Streamlined spatial and environmental expression signatures characterize the minimalist duckweed *Wolffia australiana*

AUTHORS

Tom Denyer[1], Pin-Jou Wu[1], Kelly Colt[2], Bradley W. Abramson[2,3], Zhili Pang[4], Pavel Solansky[1], Allen Mamerto[2], Tatsuya Nobori[2,5,6], Joseph R. Ecker[2,5,6], Eric Lam[4*], Todd P. Michael[2*], Marja C.P. Timmermans[1*]

[1] Center for Plant Molecular Biology, University of Tübingen, Auf der Morgenstelle 32, Tübingen 72076, Germany

[2] Plant Molecular and Cellular Biology Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037, USA

[3] Applied Sciences and Life Sciences Laboratory, Noblis, Reston, Virginia, 20191 USA

[4] Department of Plant Biology, Rutgers, The State University of New Jersey, New Brunswick, New Jersey 08901, USA

[5] Howard Hughes Medical Institute, The Salk Institute for Biological Studies, La Jolla, California 92037, USA

[6] Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037, USA

*Correspondence:

Eric Lam, eric.lam@rutgers.edu

Todd P. Michael, tmichael@salk.edu

Marja C.P. Timmermans, marja.timmermans@uni-tuebingen.de

RUNNING TITLE

Whole plant scRNA-seq atlas of the Duckweed *Wolffia*

KEY WORDS

Wolffia australiana, duckweed, single cell RNA sequencing, PHYTOMap, cell type specific responses, submergence, time-of-day, diurnal rhythms

27 **ABSTRACT**

28 Single cell genomics permits a new resolution in the examination of molecular and cellular dynamics, allowing

29 global, parallel assessments of cell types and cellular behaviors through development and in response to

30 environmental circumstances, such as interaction with water and the light-dark cycle of the Earth. Here, we

31 leverage the smallest, and possibly most structurally reduced plant, the semi-aquatic *Wolffia australiana* to

32 understand dynamics of cell expression in these contexts at the whole plant level. We examined single cell

33 resolution RNA-sequencing data, and found *Wolffia* cells divide into four principal clusters representing the above

34 and below water-situated parenchyma and epidermis. While these tissues share transcriptomic similarity with

35 model plants, they display distinct adaptations that *Wolffia* has made for the aquatic environment. Within this

36 broad classification, discrete sub-specializations are evident with select cells showing unique transcriptomic

37 signatures associated with developmental maturation and specialized physiologies. Assessing this simplified

38 biological system temporally at two key time-of-day (TOD) transitions, we identify additional TOD-responsive genes

39 previously overlooked in whole plant transcriptomic approaches and demonstrate that the core circadian clock

40 machinery and its downstream responses can vary in cell-specific manners, even in this simplified system.

41 Distinctions between cell types and their responses to submergence and/or TOD are driven by expression changes

42 of unexpectedly few genes, characterizing *Wolffia* as a highly streamlined organism with the majority of genes

43 dedicated to fundamental cellular processes. *Wolffia* provides a unique opportunity to apply reductionist biology to

44 elucidate signaling functions at the organismal level, for which this work provides a powerful resource.

45

46

47 **INTRODUCTION**

48 Multicellular organisms comprise specialized tissues accommodating diverse cell types. This variety is required to

49 achieve the array of functions necessary for an organism to develop and thrive in dynamic environments, and

50 comes about through precise coordination of complex gene regulatory networks that integrate responses to

51 internal and external cues. Plants, in particular, have evolved intricate mechanisms to accommodate diverse biotic

52 and abiotic environmental inputs, leading to changes in their metabolism, physiology, and development at cellular

53 to organismal levels. Understanding commonalities and variations in these responses on a molecular level is vital to

54 begin decoding the fundamental drivers of tissue distinction and how responses to various inputs may be

55 coordinated at the organismal level to achieve a desired phenotype. Advances in single cell resolution RNA-

56 sequencing (scRNA-seq) have opened the possibilities for systems approaches to query cellular specialization

57 during development and their specific responses to environmental cues, allowing for transcriptomic profiling at

58 high spatiotemporal resolution (Denyer and Timmermans, 2022; Seyfferth et al., 2022). In plants, scRNA-seq studies

59 published to date have largely focused on aspects of *Arabidopsis* development at the organ or tissue level (e.g.,

60 Denyer et al., 2019; Kim et al., 2021; Zhang et al., 2021; Shahan et al. 2022). However, while many of these studies

61 have examined in depth, for example, cell lineage progressions within a developing tissue, no studies have yet

62 profiled at the whole plant level to sufficient depth to enable observations at the systems level. This is principally

63 due to the high level of structural complexity and large cell numbers in most plant models.

64 A seminal study published in 2017 did profile the transcriptome for every individual cell within the model

65 nematode *C. elegans* (Cao et al., 2017). A key characteristic that enabled this feat is the simplicity of body plan and

66 low number of cells and cell types within this model animal. A plant paralogue to such a model can be found in the

67 *Wolffia* genus of the *Lemnaceae* family. These plants, ranging in size from <1 to several millimeters (mm), have a

68 highly reduced architecture (Lam and Michael, 2022). *Wolffia* plants have no roots or vasculature system but

69 exhibit clear developmental distinctions between top and bottom portions relative to the air/water interphase. For

70 example, the upper epidermal surface of *Wolffia* fronds is characterized by the presence of stomata and pigment

71 cells, which can contain phenolic compounds that turn brown in response to UV damage (Li et al., 2023). Similarly,

72 the above water ~3-8 parenchymal cell layers comprise highly chlorophyllous and relatively compact cells, whereas

73 the underwater parenchymal tissue, which forms the bulk of the frond, contains highly vacuolated cells with

74 significantly lower chlorophyll levels (Figure 1A). Further benefitting *Wolffia* as a model is its compact genome with

75 over 90% of conserved core eudicot gene functions represented by small gene families (Michael et al., 2020; Lam

76 and Michael, 2022), which can facilitate an understanding of fundamental principles of plant responses to its

77 environmental circumstances.

78   Most, if not all green organisms partition their biological activities to a specific time-of-day (TOD), and they do this

79   both developmentally and ecologically to ensure synchronization with the daily light-dark cycles on Earth (Michael

80   et al., 2003; Sanchez and Kay, 2016; Steed et al., 2021; Oravec and Greenham, 2022). TOD regulation of specific

81   biological activities is, in part, controlled internally by the circadian clock, which is highly conserved from single cell

82   algae to higher plants (Michael, 2022a; Laosuntisuk et al., 2023). In general, the internal circadian clock, in

83   conjunction with daily changes in light and temperature, controls up to 90% of the transcriptome - focusing

84   biological activities to a specific TOD. For instance, expression of genes associated with photosynthesis and

85   phytohormones peaks in the morning, nitrogen and carbon assimilation in the late afternoon, cold acclimation and

86   defense at dusk, cell cycle in the early evening, and lignin and protein biosynthesis in the middle of the night (Figure

87   S1, Supplemental Text, Bläsing et al., 2005; Covington and Harmer 2007; Filichkin et al., 2011; Ferrari et al., 2019).

88   In Wolffia, only ~13% of genes are TOD regulated under the diurnal conditions of light/dark cycles (Supplemental

89   Text, Michael et al., 2020). The lower percentage of TOD-regulated genes in Wolffia may reflect its compact

90   genome. Alternatively, it is becoming increasingly clear that TOD expression can vary by cell type (Endo, 2016; Swift

91   et al., 2022), and that bulk RNA-seq analysis may miss such cell type specific timing information.

92

93   **RESULTS**

94   ***Wolffia* cells can be broadly divided into four distinct subpopulations**

95   We transcriptionally profiled Wolffia at single cell resolution to describe, in an un-biased manner, cell types and

96   states within this budding frond system. We first optimized a protoplast isolation method, considering the distinct

97   cell wall composition of Wolffia and the presence of a cuticle as an adaptation to life on water (Borisjuk et al.,

98   2018). Viable protoplasts from clonally propagating plants (accession wa8730) were then prepared and processed

99   through the '10x Genomics' scRNA-seq cell capture and library production pipeline. Plants were gown under

100   intermediate day length (12 hours light and 12 hours dark), and cells profiled at both dawn (lights on) and dusk

101   (lights off), to capture cell type dependent TOD-responsive gene expression. The dusk dataset comprises 4,327 cells

102   isolated across two replicates, which were further filtered to 3,151 cells of high quality (mean 1,333 genes/cell

103   detected, mean 3,907 UMIs/cell). This number approaches 1× coverage of all cells within this organism and is likely

4

104 to capture a broad spectrum of cell types and developmental transitions. Indeed, a total of 12,825 unique genes

105 were identified in the dataset, corresponding to 92 % of all annotated genes in the *Wolffia* genome (Michael et al.,

106 2020; Ernst et al., 2023). In addition, replicate libraries are highly congruent (Pearson correlation coefficient R

107 >0.99, Figures S2A-B), and pseudo-bulked scRNA-seq data correlates well (R >0.83) with a bulk RNA-seq dataset of

108 pooled *Wolffia* plants (Figure 1B, Michael et al., 2020), demonstrating that despite inevitable technical limitations

109 in capturing for example rare cells or large cells, the atlas provides a rich dataset of high quality.

110 The dusk atlas resolves as nine cell clusters (C0 to C8), quite distinctly presented as four major lobes in a UMAP

111 projection, referred to henceforth as 'Superclusters' CA to CD (Figure 1C, Table S1, Interactive 3D UMAP S1). To

112 understand the transcriptomic basis for this arrangement, we used differential gene expression (DEG) analysis to

113 first identify genes defining each of the four Superclusters (log$_2$ fold change >1 and adjusted p-value <0.05, Dataset

114 S1). Just over 500 differentially expressed genes (DEGs) were identified across all Superclusters. This number is

115 unexpectedly low when compared to expression variation revealed by scRNA-seq analysis across tissues in, for

116 example, the root, shoot, or leaf of other model species (e.g. Denyer et al., 2019; Kim et al., 2021; Zhang et al.,

117 2021). This distinction is even more apparent when considering DEGs expressed in over 10% of cells within a given

118 Supercluster (PCT1) and in fewer than 10% of remaining cells (PCT2), criteria commonly used to identify tissue

119 specific marker genes. Under these conditions we identified a total of 113 Supercluster marker genes (Dataset S1).

120 The lack of specification is likely in part explained by the limited structural complexity of the plant with no root or

121 vasculature, for example, but also points to a very streamlined system with the majority of genes dedicated to basic

122 cellular processes. The detection of 92% of all annotated genes in the single cell transcriptome datasets

123 corroborates this idea.

124 *Wolffia*, similar to other duckweeds, has a compact genome with fewer protein coding genes than most other plant

125 species (Michael et al., 2020; Harkess et al., 2021; Ernst et al., 2023). However, like most non-model genomes,

126 functional annotation of the predicted genes is sparse compared to model species such as *Arabidopsis*, maize, or

127 rice. Therefore, orthologues for Supercluster DEGs were identified from well characterized model species, focusing

128 primarily on *Arabidopsis*, in order to enable Gene Ontology (GO) term enrichment analysis and to discern

129 distinguishing functions for cells within each Supercluster (Datasets S1, S10). Accordingly, cells in CA are marked by

130   expression of genes predicted to function in wax biosynthesis and cuticle development, pointing to an epidermal

131   identity (Dataset S1). For example, genes encoding VERY-LONG-CHAIN ALDEHYDE DECARBONYLASE 3 (CER3) and

132   CER8, required for cuticle biosynthesis (Lü et al., 2009), are strongly expressed in cells of this cluster (Figure 1D), as

133   is WRI4 (WRINKLED4), encoding a transcription factor (TF) that promotes oil and wax biosynthesis (Park et al., 2016,

134   Figure S2C). In addition, the orthologue of CASPARIAN STRIP MEMBRANE DOMAIN PROTEIN 5 (CASP5) in

135   Arabidopsis, involved in the formation of lignified apoplastic barriers (Roppolo et al., 2011), is specifically expressed

136   in cells of this Supercluster (Figure 1D). Further, select basic helix loop helix (bHLH) and MYB TF genes linked to

137   drought stress responses are uniquely expressed in cells of the CA Supercluster. These include MYB41 and MYB49,

138   which are part of a regulatory circuit activated in response to desiccation, and MYB60, which has been linked to

139   stomatal closure under drought stress (Wang et al., 2021, Figure S2C).

140   Cells in Supercluster CC also host several DEGs with predicted functions in lipid biosynthesis and metabolism, but

141   these relate particularly to the sphingolipid class. Examples of this include the orthologue to SPHINGOID BASE

142   HYDROXYLASE 2 (SBH2, Figure 1D), as well as Wa8730a009g003220 and Wa8730a014g002170 whose orthologues

143   have been linked to sphingolipid biosynthesis in other species (Figure S2C, Dataset S10). Additional DEGs for this

144   Supercluster encode the aquaporin PIP1B and the iron transporter IRT1. Indeed, in addition to sphingolipids,

145   Supercluster CC DEGs are also enriched for transport linked GO terms (Figure S2C, Dataset S2). Curiously, two

146   markers are orthologous to YABBY2, which encodes a TF that promotes growth and abaxial cell fate in Arabidopsis

147   lateral organs (Siegfried et al., 1999, Figure 1D). Given the description of DEGs with roles in processes such as lipid

148   biosynthesis, it is likely that CC, like CA, comprises epidermal cells, albeit of a different fundamental nature. That

149   epidermal cells are well-distinguished from other major cell types is seen commonly in scRNA-seq studies of plant

150   cells (e.g., Kim et al., 2021; Zhang et al., 2021). In root atlases, an additional strong distinction is seen between hair-

151   and non-hair epidermal cells (e.g., Denyer et al., 2019; Shahan et al. 2022). The division seen here is likely of an

152   altogether different nature and the half-submergence of Wolffia may point to different functions needed for those

153   epidermal cells that interact with the two distinct environments of above or below the air/water interphase.

