awards to T.B. and B.N., a UWA Research Training Fund Stipend award to T.B., a UWA Co-funded Postgraduate award to B.N., a Kwongan Foundation Top-up award to T.B. and a UWA Fay Gale Travel Fellowship to R.J. M.D. acknowledges grant support from NSF-IOS-EDGE-1923589, NSF-IOS-2014300, and DOEBER-DE-SC0020358 to support MD and GW during this work.

Data availability *Hakea prostrata* RNAseq reads and transcriptome assemblies used in this study are available through NCBI under the BioProject PRJNA896696. Raw reads were uploaded to the Sequence Read Archive (SRA) with accession numbers SRR22135859 to SRR22135884. The reference transcriptome assembly of *H. prostrata* was uploaded to the

Transcriptome Shotgun Assembly (TSA) database with accession number GKDQ01000000.

Declarations

Competing interests The authors declare that there are no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Andriankaja M, Dhondt S, De Bodt S, Vanhaeren H, Coppens F, De Milde L, Mühlenbock P, Skiryca A, Gonzalez N, Beemster GT (2012) Exit from proliferation during leaf development in *Arabidopsis thaliana*: a not-so-gradual process. *Dev Cell* 22:64–78. <https://doi.org/10.1016/j.devcel.2011.11.011>
- Armenteros JJA, Salvatore M, Emanuelsson O, Winther O, Von Heijne G, Elofsson A, Nielsen H (2019) Detecting sequence signals in targeting peptides using deep learning. *Life Sci Alliance* 2. <https://doi.org/10.26508/lsa.201900429>
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT (2000) Gene ontology: tool for the unification of biology. *Nat Genet* 25:25–29. <https://doi.org/10.1038/75556>
- Bar M, Ori N (2014) Leaf development and morphogenesis. *Development* 141:4219–4230. <https://doi.org/10.1242/dev.106195>
- Beemster GT, De Veylder L, Vercruyse S, West G, Rombaut D, Van Hummelen P, Galichet A, Gruissem W, Inzé D, Vuylsteke M (2005) Genome-wide analysis of gene expression profiles associated with cell cycle transitions in growing organs of *Arabidopsis*. *Plant Physiol* 138:734–743. <https://doi.org/10.1104/pp.104.053884>
- Bray NL, Pimentel H, Melsted P, Pachter L (2016) Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 34:525–527. <https://doi.org/10.1038/nbt.3519>
- Buchfink B, Reuter K, Drost H-G (2021) Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods* 18:366–368. <https://doi.org/10.1038/s41592-021-01101-x>
- Close DC, Beadle CL (2003) The ecophysiology of foliar anthocyanin. *Bot Rev* 69:149–161. [https://doi.org/10.1663/0006-8101\(2003\)069\[0149:TEOFA\]2.0.CO;2](https://doi.org/10.1663/0006-8101(2003)069[0149:TEOFA]2.0.CO;2)
- Conesa A, Götz S (2008) Blast2GO: a comprehensive suite for functional analysis in plant genomics. *Int J Plant Genomics*. <https://doi.org/10.1155/2008/619832>
- Consortium. GO (2021) The gene ontology resource: enriching a GOld mine. *Nucleic Acids Res* 49:D325–D334. <https://doi.org/10.1093/nar/gkaa1113>
- Damerval C, Citerne H, Conde E Silva N, Deveaux Y, Delanoy E, Joets J, Simonnet F, Staedler Y, Schönenberger J, Yansouni J, Le Guilloux M, Sauquet H, Nadot S (2019) Unraveling the developmental and genetic mechanisms underpinning floral architecture in Proteaceae. *Front Plant Sci* 10:18–18. <https://doi.org/10.3389/fpls.2019.00018>
- Donnelly PM, Bonetta D, Tsukaya H, Dengler RE, Dengler NG (1999) Cell cycling and cell enlargement in developing leaves of *Arabidopsis*. *Dev Biol* 215:407–419. <https://doi.org/10.1006/dbio.1999.9443>
- Eddy SR (2011) Accelerated profile HMM searches. *PLoS Comp Biol* 7:1002195–1002195. <https://doi.org/10.1371/journal.pcbi.1002195>