154   A majority of DEGs for Supercluster CB are associated with light responses and photosynthesis functions and

155   include many encoding light-harvesting chlorophyll (LHC) binding proteins (Dataset S1, Figures 1D, S2C). As such, CB

156 appears to comprise parenchyma cells of the photosynthetic heart of the plant. This Supercluster is further

157 characterized by the enriched expression of several orthologues in the GIBBERELLIC ACID-

158 STIMULATED *Arabidopsis* (GASA) gene family, which encode highly conserved cysteine-rich peptides involved in

159 redox sensing and hormonal responses (Bouteraa et al., 2023). Curiously, of all DEGs for this cluster, none are

160 orthologous to *Arabidopsis* TFs. This is in stark contrast to the other three Superclusters. In addition, TFs are

161 prominent among tissue- or cell type-specific genes in studies of other model plants (e.g. Knauer et al., 2019;

162 Denyer et al., 2019; Marand et al., 2021). It is conceivable that the strong photosynthetic signature of these

163 parenchymal cells may be driven primarily by environmental cues integrating into existing gene regulatory

164 networks.

165 Supercluster CD is close to CB in the cluster cloud (Figure 1C). However, based on the 180 DEGs distinguishing it,

166 Supercluster CD is less defined by photosynthesis. The proximity of this Supercluster to CB in the cluster cloud may

167 point to a commonality of their cellular identity (parenchyma). Conversely, the reduced photosynthesis-related

168 expression profile in CD points to a distinction similar to that predicted for the epidermis, such that Supercluster CD

169 captures the below water parenchymal cells. Indeed, the highly vacuolated cells that form the bulk size of the frond

170 would share features with their photosynthetically active counterpart but contain fewer chloroplasts as their

171 submerged position makes them less amenable to photosynthesis (Figure 1A). Further among DEGs distinguishing

172 cells in this Supercluster are orthologues of a cytochrome P450 and several WRKY and ethylene response factor

173 (ERF) TF genes, whereas GO terms enriched among DEGs are principally related to biotic- and abiotic stress

174 responses (Dataset S1). For example, homologues of pathogenesis-related (PR) proteins and WRKY60 are typically

175 pathogen-induced while *MYB36* responds to both biotic- and abiotic stress stimuli (Dong et al., 2003; Liu et al.,

176 2022; Figures 1D, S2C). In addition, genes in the ABA response pathway, e.g. *ABI1* and *AHG3,* implicated in both

177 stress-related responses and development are preferentially expressed in cells of CD (Dataset S1, Yoshida et al.,

178 2006; Pasaribu et al., 2023). Finally, a number of DEGs for this Supercluster have predicted functions in

179 development and morphogenesis, including in cell homeostasis, cell wall architecture/composition and

180 cytoskeleton organization, or encode developmental TFs. This latter finding is in line with the presence of

181 meristematic activity in a defined region within the underwater parenchymal cell population (Li et al., 2023).

182   These results indicate that *Wolffia* is largely comprised of cells with either an epidermal or parenchymal origin.

183   However, there is specialization within this broad classification as evidenced by distinct separations of

184   Superclusters and clusters in the atlas. A previous annotation of the *Wolffia* genome identified several gene

185   orthogroups specific to *Wolffia* and related duckweed species (Michael et al., 2020). Their predicted functions are

186   associated with just four GO terms including sphingolipid biosynthesis, photomorphogenesis, and wax biosynthesis.

187   It is of note that these terms are prominently reflected in the characteristics of three of the Superclusters, CC, CB

188   and CA, respectively. A fourth major GO term, 'cysteine-type endopeptidase', which ordinarily connects to the

189   processing of signaling peptides or antimicrobial peptides (AMPs) in the case of defense, does not define one

190   specific Supercluster, likely due to the broad spectrum of biological processes it affects. The fact that *Wolffia*-

191   specific orthogroups describe main Superclusters identified highlights these as the key adaptations to the partially

192   submerged environment in which *Wolffia* thrives.

193

194   **Submergence is a key distinguishing characteristic within epidermal and parenchyma cells**

195   A distinctive aspect of duckweeds such as *Wolffia* is that they live partly submerged, floating on the surface of

196   water bodies. As noted above, Superclusters CA and CC may distinguish the above- and within-water (a.k.a. top and

197   bottom) portions of the epidermis, whereas Superclusters CB and CD may comprise the top and bottom

198   parenchyma, respectively (Figure 1A). To further assess this possibility, we performed a bulk RNA-seq analysis on

199   *Wolffia* plants manually bisected into above and below water parts, discernable by the stark difference in

200   chlorophyll content and morphology (Figures 1A, S1D). In total, 262 DEGs (log_2 fold change >1, adjusted p-value

201   <0.05) between these regions were identified. Projecting average relative expression of these DEGs onto the

202   dataset UMAP showed they are overwhelmingly expressed in the epidermal clusters and map specifically to

203   Superclusters CA and CC, as expected (Figure 2A, Dataset S2). The epidermis of course interacts directly with the

204   water or air environment, but this finding predicts that this interaction triggers a strong local response that drives

205   distinctive transcriptome landscapes primarily in the epidermis.

206   Even though the above- versus within-water response is less evident in the parenchymal cells at the bulk RNA-seq

207   level, the single cell data provides a unique opportunity to discern tissue type-specific adaptations to being in an air

208   versus water environment. We therefore determined genes differentially expressed between cells of Supercluster

8

209   CB versus CD, as well as CA versus CC (Dataset S3). Confirming the above hypothesis, photosynthesis-related

210   functions associated with the chloroplast-rich, above water tissues are enriched in CB, whereas Supercluster CD

211   matches below water features. Notably, DEGs between these Superclusters are overwhelmingly upregulated in the

212   submerged cells (241 vs 38). This may be explained in part by the developmentally active cells of the meristem and

213   newly emerging daughter fronds in the submerged portion of the plant, but also indicates that parenchymal cells

214   primarily activate defined pathways upon submergence. Prominent among these are genes linked to responses to

215   abiotic stress stimuli, in line with the principal characteristic of Supercluster CD (Dataset S3). Numerous genes

216   involved in ABA signaling and osmotic stress responses are upregulated in the below water cells. Additional

217   signatures of the submergence response include a reduced response to GA, and the induction of ERFs. The latter

218   points to a highly conserved role for ethylene in the response to submergence in plants (Raskin et al., 1984; Fukao

219   et al., 2006). The submerged cells are further characterized by a heightened defense response, which seems fitting

220   given that the often-stagnant ponds in which *Wolffia* grow are likely to carry higher microbial loads and thus may

221   require heightened defense functions. Of note, there is no evidence for a switch from respiration to production of

222   energy through anaerobic pathways, such as ethanolic- or lactic acid fermentation, in the below water cells

223   (Dataset S3). This is perhaps explained by the relatively shallow submergence of *Wolffia* and the morphology of its

224   below water parenchyma, which has larger- and less tightly packed cells with additional air spaces to provide

225   increased buoyancy to the plant in the absence of aerenchyma (Bernard et al., 1990).

226   In contrast to the parenchyma, cells in both the top and bottom epidermis show environment-associated

227   specializations, with 168 and 109 upregulated DEGs in Supercluster CA and CC, respectively (Dataset S3). Enriched

228   GO terms for wax biosynthetic processes in the above-water epidermis (Supercluster CA, Figure 1D) point to the

229   top of the plant requiring additional hydrophobicity provided by the waxy cuticle to repel water and keep the

230   organism floating in the correct orientation, and to present a more hardened barrier against microbial and insect

231   pests (Borisjuk et al., 2018). The presence of cuticle exclusively on the above-water epidermis was confirmed by

232   Sudan IV staining, which indicates a fairly abrupt transition between top and bottom epidermis (Figure 2B). In

233   addition, the set of DEGs upregulated in the top epidermis indicates cells responding to desiccation (see above),

234   oxidative stress, and light stress (evidenced by induced expression of genes in the anthocyanin pathway (Kovinich

235   et al., 2015)), which mirror responses in epidermal cells of *Arabidopsis* leaves (Galvez-Valdivieso et al., 2009). In

236 contrast, the below-water epidermis is more involved in coordinating the transport of (micro-) nutrients from the

237 water environment into the plant and shows increased expression of a number of genes with distinct transport-

238 related and plasmodesmatal functions (Dataset S3). These cells also show aspects of the submergence response

239 seen in below water parenchyma, including upregulation of genes predicted to mediate abiotic and biotic stress

240 responses, although this expression signature is mostly distinct from that observed in the below water parenchyma

241 (Dataset S3). The most prominent and distinctive feature of the submerged epidermal cells is in the biosynthesis of

242 sphingolipids. The prominence of sphingolipid-related pathways in *Wolffia* has been suggested to indicate a trading

243 of terpenoids with sphingolipids for defense, perhaps because the aquatic environment favors the latter (Michael

244 et al., 2020).

245 Taking these analyses together, the principal divisions in the *Wolffia* scRNA-seq atlas appear to reflect the two

246 major tissue types, parenchyma and epidermis, divided by their relative location as the plant lives at the air/water

247 interphase (Figure 1A). Although these divisions are defined by relatively few DEGs, notable is the strong

248 submergence response evident in parenchymal cells at the scRNA-seq level, which was mostly undetected in bulk

249 transcriptomic analysis of above- and within-water regions. However, consistent with its direct interaction with the

250 plant's environment(s), the epidermis is particularly responsive to life in air versus water.

251

252 **Multiplex *in situ* RNA localization as validation of cell annotations**

253 To further validate the cell annotations and their interpretation, expression of select marker genes for the

254 Superclusters were imaged using PHYTOMap, a recently developed methodology for multiplexed spatial analysis of

255 transcripts within whole-mount plant tissues (Nobori et al., 2023). PHYTOMap provides an affordable alternative to

256 many spatial transcriptomic techniques and avoids the difficulties inherent with transforming *Wolffia* to produce

257 tissue-specific reporter lines, approaches commonly used for validation of scRNA-seq annotation (e.g., Denyer et

258 al., 2019; Zhu et al., 2023). Marker genes for Supercluster CA were found to strongly express at the above water

259 region of the plant, in the epidermis as expected (for example, Wa8730a016g000410, Wa8730a002g009150 and

260 Wa8730a010g002840, Figure 2C, Table S2). In contrast, marker genes for Supercluster CB are predominantly

261 expressed in the below water epidermis (e.g. Wa8730a011g003030), while transcripts for Wa8730a005g008250,

262 Wa8730a010g000100 and Wa8730a005g006050 are principally detected in the below water mesophyll, in line with

263   expression in cells of Supercluster CD (Figure 2C). Altogether, we examined twenty genes with varying degrees of

264   tissue specificity (Figures 2C-D, S3, Table S2). Of these, just four produced no defined data, largely due to the low-

265   resolution of, particularly, the magenta fluorescence channel. In addition, we observed a degree of

266   autofluorescence from the plant tissues largely in the green channel (Figure S3). However, despite this, almost all

267   genes selected showed expression in line with expectations. These results validate our annotation of the scRNA-seq

268   atlas and additionally highlight the utility of the PHYTOMap methodology for high-throughput examination of gene

269   expression patterns in organisms such as *Wolffia* that are less amenable to transformation.

270

271   **Cell clusters identify functional specialization within epidermal and parenchymal cells**

272   Obvious subdivisions are present within each Supercluster (Table S1). This points to distinctions beyond

273   transcriptomic differences stemming from life predominantly in water or air and suggests that epidermis- and

274   parenchyma-derived tissues each encompass several discernable cell types or states. To clarify these distinctions,

275   we identified DEGs describing each cluster by comparing its transcriptome to that of all other cells in the dataset

276   (Dataset S4). As noted above for the Superclusters, cluster distinctions reflect expression variation in comparatively

277   small sets of genes, and these were often assigned widely diverse cellular functions. Within the below water

278   epidermis, for example, clusters C0 and C8 share 39 out of a total of 147 DEGs, which indicates only a subtle

279   differentiation between these cell populations. Likewise, no cellular processes stand out as a defining characteristic

280   for cells in clusters C1 and C7, within the above water epidermal Supercluster. However, a distinction is seen in

281   cluster C3. Cells in this cluster show strong differential expression of genes with cuticle related functions (Dataset

282   S4). Indeed, orthologues of *CER* genes involved in cuticle biosynthesis are particularly-strongly expressed in C3

283   compared to other clusters (Figure 1D). Likewise, several *3-KETOACYL-COA SYNTHASE* (*KCS*) genes, as well as

284   *MYB31*, known to regulate cuticle biosynthesis in tomato (Xiong et al., 2020), are DEGs for C3, (Figures 1D, 3A).

285   Between C4 and C6, distinct clusters for the submerged parenchymal cells (Table S1), the cells in C6 are

286   distinguished by differential expression of a relatively large number (264) of genes. Although the only significantly

287   enriched GO terms for this cluster relate to functions in stress responses, among the DEGs are a substantial number

288   of orthologues for TFs and other genes with roles in plant development (Dataset S4). Examples of this include those

289   encoding NAC and LBD TFs, an auxin efflux carrier, a homolog to the receptor kinase CRINKLY4, several cell wall

11

290   biosynthesis and modifying enzymes, and VLN2 and VLN3, linked to directional growth in *Arabidopsis* (van der

291   Honing et al., 2012, Figure 3A). We note that generations of *Wolffia* were profiled together, capturing the spectrum

292   of developmental progressions; it is possible that this is resolved in the data, and the prominence of development

293   related DEGs here would indicate a notable concentration of less mature cells within C6. This point is explored in

294   more detail below.

295   As expected, the DEGs for clusters C2 and C5 in Supercluster CB primarily have functions relating to photosynthesis

296   and light responses (Dataset S4). Other marker genes for C5 point to functions in water transport, cell-to-cell

297   communication, and cytoskeletal reorganization. Of particular interest, a gene orthologous to *SUGARS WILL*

298   *EVENTUALLY BE EXPORTED TRANSPORTER* genes (*SWEET2* and *SWEET6*), is a prominent marker for a localized set

299   of cells in C5 (Figure 3A). Within vascular plants, these SWEET transporters are uniquely expressed in phloem-

300   associated cells to facilitate the source-sink distribution of photosynthates (Kim et al., 2021). Within the minimal

301   body plan of *Wolffia,* which lacks recognizable phloem cells (Figure 1A), photosynthates must be distributed across

302   cell types (Ware et al., 2023). Perhaps, the subset of *SWEET*-expressing cells could be acting as surrogates for

303   phloem cells to transport photosynthetically derived sugars to the rest of the *Wolffia* plant.

304   The cell specificity seen for the *SWEET2/6* gene orthologues prompted us to probe for additional examples of genes

305   with expression confined to a localized sub-population of cells in a cluster. To this end, we filtered the cluster DEGs

306   to those detected in fewer than 25% of cells in a given cluster (PCT1 <0.25), and fewer than 5% of cells outside of

307   that cluster (PCT2 <0.05). Scrutiny of UMAPs of these genes resulted in a list of 32 diverse DEGs showing highly

308   localized expression in one of five clusters (no such examples were found for C0, C1, C5 or C8, Dataset S4). C2

309   presented just one example. Expression of Wa8730a002g009940, the orthologue of *Arabidopsis KRATOS*, which

310   restricts cell death during stress and vasculature development (Escamez et al., 2019), is localized to the tip of

311   cluster C2 (Figure S5). Similarly, expression of many of the 'cuticle' genes is largely confined to cells within a sub-

312   region of C3 (Figure S5). Almost all mark the same strip of cells on the flank of the cluster. The localized expression

313   would indicate that the upper epidermis contains specialized cells to produce cuticular waxes. However, this region

314   is not exclusively defined by cuticle-related genes, and other examples with this localized expression include

315   orthologues for a cytochrome P450 and bHLH134 (Figure S5). The linear arrangement of this group of cells is

316 curious as it mirrors a dynamic commonly seen in scRNA-seq cluster projections that capture a developmental

317 trajectory (Denyer et al., 2019, Shahan et al., 2022). However, given the small number of genes involved in this

318 example, such a trajectory would likely reflect the differentiation of a specific specialized function, which we were

319 unable to resolve with confidence.

320 Within the same Supercluster, five genes mark two distinct regions within C7, four of which express in the same

321 subset of cells, while the expression pattern of the other is quite distinct (Figure S3). Their predicted functions hint

322 at localized signaling processes. Wa8730a010g003880, for example, is linked to oxidation of Brassinosteroids,

323 suggesting differential compositions of growth hormones across cells (Shimada et al., 2001). Most of the highly

324 localized gene expression was seen in C4 and C6. In the former, expression is preferentially localized to the tip of

325 the cluster, while in C6, a host of stress related genes including those linked to ABA and drought responses, defense

326 against fungi, DNA repair upon UV damage, as well as a core circadian clock gene *PSEUDO RESPONSE REGULATOR*

327 *5*, *PRR5*, are expressed in distinct sub-regions along the cluster (Nakamichi et al., 2005, Figure S3; Dataset S4). In

328 general, although the distinctive patterns of expression can be subtle, the strong localizations across the cluster

329 cloud point to true complexities within the clusters. The 'localized' genes also relate to a wide range of processes,

330 implying distinct specializations within cells with otherwise similar transcriptome profiles in what would appear to

331 be a highly streamlined organism.

332

333 **High conservation of guard cell transcriptomes, even at broad evolutionary distance**

334 Besides these observations, notable is the highly localized expression of Wa8730a007g001710 within a small group

335 of cells close to the periphery of C4 (Figure 3B). Its protein shares homology with FAMA, a TF that drives guard cell

336 formation in *Arabidopsis* (Smit and Bergmann, 2023). Consistent with this classification, several additional

337 *Arabidopsis* stomatal genes have a *Wolffia* orthologue showing strong, if not exclusive, expression in this small

338 subgroup of cells. This includes *EPIDERMAL PATTERNING FACTOR2* (*EPF2*), expressed in proliferating stomatal

339 meristemoids (Hunt and Gray, 2009), the guard cell maturation gene *DOF4.7* (*SCAP1*), as well as *SLAC1* and *MYB10,*

340 which regulate stomata opening and closing (Cominelli et al., 2005; Negi et al., 2013; Deng et al., 2021). However,

341 transcripts for orthologues to the early stomatal development genes *MUTE* and *SPEECHLESS*, which are known to

342 be more transiently expressed, are not found in these cells, nor elsewhere in the atlas, suggesting that the stomatal

343 cells captured are mostly mature in nature. The expression profiles of *FAMA*, *SCAP1*, *SLAC1*, and *MYB10* support

344 this idea, although the point is somewhat countered by the expression of the earlier-stomatal gene, *EPF2* (Hunt and

345 Gray, 2009; Adrian et al., 2015). However, it is possible that the latter signaling peptide may serve additional,

346 divergent roles in *Wolffia*, such as interaction with receptor-like kinases in the ERECTA gene family and receptor-

347 like proteins to regulate epidermal cell size (Meng et al., 2015).

348 Despite relatively few being isolated, the presence of guard cells in the atlas further underscores the depth of our

349 scRNA-seq dataset, since *Wolffia* plants are estimated to develop just around 30 stomata in their upper epidermis

350 (Figure 3C, Lam and Michael, 2022; Li et al., 2023). Their capture also provides an opportunity to assess

351 conservation of stomatal expression profiles. We therefore generated a transcriptome for the *Wolffia* guard cells

352 and compared this to an equivalent transcriptomic dataset derived from *Arabidopsis* stomata (Lopez-Anido et al.,

353 2021). Genes preferentially expressed in stomatal cells over other cells in the *Wolffia* atlas (Dataset S5) reveal a

354 general commonality with *Arabidopsis* stomata. Indeed, nearly 36% of *Wolffia* guard cell-associated DEGs have an

355 *Arabidopsis* orthologue expressed in the stomatal lineage (Dataset S5). This data points to high conservation

356 between *Wolffia*- and *Arabidopsis* guard cells, even at this evolutionary distance, which is in line with the early

357 origin and conserved function of this specialized cell type (Chen et al., 2016). Lastly, despite being epidermal in

358 origin and location, the *Wolffia* guard cells cluster more closely with the similarly photosynthetic parenchyma.

359 Given the streamlined developmental and environmental expression signatures observed thus far, gene expression

360 associated with photosynthesis in guard cells may override that of their epidermal origin.

361

362 **The *Wolffia* atlas captures developmental time between mother and daughter fronds**

363 *Wolffia* reproduces through budding, and in some species within this genus, this can happen almost daily. Central

364 to this, *Wolffia* plants contain a distinctive conical cavity, also called a "pocket", just below the upper parenchymal

365 layer toward one side of the frond (Figure 1A; Lam and Michael 2022). The below water parenchymal cells forming

366 the floor of this cavity are distinguished by an enlarged nucleus, electron dense cytoplasm and fewer plastids,

367 features characteristic of meristematic cells (Li et al. 2023). Their apparent asymmetric growth pattern suggests

368 that they continuously give rise to new primordia that develop into daughter fronds supported by a stipe (or

369 'branch', Li et al. 2023), which, by extension, pushes out the new plantlet at a distal 'exit'. This daughter will often

370 have initiated a granddaughter frond before it separates from the mother, taking this subsequent generation with

371 it (Figure 1A). As such, the below-water parenchymal cells may reveal biological distinctions reflecting the continual

372 production of daughter fronds, in addition to capturing the large vacuolated parenchyma cells that give the plant its

373 buoyancy.

374 We next sought to discern whether the developmental progression between daughter and mother fronds was

375 captured within our dataset. We first identified DEGs via bulk RNA-seq analysis on surgically separated mother and

376 daughter fronds (Dataset S6, Figure S4). Given the range of sizes that daughter fronds can take (Figures 1A, S4A),

377 the two samples would likely show strong commonalities. However, the daughter frond sample would be expected

378 to additionally contain more developmentally active cells undergoing rapid growth, cell division, and

379 differentiation. Indeed, we identified 114 genes in our dataset that show significant differential expression

380 between daughter and mother fronds. Among them, genes predicted to encode signaling components, including

381 via auxin and cytokinin, and cell wall modifying enzymes stand out. A 'gene set activity plot' (Seurat, Satija et al.,

382 2015) shows expression of daughter enriched DEGs primarily in regions of clusters 4, 6, and 8 (Figure 3D). As

383 expected, these clusters encompass the two major tissue types (Table S1), although proportionately more daughter

384 cells are of a seeming parenchymal origin (C6). This may in part have a technical basis, as smaller sized cells are

385 captured more efficiently in the scRNA-seq process over the ordinarily large below-water parenchymal cells.

386 Nonetheless, the pronounced daughter cell expression signature within C6 is in line with the aforementioned

387 differential expression of developmentally relevant genes in this cluster (Dataset S5). Its prominence particularly at

388 the tip of C6 may signify a developmental trajectory from the tip to the center of the cluster cloud that captures

389 maturation of daughter-to-mother cell states. However, at this time, putative trajectories and rare cells of the

390 meristem inside the pocket region could not be confidently distinguished.

391 *Wolffia* orthologues of *Arabidopsis* genes marking distinct phases of the cell cycle show generally broad expression

392 in cells across the atlas, suggesting that a daughter-to-mother state transition is not explained easily by cell cycle

393 activity (Dataset S11). Indeed, cell cycle related genes are not prominent among DEGs for any particular cluster.

394 However, there are some distinctions of note. Genes connected to the G1/G0 transition, for example, may form an

395 exception. Gene set activity plots of these genes show that epidermal cells are primarily captured in this more

396   quiescent phase of the cell cycle (Figure 3E). In addition, genes linked to mitosis show somewhat stronger

397   expression in cells at the tips of particularly C6, possibly in line with the daughter state (Figure 3E). However, cells

398   displaying S phase activity do not localize in one particular point of the UMAP, which may reflect a need for

399   genome endoreduplication in expanding cells.

400

401   **Time-of-Day sampling demonstrates a streamlined system incorporating cell type specific responses**

402   Aside from submergence, TOD is a critical environmental input to which cells in plants must respond (Steed et al.,

403   2021; Swift et al., 2022). In previous work, 13% of *Wolffia* genes show a robust TOD response, a percentage far

404   below that seen in other plant species (Filichkin et al., 2011; Ferrari et al., 2019; Michael et al., 2020; Michael

405   2022b). We reasoned that performing bulk RNA-seq on whole plants may mask TOD signals that are variable across

406   cell types, especially for a plant with such a simple body plan, and thus could significantly underestimate the true

407   number of genes under TOD control. As such, we generated an additional scRNA-seq atlas for plants collected at

408   dawn, twelve hours separated from the time analyzed above (dusk). The dawn dataset, after applying the same

409   filter parameters as for the dusk data, comprises 2,435 cells with a mean of 1,215 genes per cell and 12,565 genes

410   in total detected (mean 4,417 UMIs/cell). The two replicates within this atlas are also highly congruent (R >0.99,

411   Figure S4B).

412   The dawn dataset shows a largely similar cluster landscape to that observed at dusk, with eight clusters arranged

413   into four distinct Superclusters (Figure 4A, Dataset S8, Interactive 3D UMAP S2). Integration of the dawn and dusk

414   atlases into a joint UMAP shows that cells from the two TOD datasets form separate clusters that are near-mirror

415   images of each other (Figure 4B, Interactive 3D UMAP S3). While the dawn and dusk Superclusters clearly separate,

416   the mirroring suggests they represent similar cell types. Indeed, overall gene expression in cells of each

417   Supercluster is highly correlated between dawn and dusk (Figure 4C), and the DEGs characterizing the four

418   Superclusters show substantial overlap across the dawn and dusk TOD conditions (Figure 4D, Datasets S1 and S7).

419   As such, the defining characteristics of the four Superclusters, e.g. cuticle and sphingolipid biosynthesis, do not

420   change by TOD, except that in the chlorophyll-rich above water parenchyma (CB) light responses are much

421   enhanced within the dusk dataset, as might be expected following an extensive period of light (Sanchez and Kay,

422   2016; Oravec and Greenham, 2022). Accordingly, the following broad cell identities could be assigned to the dawn

16

423 Superclusters: CA, Above water epidermis; CC, Below water epidermis; CB, Above water parenchyma; CD, Below

424 water parenchyma (Table S1). The separation of the dawn and dusk clusters highlights that TOD plays a prominent

425 role in defining the transcriptome, and although this information does not override cell identity, sampling across

426 TOD provides a more comprehensive view of tissue type-dependent expression and identifies marker genes,

427 otherwise missed (Datasets S1 and S7).