- El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE, Finn RD (2019) The Pfam protein families database in 2019. *Nucleic Acids Res* 47:D427–D432. <https://doi.org/10.1093/nar/gky995>
- Figueira L, Li W, McWilliam H, Lopez R, Xenarios I, Bougueleret L, Bridge A, Poux S, Redaschi N, Aimò L (2015) Consortium, U., UniProt: a hub for protein information. *Nucleic Acids Res* 43:D204–D212
- Francis FJ, Markakis PC (1989) Food colorants: anthocyanins. *Crit Rev Food Sci Nutr* 28:273–314. <https://doi.org/10.1080/10408398909527503>
- Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28:3150–3152. <https://doi.org/10.1093/bioinformatics/bts565>
- Geng Y, Cai C, McAdam SAM, Banks JA, Wisecaver JH, Zhou Y (2021) A *de novo* transcriptome assembly of *Ceratopteris richardii* provides insights into the evolutionary dynamics of complex gene families in land plants. *Genome Biol Evol* 13. <https://doi.org/10.1093/gbe/evab042>
- Giusti MM, Wrolstad RE (2001) Characterization and measurement of anthocyanins by UV-visible spectroscopy. *Curr Protocol Food Anal Chem*: F1. 2.1-F1. 2.13. <https://doi.org/10.1002/0471142913.faf0102s00>
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29:644–652. <https://doi.org/10.1038/nbt.1883>
- Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* 8:1494–1512. <https://doi.org/10.1038/nprot.2013.084>
- Hayes PE, Clode PL, Oliveira RS, Lambers H (2018) Proteaceae from phosphorus-impooverished habitats preferentially allocate phosphorus to photosynthetic cells: an adaptation

- improving phosphorus-use efficiency. *Plant Cell Environ* 41:605–619. <https://doi.org/10.1111/pce.13124>
- Hopper SD, Gioia P (2004) The southwest Australian floristic region: evolution and conservation of a global hot spot of biodiversity. *Annu Rev Ecol Evol Syst*: 623–650. <https://doi.org/10.1146/annurev.ecolsys.35.112202.130201>
- Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ (2019) eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* 47:D309–D314. <https://doi.org/10.1093/nar/gky1085>
- Hughes NM, Morley CB, Smith WK (2007) Coordination of anthocyanin decline and photosynthetic maturation in juvenile leaves of three deciduous tree species. *New Phytol* 175:675–685. <https://doi.org/10.1111/j.1469-8137.2007.02133.x>
- Jiao W-B, Schneeberger K (2020) Chromosome-level assemblies of multiple *Arabidopsis* genomes reveal hotspots of rearrangements with altered evolutionary dynamics. *Nat Commun* 11:1–10. <https://doi.org/10.1038/s41467-020-14779-y>
- Kalve S, De Vos D, Beecher GT (2014) Leaf development: a cellular perspective. *Front Plant Sci* 5:362. <https://doi.org/10.3389/fpls.2014.00362>
- Kim D, Paggi JM, Park C, Bennett C, Salzberg SL (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 37:907–915. <https://doi.org/10.1038/s41587-019-0201-4>
- Krogh A, Larsson B, Von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567–580. <https://doi.org/10.1006/jmbi.2000.4315>
- Kuppusamy T, Giavalisco P, Arvidsson S, Sulpice R, Stitt M, Finnegan PM, Scheible W-R, Lambers H, Jost R (2014) Lipid biosynthesis and protein concentration respond uniquely to phosphate supply during leaf development in highly phosphorus-efficient *Hakea prostrata*. *Plant Physiol* 166:1891–1911. <https://doi.org/10.1104/pp.114.248930>
- Kuppusamy T, Hahne D, Ranathunge K, Lambers H, Finnegan PM, Kuppusamy T, Hahne D, Ranathunge K, Lambers H, Finnegan PM (2021) Delayed greening in phosphorus-efficient *Hakea prostrata* (Proteaceae) is a photoprotective and nutrient-saving strategy. *Funct Plant Biol* 48:218–230. <https://doi.org/10.1071/FP19285>