428 To discern features of TOD regulation across *Wolffia* cells, Supercluster transcriptomes were compared across

429 dawn and dusk datasets (Figure 5A, Table S3). Even with a strict cut-off (Log$_2$ FC >1 and adjusted p value <0.05), a

430 considerable number of TOD-responsive DEGs were identified between equivalent cell populations: 364, 164, 157

431 and 157, for CA, CB, CC and CD, respectively (Dataset S9). The above water epidermis is the most TOD-responsive,

432 with considerably more DEGs between dawn and dusk than even the key photosynthetic cells of the parenchyma.

433 This would perhaps suggest a dynamic protective role of the epidermis to the change of light during the light/dark

434 cycle. Of the TOD responsive genes detected, 76 are TOD regulated in all four Superclusters. These 76 genes display

435 the same TOD expression pattern across each of the Superclusters, indicating that at least in *Wolffia,* some genes

436 show a coordinated TOD response across the organism regardless of cell type and growth environment (Dataset S9,

437 Figure 5A, Table S3). Similarly, genes that are TOD regulated in two or three Superclusters share either a day- or

438 night-dependent pattern of expression. Consistent with the findings from bulk RNA-seq (Michael et al., 2020), as

439 well as from other plant species (Ferrari et al., 2019), across the four main cell clusters, dawn DEGs are enriched for

440 GO terms relating to photosynthesis and light responses, while dusk DEGs are enriched for GO terms relating to

441 ribosomal- and polysome activity, which would point to increased translation late in the day (Supplemental Text,

442 Datasets S8-9, Dodd et al., 2005; Ferrari et al., 2019). Further, whereas the effects of light on photosynthesis and

443 plastid development are seen in all Superclusters, this signature is most prominent in cells of the above-water

444 tissues.

445 Among the TOD co-regulated genes are core components of the plant circadian clock, which forms a negative

446 feedback loop that interacts with an array of light signaling and developmental genes to control global TOD

447 expression in plants (Sanchez and Kay, 2016; Oravec and Greenham, 2022; Michael 2022a, Dataset S11). At its core,

448 the circadian clock is regulated by the SHAQKYF-type-MYB (sMYB) TFs, LATE ELONGATED HYPOCOTYL (LHY), and

17

449 REVEILLE (RVE), whose transcripts are highly expressed in cells of the dawn- but not the dusk-clusters, as expected

450 (Figure 5B; Oravec and Greenham, 2022). In contrast, the clock genes *GIGANTEA* (*GI*) and *FLAVIN-BINDING, KELCH*

451 *REPEAT FBOX* (*FKF1*) are primarily expressed in cells collected at dusk, as seen in bulk RNA-seq (Figure 5B). Further,

452 validating detection of TOD specific expression behaviors in the scRNA-seq datasets, *in situ* imaging of *RVE*

453 expression using PHYTOMap resulted in a strong expression signal only at dawn, as seen in other species as well as

454 in bulk RNA-seq (Figure 5C). Likewise, transcripts for Wa8730a010g005400, an orthologue of *Arabidopsis FATTY*

455 *ACID DESATURASE 2* (*FAD2*) identified as strongly upregulated in the dawn dataset, were preferentially detected in

456 morning samples, whereas the Wa8730a010g000190 expression signal was detected primarily at dusk, as predicted

457 (Figure 5C, Table S2). However, several 'core' circadian genes (Dataset S11) were found to be differentially

458 expressed in one or more clusters, suggesting that the circadian machinery may be distinct per tissue/cell type.

459 Orthologues for clock associated Pseudoresponse Regulators (PRRs) and for members of the ZEITLUPE (ZTL) family

460 (Wa8730a009g001000, Wa8730a003g007690), for example, cycle exclusively in CA and CA and CB, respectively

461 (Dataset S9, Somers et al., 2000; Para et al., 2007). This finding that both core circadian genes as well as diurnal-

462 regulated genes can show cell type specificity is consistent with results from both a different duckweed system as

463 well as from a range of other plant species (Watanabe et al., 2021; Endo, 2016; Swift et al., 2022).

464 Of the 456 TOD responsive genes detected by scRNA-seq, 289 (> 60%) do not overlap with high confidence 'cycling'

465 genes detected by bulk RNA-seq (Michael et al., 2020, Cutoff R >0.8). The considerable excess of TOD-responsive

466 genes identified by scRNA-seq, highlights the value of this approach for detecting tissue-dependent environmental

467 responses (Jean-Baptiste et al., 2019; Shulse et al., 2019). On an individual basis, 233, 60, 68 and 58 DEGs from CA,

468 CB, CC and CD, respectively, were not previously detected at the bulk RNA-seq level. Beyond these novel TOD

469 regulated genes, each Supercluster harbors 195, 22, 32 and 26 DEGs, respectively, that are specific to that

470 Supercluster (Figure 5A; Dataset S9), demonstrating responses to TOD input that are not only tissue-type specific

471 but interact uniquely with environmental cues stemming from life within or above water (Greenham et al., 2017).

472 These Supercluster-specific TOD responses can be considerable (e.g. Figure 5D). A prominent example is

473 Wa8730a008g003780 (*IRT2*), which unlike the aforementioned *IRT1*-orthologue (Figure S1C), is expressed

474 predominantly in the above-water epidermis at dawn (Figure 5E). *BROAD-RANGE SUGAR PHOSPHATE*

475 *PHOSPHATASE* (*SGPP*, Wa8730a011g001940) shows a similar expression profile, while Wa8730a017g000040

476 displays the opposite behavior in the epidermis (Figure 5F). Other genes showing strong distinctions in this regard

477 include the dawn-upregulated orthologues of *Arabidopsis ALPHA CARBONIC ANHYDRASE1* (*ACA1*) and *ACA3*

478 (Wa8730a020g000810 and Wa8730a012g002610), specifically cycling in the above- and within- water epidermis,

479 respectively (Figure 5F). The epidermis specific expression of these ACA genes at dawn highlights the important role

480 this enzyme plays in facilitating carbon capture from the air and water environments through conversion of CO_2

481 into carbonic acid. Other DEGs in the above-water epidermis at dawn are associated with stress and temperature

482 responses, whereas the below-water epidermis shows regulation of the response to ultraviolet light (Supplemental

483 Text, Datasets S8-9).

484 Fewer genes cycle specifically in the parenchyma, though Wa8730a011g002740 stands out in that it is upregulated

485 in the below-water parenchyma at dawn (Figure 5F). Also the karrikin response appears to be TOD regulated in this

486 way, whereas the above-water parenchyma has DEGs with functions in various metabolic and biosynthetic

487 pathways at the same TOD (Dataset S9). Importantly, aside from these select examples, the genes that are TOD

488 regulated in individual Superclusters have diverse functions indicating TOD effects on a range of processes, rather

489 than explicit up- or downregulation of overt regulatory networks.

490 Taken together, this single cell level analysis of TOD responsive gene expression is consistent with the relatively

491 streamlined nature of *Wolffia* gene regulation. It also highlights the benefits of the methodology for discerning

492 expression signatures distinguishing cell types and their responses to environmental cues, as well as demonstrating

493 the considerable cell type-specific changes in this regard. Indeed, while select biological processes are generally

494 controlled in a TOD fashion as expected (Supplemental Text, Datasets S8-9), many genes cycle in a manner

495 dependent on cell type and/or life in water versus air. These are associated with widely diverse cellular functions

496 suggesting a pleiotropic response (Supplemental Text, Datasets S8-9). That the above-water epidermis is the most

497 responsive of all cell types was unexpected and might form a feature linked to the very specific habitat and lifestyle

498 of *Wolffia*.

499

500

501

502 **DISCUSSION**

503 *Wolffia* holds great promise as a key model organism for the exploration of manifold biological questions. The

504 experimental and practical opportunities offered by *Wolffia* as a biofuel crop and novel food, for example, are in no

505 small way connected to its basic body plan, small size, and rapid regeneration (Lam and Michael, 2022). Further, its

506 compact genome offers opportunities to discover basic cellular processes defining cell types and how these

507 respond to environmental change. In particular, the half-submerged lifestyle of *Wolffia* offers a novel opportunity

508 to examine how a plant responds to critical environmental inputs stemming from life in air versus water. We

509 examined this, and the impact of TOD, on individual cell types of the entire plant using single cell resolution

510 transcriptomics. The division of *Wolffia* cells into two major tissue types, specialized by the submerged water

511 environment, demonstrates the most fundamental divisions of this minimalist plant. Transcriptomic comparisons

512 between these cell populations reveals a very streamlined organism with relatively few DEGs between cell types,

513 submerged and above water tissues, mother and daughter fronds, or across different TOD. As we detected

514 transcripts for most annotated genes in the genome, we can surmise that the majority are broadly and evenly

515 expressed 'housekeeping' genes. Nonetheless, DEGs between these various divisions identify basic cellular

516 functions underlying the distinct biology of these tissue types. Most notably, we see that gene orthogroups specific

517 to the duckweeds are linked to specific tissue-dependent adaptations for a species that floats on water.

518 Beyond the fundamental tissue type divisions, the atlas captures additional cellular complexity and specialization,

519 including cells with a strong developmental signature within the below water parenchyma. Developmental activity

520 is clear from the abundant expression of TFs and other key genes underlying patterning and growth, as well as a

521 strong daughter frond and mitotic "signal", particularly in cluster C6. Its cells likely include those of the youngest

522 daughter fronds, and perhaps the rare cells of the meristem and stipe. Further, expression dynamics across C6

523 seemingly points to the capture of a developmental progression within the cluster structure. While further scrutiny

524 of putative trajectories and rare cell types would benefit from increased sequencing depth and cell numbers, the

525 current data offers a first opportunity to explore development in this minimalist plant, unique as it is in the rapidity

526 of its replication and growth.

20

527 Aside from developmental progression, the atlas identifies transcriptomes for guard cells of the stomata and

528 particular epidermal cells with a penchant for cuticle production. In addition, we see evidence for potential sub-

529 specializations that could compensate for the simple body plan of *Wolffia*, relative to other plants. The highly

530 localized expression of SWEET transporters within the parenchyma, for example, potentially compensates for the

531 lack of a vascular system. Similarly, defence genes in cells of the below water parenchyma and epidermis point to

532 an acute response to potential waterborne antagonists. The stomatal transcriptome compiled from our data

533 highlights the conservation of certain cell types across species as distantly related as the model eudicot *Arabidopsis*

534 and this non-grass monocot. However, it also exposes the power of looking at new model systems representing

535 unique biology and positions in the green lineage, with the *Wolffia* Superclusters displaying unique gene networks

536 adapted to the aquatic environment.

537 Our analysis further highlights the importance of taking TOD responses into consideration when developing an

538 understanding of the systems level organization of cells, as well as distinct aspects of networks resulting from the

539 primary driver of plant biology, the sun. The observed TOD expression changes predict, that the clock changes from

540 tissue type to tissue type, even in this minimal plant system as in other systems (Endo, 2016; Swift et al., 2022), and

541 that responses do so as well. Photosynthesis, plastid development, and translation are major processes TOD

542 regulated across the plant, but many genes respond to TOD in a manner dependent on cell type and life in water

543 versus air. These genes are predicted to have widely diverse functions suggesting effects of TOD on a broad range

544 of specialized cellular processes. In particular, the above water-epidermis is considerably more responsive to TOD

545 than other cell types. This could be a consequence of this cell type being at the vanguard of the changing light and

546 highlights the relative prominence of local protective measures to this environmental cue.

547 Despite its virtues, a limitation of *Wolffia* as a model arises from current challenges in its routine and rapid

548 transformation for molecular analyses. Therefore, to validate our cluster annotation, a critical element of scRNA-

549 seq studies, we utilized PHYTOMap, a recently developed, multiplexed, *in situ* hybridisation methodology (Nobori

550 et al., 2023). Besides the limitation of select fluorescence channels to vividly capture rarer- or lower-expressed

551 transcripts, this proved highly successful, offering a practical option going forward to examine spatial patterns of

552 expression in model organisms like *Wolffia* that are less amenable to analysis of transgenic reporters.

21

553   Global studies of gene regulation at the systems level in multicellular organisms are extremely complex due to, in

554   part, the intricate networks of TF activity and the myriad tissue types involved. Notwithstanding the

555   aforementioned limitations of sequencing depth in resolving the rarest dynamics within this species, the

556   streamlined gene expression variation identified in this work points to *Wolffia* being a tractable model for studies

557   of regulatory dynamics at the organismal level with examination of chromatin dynamics at single cell resolution, for

558   example, as an attractive next step. This model equally offers amenable opportunities for integrating metabolic- or

559   proteomic examination, utilizing the amalgamation of nascent technologies and methodologies (Clark et al., 2021;

560   Zhao and Rhee, 2022). Such complex, layered data transposed onto an organism of such elementary structure, but

561   harbouring fascinating generational dynamics, and a unique sub-aquatic lifestyle convenient for environmental

562   studies, offers an opportunity to examine complex dynamics. The atlas we present here, of 5,604 high quality cells,

563   broadly representative of the entire organism, offers a strong reference to explore a wealth of fundamental

564   questions going forward. Key curiosities in this regard are the differences between above- and below-water-

565   situated tissues, and the varying roles the different tissues undertake in these contexts. Particularly interesting,

566   perhaps, are the variable defense mechanisms of the below water parenchyma and epidermis, and the differential

567   responses to TOD. Similarly, there is clearly more to be learned regarding the unique development of *Wolffia*. To

568   this end, the atlas can be easily explored on our interactive browser at https://www.zmbp-resources.uni-

569   tuebingen.de/timmermans/plant-single-cell-browser/ (Ma et al., 2020).

570

571   **METHODS**

572   **Plant growth**

573   *W. australiana* line 8730 (Australia, New South Wales) was used for all experiments. This line was chosen due to it

574   having the best genome assembly available at the time, and for the availability of bulk RNA-seq TOD time course

575   data (Michael et al., 2020). Plants were grown at intermediate days of 12 h light (100 µE)/12 h dark cycles at a

576   constant 24 °C in 100 ml nutrient medium (0.5 × Schenk & Hildebrandt (Duchefa), 0.1% sucrose, pH 6.7). Plants

577   were sub-cultured weekly by transferring ±10 fronds to fresh medium.

578

22

579 **Protoplast isolation for scRNA-seq library construction**

580 Approximately 200-300 plants, less than one-month old, were collected for each sample. Dusk and dawn samples

581 were collected in parallel from separate growth chambers with contrasting light/dark cycles. Dusk cycles were

582 collected at the end of the light period (lights off, Zeitgeber (ZT) 12) and dawn cycles at the end of the dark period

583 (lights on, ZT0). To check TOD dynamics were retained despite this, gene sets previously shown to cycle at

584 particular times in *Wolffia* were assessed in the resultant data (Michael et al., 2020, Figure S5C). To avoid problems

585 associated with the thick, waterproof cuticle of the plants, individual samples were diced with a razor blade to

586 improve access the innermost cells. Diced tissue samples were placed into a freshly made 'protoplast enzyme mix'

587 with a higher concentration of commonly used cell wall digestion enzymes (0.1 M KCl, 0.02 M $MgCl_2$, 0.1% BSA

588 (Sigma Aldrich), 0.08 M MES (Duchefa), 0.6 M Mannitol (Duchefa), pH 5.5, adjusted with Tris, 1.5% Cellulase R-10,

589 1% Maceroenzyme, 0.5% Pectolyase; all enzymes Duchefa). Tissue samples were digested on an orbital shaker set

590 at < 200 rpm for maximum 90 minutes to minimize transcriptional noise. Dawn samples were kept in darkness for

591 this step. Digested cells were passed through a 100 μm sieve followed by a 40 μm cell strainer (pluriSelect) and

592 resuspended in 10 ml 'wash buffer' ('protoplast enzyme mix' without enzymes) and centrifuged (1,000 × G, 20 °C).

593 The supernatant was then removed, and the pellet resuspended in 10 ml wash buffer. This was centrifuged again

594 (500 × G, 20 °C), the supernatant removed, and the pellet resuspended in 500 μl wash buffer. By necessity, wash

595 steps for both samples were performed in light with the whole process minimized to 30 minutes. Protoplasts were

596 validated under a light microscope and quantified using a haemocytometer. Concentrations were adjusted with

597 wash buffer to a density of approximately 800-900 cells per ml.

598

599 **Single cell RNA-seq library preparation and sequencing**

600 Single cell RNA-seq libraries were prepared from fresh protoplasts according to the 10x Genomics Single Cell 3'

601 Reagent Kit v 3.1 protocol. Cells were loaded onto the 10x Chromium controller within 1 H of the end of digestion.

602 ~60% excess cells over target were added, as per the 10x protocol. 11 cycles were used for cDNA amplification, and

603 12 for final PCR amplification of the adapter-ligated libraries. Final library size and quality was checked on a DNA

604 High Sensitivity Bioanalyzer chip (Agilent), and libraries were quantified using the NEBNext Library Quantification

605 Kit for Illumina and sequenced to a depth of >20,000 reads per cell.

23

606

607 **RNA-seq library preparation and sequencing**

608 For bulk RNA-seq comparisons of above- and below-water tissues, and mother-daughter fronds, plants were

609 bisected with needles under a light microscope and appropriate tissues collected in liquid nitrogen and ground to a

610 fine powder. RNA was extracted using TRIzol (Thermo Fisher Scientific), following the manufacturer's protocol. RNA

611 samples were quantified by NanoDrop, and quality assured based on RNA Bioanalyzer chip traces (Agilent). mRNA

612 was enriched by Oligo(dT) pull-down using the NEBNext Poly(A) mRNA Magnetic Isolation Module, and RNA

613 libraries constructed using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina with NEB Multiplex

614 oligos. Final library size and quality was checked on a DNA High Sensitivity Bioanalyzer chip, and libraries were

615 quantified using the NEBNext Library Quantification Kit for Illumina and sequenced to a depth of 10M reads per

616 sample.

617

618 **PHYTOMap sample preparation and image capture**

619 *Wolffia* was grown under 12 h/12 h light cycles, with lights on at 7:00 AM and off at 7:00 PM. Live fronds were used

620 for the experiment. The dawn set was harvested at 8:20 AM and the dusk set at 5:00 PM, immediately starting the

621 initial FAA fixation step upon harvesting. The protocol in Nobori et al., 2023 was followed with some modifications

622 to accommodate increased volume of sample required for each set of probes used, as well as increased incubation

623 times to allow solutions to fully penetrate the fronds. Yakult enzymes in the Cell Wall Digestions Steps were

624 substituted with cellulase (Sigma cat. No. SAE0020), macerozyme (PlantMedia cat. No. 21560017-1), and pectinase

625 (Sigma cat. No. P2401-1KU). Additionally, the fronds did not stick to poly-D-lysine coated dishes, so the tissue was

626 processed in 1.5 ml tubes for all incubation steps. The SNAIL probe concentration was also increased from 10 nM in

627 the original protocol to 500 nM per oligo in the final pool to account for the larger volume of *Wolffia* fronds.

628 Samples were prepared for imaging on microscope slides with a 22 mm × 50 mm coverglass. The additional size of

629 the coverglass allowed better adhesion to the slide. Smaller coverslips were more easily disrupted due to the

630 thickness of the duckweed fronds. Imaging was performed on an Olympus FV3000 confocal microscope. Images

631 were taken using a 20 × zoom with image resolution of 254 × 254 pixels per tile. Four tiled images per frond were

632 taken and stitched together for a total of 1019 × 1019-pixel image sizes to encompass a whole frond in one image.

633 Z-stacks were generated for each set of tiles to span the entire depth of each frond imaged, number of layers

634 dependent on frond thickness. The following channel settings were used for each fluorophore: AF488, 499 nm

635 excitation, 504-554 emission; Cy3, 554 nm excitation, 559-650 emission; Cy5, 649 nm excitation, 657-735nm

636 emission; and AF740, 752 nm excitation, 760-839 nm emission.

637

638 **Histological staining**

639 *Wolffia* plants were embedded in 7% agarose and sectioned to a depth of 75 µM by Vibratome (Leica). Staining was

640 performed with 0.1% Sudan IV (Sigma) for ~10 minutes, as per Nadiminti et al., 2015.

641

642 **Bulk RNA-seq analysis**

643 Sequence reads (Paired End, 150 bp) were aligned to the *W. australiana* line 8730 (Wa8730.asm201904v2.fasta)

644 reference genome with STAR (version 2.7.10a; Ernst et al., 2023; Dobin et al., 2013). All common names and

645 accession of genes referred to in the manuscript can be found in Dataset S12. Read counts of genes were calculated

646 on uniquely mapped reads using 'featureCounts' (version 2.0.1; Liao et al., 2014) with *W. australiana* annotation

647 (Wa8730.gtf). Differential expression (DE) analysis was conducted using DESeq2 (version 1.34.0; Love et al., 2014).

648 Differential expressed genes (DEGs) were identified by having a Log_2 fold-change > 1 and an adjusted p-value <

649 0.05. For correlation analysis of gene expression between bulk RNA-seq and scRNA-seq, Log_2 (mean RPM+1)

650 expression values were calculated for each gene and the Pearson correlation coefficient determined in R (v4.1.2; R

651 Core Team 2021).

652

653 **Generation of single cell expression matrices**

654 Cell Ranger version (5.0.1) (10x Genomics) was used for initial scRNA-seq data processing (default settings).

655 CellRanger Count was used to align sequencing reads to the *W. australiana* reference genome. Cells were filtered

656 with a quality cut-off of > 650 UMIs/cell. Gene expression matrices were created for each readset (dusk-1, dusk-2,

657 dawn-1, dawn-2) and readsets were merged with cellranger aggr. The final .h5 file was converted to a csv

658 expression matrix via cellranger mat2csv and used for further analysis. See Table S4 for further metrics pertaining

659 to scRNA-seq data analysis and quality.

660

661 **Dimensionality reduction, UMAP visualization, and cell clustering analysis**

662 All standard bioinformatic analyses were performed using R Statistical Software (v4.1.2; R Core Team 2021). The

663 Seurat R package (version 4.3.0, Satija et al., 2015), was used for downstream analysis with default parameters

664 unless otherwise specified. Gene expression values were normalized with 'LogNormalize' function. The top 2000

665 highly variable genes were identified with 'FindVariableFeatures' and used for PCA calculation. First 10 PCs were

666 used for Louvain clustering of the cells and UMAP dimensionality reduction. Interactive 3D UMAPs for the dusk-,

667 dawn-, and combined datasets can be found in Supplemental materials. Superclusters were manually assigned by

668 analyzing UMAP visualizations. Select clusters were sub-clustered using 50 PCs. Pearson's correlation of gene

669 expression between the dusk and dawn was calculated on genes expressed in > 10 % of cells in both superclusters.

670 DEGs were identified with 'FindAllMarkers' with Log_2 Fold change threshold > 1, adjusted p-value < 0.05 and PCT1

671 (percentage of cells) >10%. Gene set activity plots were calculated using 'AddModuleScore'. For integration and

672 comparison of dusk and dawn datasets, Seurat objects were generated independently and then integrated using a

673 Seurat integration pipeline using 2000 anchor features and the canonical correlation analysis (CCA) integration

674 method. After integration, 20 PCs were used as input features for the 'RunUMAP' and 'FindNeighbors' functions.

675 TOD DEGs between dusk and dawn of each supercluster was performed with 'FindMarker' using integrated data

676 with Log_2 fold-change > 1, adjusted p-value < 0.05 and PCT >10 % in either PCT1 or PCT2.

677

678 **Stomatal transcriptome analysis**

679 The cluster of cells expressing the stomatal marker Wa8730a007g001710 (AT3G24140, FAMA) was manually

680 selected. Differential expression analysis between stomatal cells and remaining cells in the dusk/dawn integrated

681 data object was conducted using 'FindMarkers' with the same parameters as above.

682

683 **Gene Ontology (GO) enrichment analysis**

684 Gene ontology (GO) term overrepresentation analysis was carried using the *Wolffia* gene annotation and

685 GOATOOLS (Klopfenstein et al., 2018). Significant GO terms (P >0.05) were identified from specific gene lists with

686 the flags (--pval=0.05 --method=fdr_bh --pval_field=fdr_bh).

26

687

688 **Identification of *Wolffia* orthologs**

689 Orthology was determined with OrthoFinder (Emms and Kelly, 2019). Marker genes from model species based on

690 existing literature were used to search the ortho-groups for lm5633 orthologues. These marker genes in model

691 organisms were compared with the OrthoFinder gene table and a corresponding lm5633 orthologue was observed.

692 Generally, annotations with a geneID from eggnog mapper were more reliable than observing marker genes in

693 larger gene families (>3 copies per genomes).

694

695 **DATA ACCESS**

696 All raw and processed sequencing data generated in this study is available through the NCBI short read archive

697 (https://www.ncbi.nlm.nih.gov/sra/), as well as under the NCBI BioProject PRJNA1124135. Please see Supplemental

698 Table S5 for individual SRA accession numbers. The scRNA-seq atlas can be easily explored via the interactive

699 browser at https://www.zmbp-resources.uni-tuebingen.de/timmermans/plant-single-cell-browser/. The scRNA-seq

700 Seurat objects and associated matrices are accessible via the same Browser

701

702 **COMPETING INTEREST STATEMENT**

703 The authors declare no conflict of interest.

704

705 **ACKNOWLEDGMENTS**

708

709 **AUTHOR CONTRIBUTIONS**

710   E.L., T.P.M., M.C.P.T. and T.D. conceived and designed the study. T.D. produced all libraries and performed all wet

711   lab experimentation, except for the PHYTOMap analysis which was carried out by K.C., A.M. and T.N.. K.C. and

712   B.W.A. carried out sequencing and performed an initial analysis of the dusk single cell data. P-J.W. performed the

713   single cell and bulk RNA-seq data analyses. T.P.M. performed the OrthoFinder and TOD analyses. P.S. generated the

714   single cell browser interface for accommodating the *Wolffia* single cell atlases. T.D. and M.C.P.T wrote the

715   manuscript with contributions from E.L. and T.P.M. who also proofread the final version.