- Kursar TA, Coley PD (1992a) The consequences of delayed greening during leaf development for light absorption and light use efficiency. *Plant Cell Environ* 15:901–909. <https://doi.org/10.1111/j.1365-3040.1992.tb01022.x>
- Kursar TA, Coley PD (1992b) Delayed development of the photosynthetic apparatus in tropical rain forest species. *Funct Ecol*: 411–422. <https://doi.org/10.2307/2389279>
- Kursar TA, Coley PD (1992c) Delayed greening in tropical leaves: an antiherbivore defense? *Biotropica* 24:256–262. <https://doi.org/10.2307/2388520>
- Lambers H (2022) Phosphorus acquisition and utilization in plants. *Annu Rev Plant Biol* 73:11–126. <https://doi.org/10.1146/annurev-arplant-102720-125738>
- Lambers H, Cawthray GR, Giavalisco P, Kuo J, Laliberté E, Pearce SJ, Scheible WR, Stitt M, Teste F, Turner BL (2012) Proteaceae from severely phosphorus-impooverished soils extensively replace phospholipids with galactolipids and sulfolipids during leaf development to achieve a high photosynthetic phosphorus-use-efficiency. *New Phytol* 196:1098–1108. <https://doi.org/10.1111/j.1469-8137.2012.04285.x>
- Lambers H, Clode PL, Hawkins HJ, Laliberté E, Oliveira RS, Reddell P, Shane MW, Stitt M, Weston P (2015a) Metabolic adaptations of the non-mycotrophic Proteaceae to soils with low phosphorus availability. *Annu Plant Rev* 48:289–336. <https://doi.org/10.1002/9781118958841.ch11>
- Lambers H, Finnegan PM, Jost R, Plaxton WC, Shane MW, Stitt M (2015b) Phosphorus nutrition in Proteaceae and beyond. *Nature Plants* 1:1–9. <https://doi.org/10.1038/nplants.2015.10910.1038/nplants.2015.109>
- Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, Kebrom TH, Provart N, Patel R, Myers CR (2010) The developmental dynamics of the maize leaf transcriptome. *Nat Genet* 42:1060–1067. <https://doi.org/10.1038/ng.703>
- Lim SL, D'Agui HM, Enright NJ, He T (2017) Characterization of leaf transcriptome in *Banksia hookeriana*. *Genomics Proteomics Bioinformatics* 15:49–56. <https://doi.org/10.1016/j.gpb.2016>
- Loudya N, Mishra P, Takahagi K, Uehara-Yamaguchi Y, Inoue K, Bogre L, Mochida K, López-Juez E (2021) Cellular and transcriptomic analyses reveal two-staged chloroplast biogenesis underpinning photosynthesis build-up in the wheat leaf. *Genome Biol* 22:1–30. <https://doi.org/10.1186/s13059-021-02366-3>
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:1–21. <https://doi.org/10.1186/s13059-014-0550-8>
- Maere S, Heymans K, Kuiper M (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 21:3448–3449. <https://doi.org/10.1093/bioinformatics/bti551>
- Manni M, Berkeley MR, Seppely M, Simão FA, Zdobnov EM (2021) BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol* 38:4647–4654. <https://doi.org/10.1093/molbev/msab199>
- Matusiewicz H, Golik B (2004) Determination of major and trace elements in biological materials by microwave induced plasma optical emission spectrometry (MIP-OES) following tetramethylammonium hydroxide (TMAH) solubilization. *Microchem J* 76:23–29. <https://doi.org/10.1016/j.microc.2003.10.007>
- Miyazawa S-I, Makino A, Terashima I (2003) Changes in mesophyll anatomy and sink-source relationships during leaf development in *Quercus glauca*, an evergreen tree showing delayed leaf greening. *Plant Cell Environ* 26:745–755. <https://doi.org/10.1046/j.1365-3040.2003.01011.x>
- Nelissen H, Gonzalez N, Inze D (2016) Leaf growth in dicots and monocots: so different yet so alike. *Curr Opin Plant Biol* 33:72–76. <https://doi.org/10.1016/j.pbi.2016.06.009>