716

726

727   **FIGURE LEGENDS**

728   **Figure 1.** A scRNA-seq atlas of *Wolffia australiana* cells reveals two core cell types optimized for life in air versus

729   water. (A) *Wolffia australiana* comprises a mother (M) frond from which successive daughter (D) fronds bud off.

730   Within the plant, a granddaughter (GD) frond is often also seen. A developmental progression of a daughter frond

731   (anticlockwise from top left) is shown on the bottom right. The plant can generally be divided into epidermal and

732   parenchymal cells, with stomata dispersed across the above water epidermis, and the parenchyma divided into

733   photosynthetic palisade parenchyma at the top of the plant, and spongy parenchyma in the hull (top right). A

734   conical cavity is seen near the center of the frond below the surface of the water. (B) Pearson correlation analysis

28

shows that gene expression values in merged single cell and bulk-tissue RNA-seq of whole *Wolffia* plants are highly

correlated. (C) UMAP visualization of the dusk scRNA-seq atlas reveals nine distinct cell clusters, organized around

four Superclusters (inset) corresponding to above water epidermis (CA), below water epidermis (CC), above water

parenchyma (CB), and below water parenchyma (CD). (D) UMAP projections of normalized expression profiles of

select DEGs within the atlas showing cluster and Supercluster specificity. Names of *Arabidopsis* orthologues are

given (see Dataset S10).

Figure 2. Tissue type specific responses to life in air versus water are validated by PHYTOMap multiplex *in situ* RNA

detection. (A) UMAP projections of 'gene set activity' (Seurat) profiles of above- and below-water enriched DEGs as

determined by bulk RNA-seq show the response to submerge is primarily detected in the epidermis. (B) Histological

sections of *Wolffia* stained with Sudan IV shows stained cuticle (purple arrowhead) on the exterior of the above-

water epidermis (top) that is lacking from the below-water epidermis (bottom). Blue arrowhead: approximate

position of the waterline on the flank of the plant. (C) Select Supercluster markers imaged using PHYTOMap show

strong cell type specificity in line with scRNA-seq predictions. Four genes are imaged in the same plant in four

fluorescence channels. From left-to-right: combined expression profiles, images in the blue, red, and green, and

magenta channels. White dashed lines: approximate water line; §+: identical exterior or interior sections across

different imaging channels. (D) PHYTOMap images showing the combined expression profiles of the genes

indicated below. Composite images are a projection of 13-15 z-stack sections of 2-3 μm each. Information on genes

examined can be found in Table S2 and Supplemental Figure S2. Scale bar, 250 μm. Where available, names of

Arabidopsis orthologues are given (see Dataset S1).

Figure 3. Specialized cell types and developmental stages are distinguished within major tissue type divisions. (A)

UMAP projections showing normalized expression profiles of select DEGs within the atlas showing cluster

specificity. Names of *Arabidopsis* orthologues are given (Dataset S10). (B) UMAP showing expression of

Wa8730a007g001710, orthologue of *Arabidopsis* FAMA, is concentrated in a small subset of cells within

Supercluster CD. (C) Stomata, marked by autofluorescence (green), are dispersed across the upper, above water,

epidermis. White arrows highlight select guard cells. Scale bar, 250 μm. (D) UMAP visualization of 'gene set activity'

(Seurat) of daughter-enriched DEGs determined by bulk RNA-seq shows that developmentally active daughter cells

29

761 are localized primarily within specific clusters, particularly within a region of C6. (E) UMAP visualizations of 'gene

762 set activity' of genes associated with distinct phases of the cell cycle (Dataset S11) show that while cells undergoing

763 DNA replication are largely dispersed across the atlas, mitotic cells primarily localize to C6, and to a lesser extent,

764 C4, mirroring the distribution of daughter cells.

765 **Figure 4.** Dusk and dawn transcriptomes are well-conserved though many genes show TOD responses. (A) UMAP

766 visualization of the dawn scRNA-seq atlas reveals eight distinct clusters, organized around four Superclusters

767 (inset). (B) Integration of dusk and dawn datasets reveals a mirror image arrangement of Superclusters, indicating a

768 strong conditional transcriptome distinction that do not override cell identity. (C) Gene expression correlation

769 between dusk and dawn replicates of distinct Superclusters reveals high correlation. Genes expressed in more than

770 10% of the cells in each Supercluster were used in each instance. (D) Venn diagrams showing just a partial

771 conservation of DEGs between Superclusters of dusk and dawn conditions indicating that sampling across TOD

772 identifies unique tissue-specific DEGs. Identies of Superclusters CA, CB, CC and CD are matched across datasets.

773 **Figure 5.** TOD responses vary depending on cell type and on life in water versus air. (A) Venn diagram illustrating

774 the overlap of TOD DEGs across Superclusters, 76 genes are TOD regulated in all four Supercluster comparisons, but

775 many genes show a tissue-dependent TOD response. (B) UMAP projections of normalized expression for the core

776 circadian clock genes *REVEILLE* (*RVE*), *GIGANTEA* (*GI*), *FLAVIN-BINDING, KELCH REPEAT, FBOX* (*FKF1*), and *LATE*

777 *ELONGATED HYPOCOTYL* (*LHY*), reveal the expected strong TOD responses. (C) Expression of select dawn or dusk

778 markers imaged at dawn (top) and dusk (bottom) using PHYTOMap show strong preferential expression at dawn

779 (left two), or dusk (right). Composite images are projections of 13-15 z-stack sections of 2-3 μm each. Information

780 on genes examined can be found in Table S2. Scale bars, 250 μm. (D) UMAP projections of example genes showing

781 differential expression (DE) between dusk and dawn conditions as well as a level of cell-type specificity. Expression

782 for each gene is plotted on an integrated plot of datasets: a - Wa8730a017g001020, b - Wa8730a016g003000, c -

783 Wa8730a017g003930, d - Wa8730a005g008830, e - Wa8730a006g006550, f - Wa8730a003g006530, g -

784 Wa8730a006g005980, h - Wa8730a019g001460, i - Wa8730a020g001730, j - Wa8730a010g005400, k -

785 Wa8730a012g004550, l - Wa8730a002g000180, m - Wa8730a006g005670, n - Wa8730a003g004150, o -

786 Wa8730a002g007540, p - Wa8730a002g007630, q - Wa8730a015g001450, r - Wa8730a018g000010, s -

787　Wa8730a002g010730, t - Wa8730a010g003830. (E) UMAP projection of IRT2 (Wa8730a011g002740) expression in

788　the integrated dusk/dawn atlas showing a strong tissue-specific TOD response. (F) UMAP projections of further

789　genes showing tissue specific TOD responses. *SGPP* - 8730a011g001940, *ACA1* - Wa8730a020g000810, *ACA3* -

790　Wa8730a017g000040.

791

792　**REFERENCES**

793　Adrian J, Chang J, Ballenger CE, Bargmann BO, Alassimone J, Davies KA, Lau OS, Matos JL, Hachez C, Lanctot A, et al.

794　　2015. Transcriptome dynamics of the stomatal lineage: birth, amplification, and termination of a self-renewing

795　　population. *Developmental Cell* **33**, 107–118.

796　Bernard FA, Bernard JM, Denny P. 1990. Flower structure, anatomy and life history of *Wolffia australiana* (Benth.)

797　　den Hartog & van der Plas. *Bulletin of Torrey Botanical Club* **117**, 18-26.

798　Bezrutczyk M, Zöllner NR, Kruse CPS, Hartwig T, Lautwein T, Köhrer K, Frommer WB, Kim JY. 2021. Evidence for

799　　phloem loading via the abaxial bundle sheath cells in maize leaves. *The Plant Cell* **33**, 531–547.

800　Bläsing OE, Gibon Y, Günther M, Höhne M, Morcuende R, Osuna D, Thimm O, Usadel B, Scheible WR, Stitt M. 2005.

801　　Sugars and circadian regulation make major contributions to the global regulation of diurnal gene expression in

802　　*Arabidopsis*. *Plant Cell* **17**, 3257-81.

803　Borisjuk N, Peterson AA, Lv J, Qu G, Luo Q, Shi L, Chen G, Kishchenko O, Zhou Y, Shi J. 2018. Structural and

804　　biochemical properties of duckweed surface cuticle. *Frontiers In Chemistry* **6**, 317.

805　Bouteraa MT, Ben Romdhane W, Baazaoui N, Alfaifi MY, Chouaibi Y, Ben Akacha B, Ben Hsouna A, Kačániová M,

806　　Ćavar Zeljković S, Garzoli S, Ben Saad R. 2023. GASA proteins: review of their functions in plant environmental

807　　stress tolerance. *Plants (Basel, Switzerland)* **12**, 2045.

808　Cao J, Packer JS, Ramani V, Cusanovich DA, Huynh C, Daza R, Qiu X, Lee C, Furlan SN, Steemers FJ, et al. 2017.

809　　Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661–667.

810   Chen ZH, Chen G, Dai F, Wang Y, Hills A, Ruan YL, Zhang G, Franks PJ, Nevo E, Blatt MR. 2017. Molecular evolution
811   of grass stomata. *Trends In Plant Science* **22**, 124–139.

812   Cominelli E, Galbiati M, Vavasseur A, Conti L, Sala T, Vuylsteke M, Leonhardt N, Dellaporta SL, Tonelli C. 2005. A
813   guard-cell-specific MYB transcription factor regulates stomatal movements and plant drought tolerance.
814   *Current Biology* **15**, 1196–1200.

815   Covington MF, Harmer SL. 2007. The circadian clock regulates auxin signaling and responses in *Arabidopsis*. *PLoS*
816   *biology* **5**, e222.

817   Deng YN, Kashtoh H, Wang Q, Zhen GX, Li QY, Tang LH, Gao HL, Zhang CR, Qin L, Su M, et al. 2021. Structure and
818   activity of SLAC1 channels for stomatal signaling in leaves. *PNAS* **118**, e2015151118.

819   Denyer T, Ma X, Klesen S, Scacchi E, Nieselt K, Timmermans MCP. 2019. Spatiotemporal developmental trajectories
820   in the *Arabidopsis* root revealed using high-throughput single-cell rna sequencing. *Developmental Cell* **48**, 840–
821   852.

822   Denyer T, Timmermans MCP. 2022. Crafting a blueprint for single-cell RNA sequencing. *Trends In Plant Science* **27**,
823   92–103.

824   Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast
825   universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.

826   Dodd AN, Salathia N, Hall A, Kévei E, Tóth R, Nagy F, Hibberd JM, Millar AJ, Webb AA. 2005. Plant circadian clocks
827   increase photosynthesis, growth, survival, and competitive advantage. *Science* **309**, 630-633.

828   Dong J, Chen C, Chen Z. 2003. Expression profiles of the *Arabidopsis* WRKY gene superfamily during plant defense
829   response. *Plant Molecular Biology* **51**, 21–37.

830   Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome*
831   *Biology* **20**, 238.

832   Endo M. 2016. Tissue-specific circadian clocks in plants. *Current Opinion in Plant Biology* **29**, 44–49.

833   Ernst E, Abramson B, Acosta K, Hoang PTN, Mateo-Elizalde C, Schubert V, Pasaribu B, Hartwick N, Colt K, Aylward A,

834   et al. 2023. The genomes and epigenomes of aquatic plants (*Lemnaceae*) promote triploid hybridization and

835   clonal reproduction. *bioRxiv* doi: 10.1101/2023.08.02.551673

836   Escamez S, Stael S, Vainonen JP, Willems P, Jin H, Kimura S, Van Breusegem F, Gevaert K, Wrzaczek M, Tuominen H.

837   2019. Extracellular peptide Kratos restricts cell death during vascular development and stress in *Arabidopsis*.

838   *Journal Of Experimental Botany* **70**, 2199–2210.

839   Ferrari C, Proost S, Janowski M, Becker J, Nikoloski Z, Bhattacharya D, Price D, Tohge T, Bar-Even A, Fernie A, et al.

840   2019. Kingdom-wide comparison reveals the evolution of diurnal gene expression in *Archaeplastida*. *Nature*

841   *Communications* **10**, 737.

842   Filichkin SA, Breton G, Priest HD, Dharmawardhana P, Jaiswal P, Fox SE, Michael TP, Chory J, Kay SA, Mockler TC.

843   2011. Global profiling of rice and poplar transcriptomes highlights key conserved circadian-controlled pathways

844   and cis-regulatory modules. *PloS One* **6**, e16907.

845   Fukao T, Xu K, Ronald PC, Bailey-Serres J. 2006. A variable cluster of ethylene response factor-like genes regulates

846   metabolic and developmental acclimation responses to submergence in rice. *The Plant Cell* **18**, 2021–2034.

847   Galvez-Valdivieso G, Fryer MJ, Lawson T, Slattery K, Truman W, Smirnoff N, Asami T, Davies WJ, Jones AM, Baker

848   NR, Mullineaux PM. 2009. The high light response in *Arabidopsis* involves ABA signaling between vascular and

849   bundle sheath cells. *The Plant Cell* **21**, 2143–2162.

850   Glenz R, Kaiping A, Göpfert D, Weber H, Lambour B, Sylvester M, Fröschel C, Mueller MJ, Osman M, Waller F. 2022.

851   The major plant sphingolipid long chain base phytosphingosine inhibits growth of bacterial and fungal plant

852   pathogens. *Scientific Reports* **12**, 1081.

853   Greenham K, Guadagno CR, Gehan MA, Mockler TC, Weinig C, Ewers BE, McClung CR. 2017. Temporal network

854   analysis identifies early physiological and transcriptomic indicators of mild drought in *Brassica rapa*. *Elife* **6**,

855   e29655

856   Harkess A, McLoughlin, Bilkey N, Elliot K, Emenecker R, Mattoon E, Miller K, Czymmek RD, Meyers BC, Michael TP.

857   2021. Improved Spirodela polyrhiza genome and proteomic analyses reveal a conserved chromosomal

858   structure with high abundance of chloroplastic proteins favoring energy production. *Journal of Experimental*

859   *Biology* **72**, 2491-2500.

860   Hunt L, Gray JE. 2009. The signaling peptide EPF2 controls asymmetric cell divisions during stomatal development.

861   *Current Biology* **19**, 864–869.

862   Jean-Baptiste K, McFaline-Figueroa JL, Alexandre CM, Dorrity MW, Saunders L, Bubb KL, Trapnell C, Fields S,

863   Queitsch C, Cuperus JT. 2019. Dynamics of gene expression in single root cells of *Arabidopsis thaliana*. *The*

864   *Plant Cell* **31**, 993–1011.

865   Kim JY, Symeonidi E, Pang TY, Denyer T, Weidauer D, Bezrutczyk M, Miras M, Zöllner N, Hartwig T, Wudick MM, et

866   al. 2021. Distinct identities of leaf phloem cells revealed by single cell transcriptomics. *The Plant Cell* **33**, 511–

867   530.

868   Klopfenstein DV, Zhang L, Pedersen BS, Ramírez F, Warwick Vesztrocy A, Naldi A, Mungall CJ, Yunes JM, Botvinnik

869   O, Weigel M, et al. 2018. GOATOOLS: A Python library for Gene Ontology analyses. *Scientific Reports* **8**, 10872.

870   Knauer S, Javelle M, Li L, Li X, Ma X, Wimalanathan K, Kumari S, Johnston R, Leiboff S, Meeley R, et al. 2019. A high-

871   resolution gene expression atlas links dedicated meristem genes to key architectural traits. *Genome Research*

872   **29**, 1962–1973.

873   Kovinich N, Kayanja G, Chanoca A, Otegui MS, Grotewold E. 2015. Abiotic stresses induce different localizations of

874   anthocyanins in *Arabidopsis*. *Plant Signaling & Behavior* **10**, e1027850.

875   Lam E, Michael TP. 2022. *Wolffia*, a minimalist plant and synthetic biology chassis. *Trends In Plant Science* **27**, 430–

876   439.

877   Laosuntisuk K, Elorriaga E, Doherty CJ. 2023. The game of timing: Circadian rhythms intersect with changing

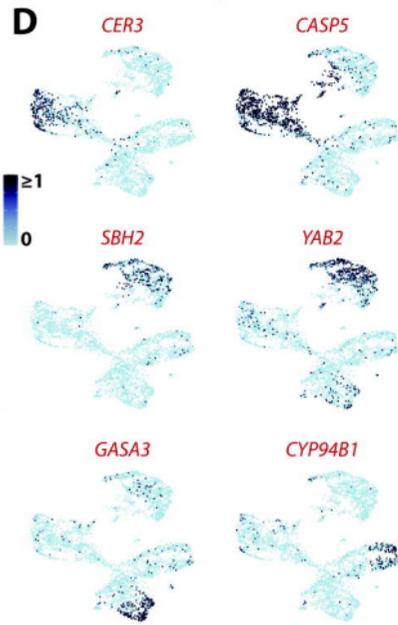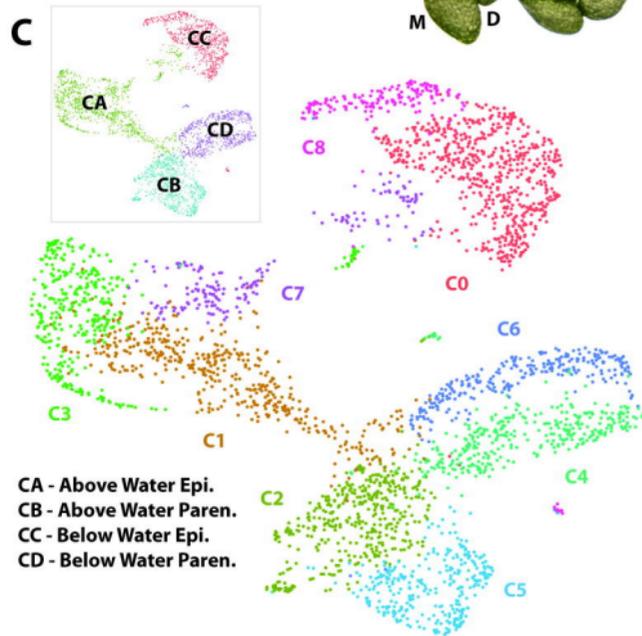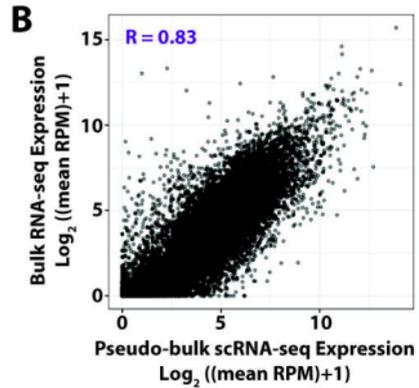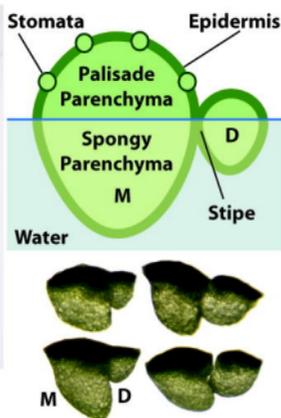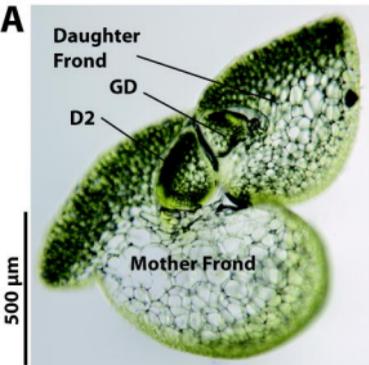878   environments. *Annual review of plant biology* **74**, 511–538.

879   Li F, Yang JJ., Sun ZY, Wang L, Qi LY, Liu YQ, Zhang HM, Dang LF, Wang SJ, Luo CX, et al. 2023. Plant-on-chip: Core

880   morphogenesis processes in the tiny plant *Wolffia australiana*. *PNAS nexus* **2**, pgad141.

881   Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general-purpose program for assigning sequence reads

882   to genomic features. *Bioinformatics* **30**, 923–930.

883   Liu T, Chen T, Kan J, Yao Y, Guo D, Yang Y, Ling X, Wang J, Zhang B. 2022. The GhMYB36 transcription factor confers

884   resistance to biotic and abiotic stress by enhancing PR1 gene expression in plants. *Plant Biotechnology Journal*

885   **20**, 722–735.

886   Lopez-Anido CB, Vatén A, Smoot NK, Sharma N, Guo V, Gong Y, Anleu Gil MX, Weimer AK, Bergmann DC. 2021.

887   Single-cell resolution of lineage trajectories in the *Arabidopsis* stomatal lineage and developing leaf.

888   *Developmental Cell* **56**, 1043–1055.

889   Lou P, Wu J, Cheng F, Cressman LG, Wang X, McClung CR. 2012. Preferential retention of circadian clock genes

890   during diploidization following whole genome triplication in *Brassica rapa*. *Plant Cell* **24**:2415-26.

891   Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with

892   DESeq2. *Genome Biology* **15**, 550.

893   Lü S, Song T, Kosma DK, Parsons EP, Rowland O, Jenks MA. 2009. *Arabidopsis* CER8 encodes LONG-CHAIN ACYL-COA

894   SYNTHETASE 1 (LACS1) that has overlapping functions with LACS2 in plant wax and cutin synthesis. *The Plant*

895   *Journal* **59**, 553–564.

896   Ma X, Denyer T, Timmermans MCP. 2020. PscB: A browser to explore plant single cell rna-sequencing data sets.

897   *Plant Physiology* **183**, 464–467.

898   Marand AP, Chen Z, Gallavotti A, Schmitz RJ. 2021. A cis-regulatory atlas in maize at single-cell resolution. *Cell* **184**,

899   3041–3055.

900   Meng X, Chen X, Mang H, Liu C, Yu X, Gao X, Torii KU, He P, Shan L. 2015. Differential function of *Arabidopsis* SERK

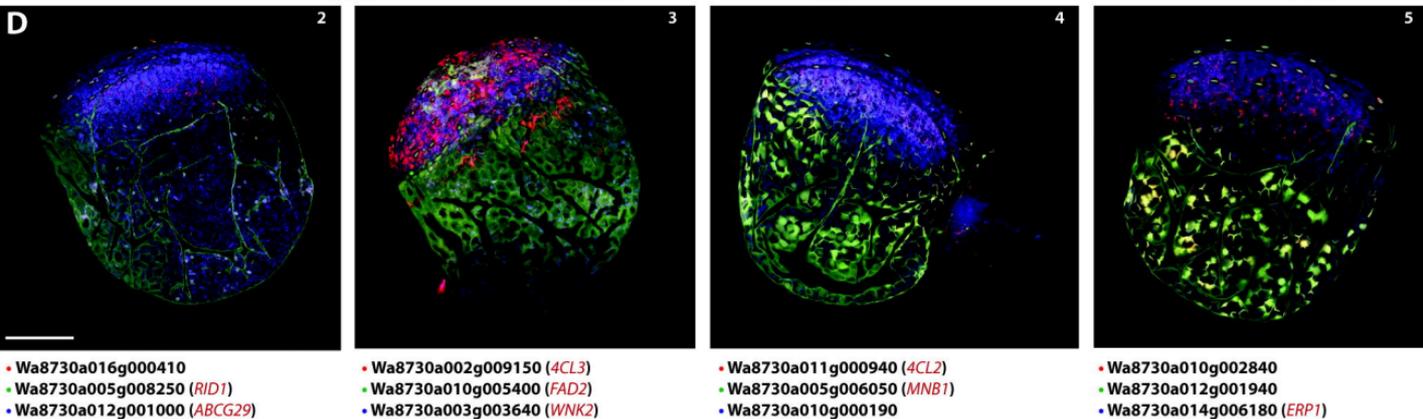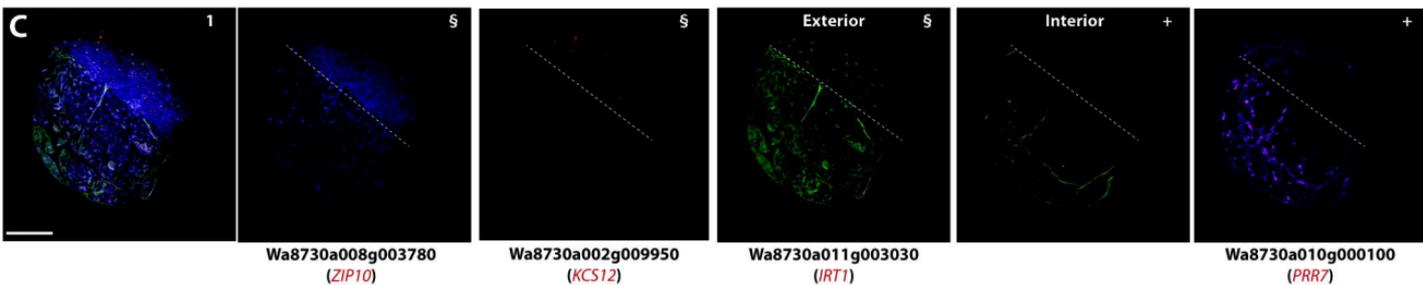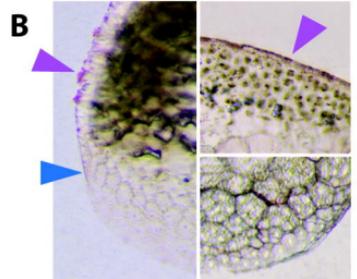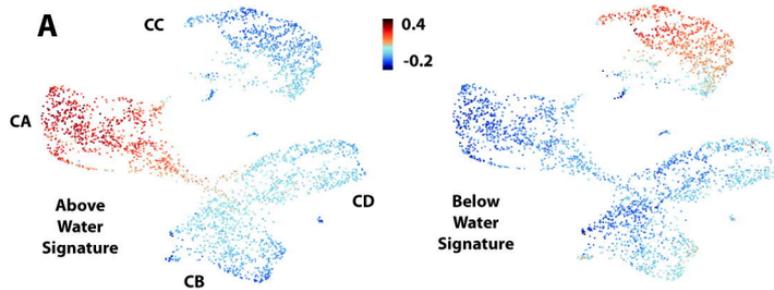901   family receptor-like kinases in stomatal patterning. *Current Biology* **25**, 2361–2372.

902 Michael TP, Salomé PA, Yu HJ, Spencer TR, Sharp EL, McPeek MA, Alonso JM, Ecker JR, McClung CR. 2003. Enhanced

903   fitness conferred by naturally occurring variation in the circadian clock. *Science* **302**, 1049–1053.