- Petersen TN, Brunak S, Von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8:785–786. <https://doi.org/10.1038/nmeth.1701>
- Pipitone R, Eicke S, Pfister B, Glauser G, Falconet D, Uwizeye C, Pralon T, Zeeman SC, Kessler F, Demarsy E (2021) A multifaceted analysis reveals two distinct phases of chloroplast biogenesis during de-etiolation in *Arabidopsis*. *Elife* 10:e62709. <https://doi.org/10.7554/eLife.62709>

- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–D596. <https://doi.org/10.1093/nar/gks1219>
- R Core Team (2022) A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Reich PB, Walters MB, Ellsworth DS (1997) From tropics to tundra: global convergence in plant functioning. *Proc Natl Acad Sci* 94:13730–13734. <https://doi.org/10.1073/pnas.94.25.13730>
- Robles P, Quesada V (2017) Emerging roles of mitochondrial ribosomal proteins in plant development. *Int J Mol Sci* 18:2595. <https://doi.org/10.3390/ijms18122595>
- Robles P, Quesada V (2022) Unveiling the functions of plastid ribosomal proteins in plant development and abiotic stress tolerance. *Plant Physiol Biochem*. <https://doi.org/10.1016/j.plaphy.2022.07.029>
- Romani I, Tadini L, Rossi F, Masiero S, Pribil M, Jahns P, Kater M, Leister D, Pesaresi P (2012) Versatile roles of Arabidopsis plastid ribosomal proteins in plant growth and development. *Plant J* 72:922–934. <https://doi.org/10.1111/tpj.12000>
- RStudio Team (2021) RStudio: Integrated Development for R. In: R Team (ed). RStudio, PBC
- Sayers EW, Beck J, Bolton EE, Bourexis D, Brister JR, Canese K, Comeau DC, Funk K, Kim S, Klimke W (2021) Database resources of the national center for biotechnology information. *Nucleic Acids Res* 49:D10. <https://doi.org/10.1093/nar/gk11031>
- Schippers JH, Mueller-Roeber B (2010) Ribosomal composition and control of leaf development. *Plant Sci* 179:307–315. <https://doi.org/10.1016/j.plantsci.2010.06.012>
- Shane MW, Lambers H (2005) Cluster roots: a curiosity in context. *Plant Soil* 274:101–125. <https://doi.org/10.1007/s11104-004-2725-7>
- Shane MW, De Vos M, de Roock S, Cawthray GR, Lambers H (2003) Effects of external phosphorus supply on internal phosphorus concentration and the initiation, growth and exudation of cluster roots in *Hakea prostrata* R. *Br Plant Soil* 248:209–219. <https://doi.org/10.1023/A:1022320416038>
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504. <https://doi.org/10.1101/gr.1239303>
- Song L, Florea L (2015) Rcorrector: efficient and accurate error correction for Illumina RNA-seq reads. *GigaScience* 4: 13742–13015–10089–y. <https://doi.org/10.1186/s13742-015-0089-y>
- Sulpice R, Ishihara H, Schlereth A, Cawthray GR, Encke B, Givalisco P, Ivakov A, Arrivault S, Jost R, Krohn N (2014) Low levels of ribosomal RNA partly account for the very high photosynthetic phosphorus-use efficiency of *P. rostrata* species. *Plant Cell Environ* 37:1276–1298. <https://doi.org/10.1111/pce.12240>
- Tenenbaum D, Maintainer B (2022) KEGGEST: client-side REST access to the Kyoto Encyclopedia of Genes and Genomes (KEGG), R Package Version 1.36.2
- Turner BL, Laliberté E (2015) Soil development and nutrient availability along a 2 million-year coastal dune chronosequence under species-rich mediterranean shrubland in southwestern Australia. *Ecosystems* 18:287–309. <https://doi.org/10.1007/s10021-014-9830-0>
- Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J (2016) The Perseus computational platform for comprehensive analysis of (prote) omics data. *Nat Methods* 13:731–740. <https://doi.org/10.1038/nmeth.3901>
- Veneklaas EJ, Lambers H, Bragg J, Finnegan PM, Lovelock CE, Plaxton WC, Price CA, Scheible WR, Shane MW, White PJ (2012) Opportunities for improving phosphorus-use efficiency in crop plants. *New Phytol* 195:306–320. <https://doi.org/10.1111/j.1469-8137.2012.04190.x>
- Vogado NO, Cheesman AW, Cernusak LA (2022) Delayed greening during leaf expansion under ambient and elevated CO₂ in tropical tree seedlings. *Austral Ecol* 47:530–540. <https://doi.org/10.1111/aec.13139>
- Wada M, Kong S-G (2018) Actin-mediated movement of chloroplasts. *J Cell Sci* 131:jcs.210310. <https://doi.org/10.1242/jcs.210310>
- Wang L, Czedit-Eysenberg A, Mertz RA, Si Y, Tohge T, Nunes-Nesi A, Arrivault S, Dedow LK, Bryant DW, Zhou W (2014) Comparative analyses of C4 and C3 photosynthesis in developing leaves of maize and rice. *Nat Biotechnol* 32:1158–1165. <https://doi.org/10.1038/nbt.3019>
- Wang G, Oh D-H, Dassanayake M (2020) GOMCL: a toolkit to cluster, evaluate, and extract non-redundant associations of Gene Ontology-based functions. *BMC Bioinformatics* 21:1–9. <https://doi.org/10.1186/s12859-020-3447-4>
- Wellburn AR (1994) The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *J Plant Physiol* 144:307–313. [https://doi.org/10.1016/S0176-1617\(11\)81192-2](https://doi.org/10.1016/S0176-1617(11)81192-2)
- Western Australian Herbarium (1998) Florabase--the Western Australian Flora. Department of Biodiversity, Conservation and Attractions. <https://florabase.dpaw.wa.gov.au/>. Accessed 30 Jul 2022
- Wickham H (2016) Data analysis. ggplot2. Springer, Cham. https://doi.org/10.1007/978-3-319-24277-4_9
- Wingler A, Lea PJ, Quick WP, Leegood RC (2000) Photorespiration: metabolic pathways and their role in stress protection. *Philos Trans R Soc Lond B Biol Sci* 355:1517–1529. <https://doi.org/10.1098/rstb.2000.0712>
- Woodall GS, Stewart GR (1998) Do anthocyanins play a role in UV protection of the red juvenile leaves of *Syzygium*? *J Exp Bot* 49:1447–1450. <https://doi.org/10.1093/jxb/49.325.1447>
- Wright IJ, Reich PB, Westoby M, Ackerly DD, Baruch Z, Bongers F, Cavender-Bares J, Chapin T, Cornelissen JHC, Diemer M, Flexas J, Garnier E, Groom PK, Gulias J, Hikosaka K, Lamont BB, Lee T, Lee W, Lusk C, Midgley JJ, Navas M-L, Niinemets Ü, Oleksyn J, Osada N, Poorter H, Poot P, Prior L, Pyankov VI, Roumet C, Thomas SC, Tjoelker MG, Veneklaas EJ, Villar R (2004) The worldwide leaf economics spectrum. *Nature* 428:821–827. <https://doi.org/10.1038/nature02403>
- Wrolstad RE, Durst RW, Lee J (2005) Tracking color and pigment changes in anthocyanin products. *Trends Food Sci Technol* 16:423–428. <https://doi.org/10.1016/j.tifs.2005.03.019>
- Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L (2021) clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innov* 2:100141. <https://doi.org/10.1016/j.xinn.2021.100141>

- Yan L, Zhang X, Han Z, Pang J, Lambers H, Finnegan PM (2019) Responses of foliar phosphorus fractions to soil age are diverse along a 2 Myr dune chronosequence. *New Phytol* 223:1621–1633. <https://doi.org/10.1111/nph.15910>
- Zhao C, He L, Xia H, Zhou X, Geng Y, Hou L, Li P, Li G, Zhao S, Ma C, Tang R, Pandey MK, Varshney RK, Wang X (2021) *De novo* full length transcriptome analysis of *Arachis glabrata* provides insights into gene expression dynamics in response to biotic and abiotic stresses. *Genomics* 113:1579–1588. <https://doi.org/10.1016/j.ygeno.2021.03.030>

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