904 Michael TP, Ernst E, Hartwick N, Chu P, Bryant D, Gilbert S, Ortleb S, Baggs EL, Sree KS, Appenroth KJ, et al. 2020.

905   Genome and time-of-day transcriptome of *Wolffia australiana* link morphological minimization with gene loss

906   and less growth control. *Genome Research* **31**, 225–238.

907 Michael TP. 2022a. Core circadian clock and light signaling genes brought into genetic linkage across the green

908   lineage. *Plant Physiology* **190**, 1037–1056.

909 Michael TP. 2022b. Time of Day analysis over a field grown developmental time course in rice. *Plants* **12**, 166.

910 Nadiminti PP, Rookes JE, Boyd BJ, Cahill DM. 2015. Confocal laser scanning microscopy elucidation of the

911   micromorphology of the leaf cuticle and analysis of its chemical composition. *Protoplasma* **252**, 1475–1486.

912 Nakamichi N, Kita M, Ito S, Sato E, Yamashino T, Mizuno T. 2005. The *Arabidopsis* pseudo-response regulators,

913   PRR5 and PRR7, coordinately play essential roles for circadian clock function. *Plant & Cell Physiology* **46**, 609–

914   619.

915 Negi J, Moriwaki K, Konishi M, Yokoyama R, Nakano T, Kusumi K, Hashimoto-Sugimoto M, Schroeder JI, Nishitani K,

916   Yanagisawa S, Iba K. 2013. A Dof transcription factor, SCAP1, is essential for the development of functional

917   stomata in *Arabidopsis*. *Current Biology* **23**, 479–484.

918 Nobori N, Oliva M, Lister R, Ecker JR. 2023. Multiplexed single-cell 3D spatial gene expression analysis in plant tissue

919   using PHYTOMap. *Nature Plants* **9**, 1026-1033.

920 Oravec MW, Greenham K. 2022. The adaptive nature of the plant circadian clock in natural environments. *Plant*

921   *Physiology* **190**, 968-980

922 Para A, Farré EM, Imaizumi T, Pruneda-Paz JL, Harmon FG, Kay SA. 2007. PRR3 Is a vascular regulator of TOC1

923   stability in the *Arabidopsis* circadian clock. *Plant Cell* **19**, 3462-73.

924 Park CS, Go YS, Suh MC. 2016. Cuticular wax biosynthesis is positively regulated by WRINKLED4, an AP2/ERF-type

925   transcription factor, in *Arabidopsis* stems. *The Plant Journal* **88**, 257–270.
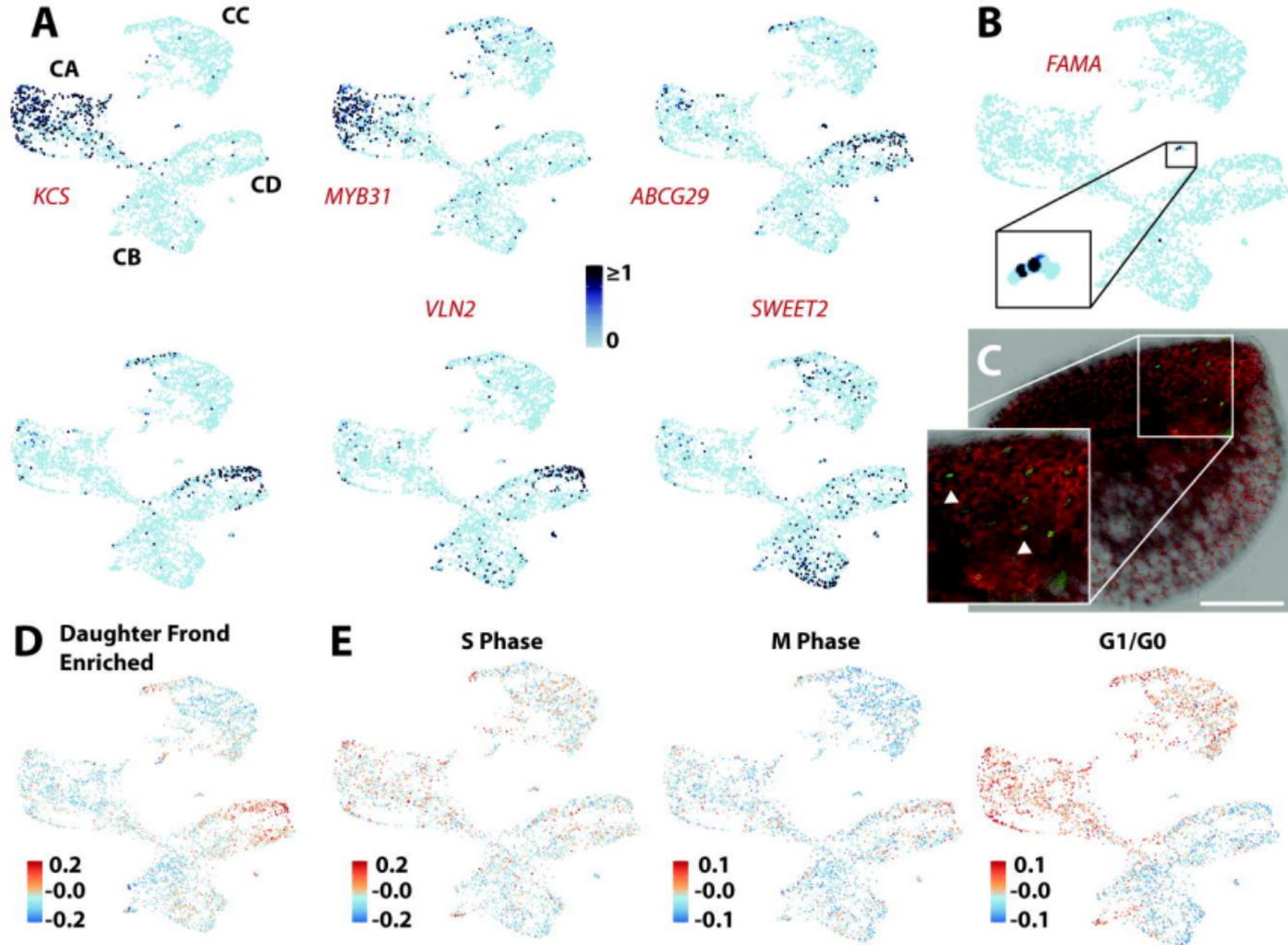
926   Pasaribu B, Acosta K, Aylward A, Liang Y, Abramson BW, Colt K, Hartwick NT, Shanklin J, Michael TP, Lam E. 2023.

927   Genomics of turions from the greater duckweed reveal its pathways for dormancy and re-emergence strategy.

928   *The New phytologist* **239**, 116–131.

929   R Core Team. 2021. A language and environment for statistical computing. *R Foundation for Statistical Computing*,

930   Vienna, Austria.

931   Raskin I, Kende H. 1984. Regulation of growth in stem sections of deep-water rice. *Planta* **160**, 66–72.

932   Roppolo D, De Rybel B, Dénervaud Tendon V, Pfister A, Alassimone J, Vermeer JE, Yamazaki M, Stierhof YD,

933   Beeckman T, Geldner N. 2011. A novel protein family mediates Casparian strip formation in the endodermis.

934   *Nature* **473**, 380–383.

935   Sanchez SE, Kay SA. 2016. The plant circadian clock: From a simple timekeeper to a complex developmental

936   manager. *Cold Spring Harbor Perspectives in Biology* **8**, a027748.

937   Satija R, Farrell JA, Gennert D, Schier AF, Regev A. 2015. Spatial reconstruction of single-cell gene expression data.

938   *Nature Biotechnology* **33**, 495–502.

939   Seyfferth C, Renema J, Wendrich JR, Eekhout T, Seurinck R, Vandamme N, Blob B, Saeys Y, Helariutta Y, Birnbaum

940   KD, De Rybel B. 2021. Advances and opportunities in single-cell transcriptomics for plant research. *Annual*

941   *Review of Plant Biology* **72**, 847–866.

942   Shahan R, Hsu CW, Nolan TM, Cole BJ, Taylor IW, Greenstreet L, Zhang S, Afanassiev A, Vlot AHC, Schiebinger G, et

943   al. 2022. A single-cell *Arabidopsis* root atlas reveals developmental trajectories in wild-type and cell identity

944   mutants. *Developmental Cell* **57**, 543–560.

945   Shimada Y, Fujioka S, Miyauchi N, Kushiro M, Takatsuto S, Nomura T, Yokota T, Kamiya Y, Bishop GJ, Yoshida S.

946   2001. Brassinosteroid-6-oxidases from *Arabidopsis* and tomato catalyze multiple C-6 oxidations in

947   brassinosteroid biosynthesis. *Plant Physiology* **126**, 770–779.

948   Shulse CN, Cole BJ, Ciobanu D, Lin J, Yoshinaga Y, Gouran M, Turco GM, Zhu Y, O'Malley RC, Brady SM, & Dickel DE.

949   (2019). High-throughput single-cell transcriptome profiling of plant cell types. *Cell Reports* **27**, 2241–2247.
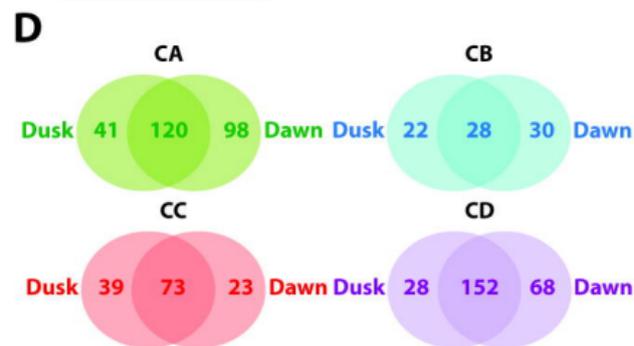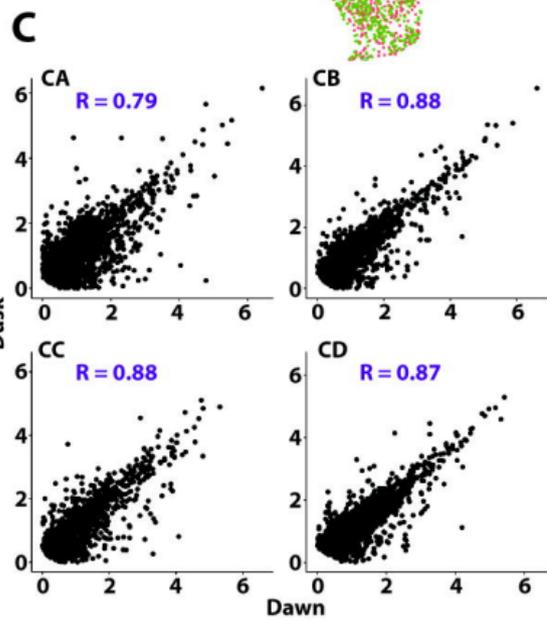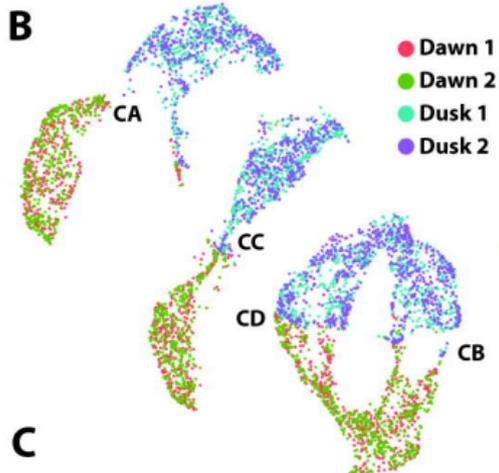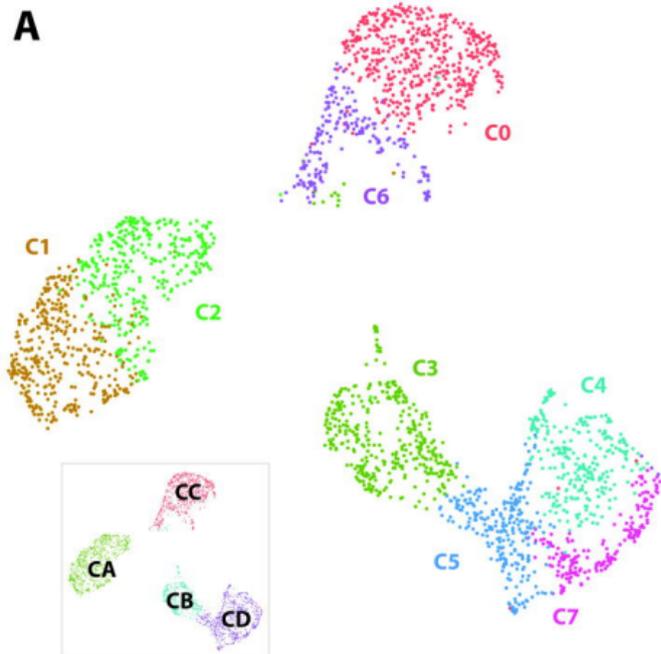
950   Siegfried KR, Eshed Y, Baum SF, Otsuga D, Drews GN, Bowman JL. 1999. Members of the YABBY gene family specify
951   abaxial cell fate in *Arabidopsis*. *Development* **126**, 4117–4128.

952   Smit ME, Bergmann DC. 2023. The stomatal fates: Understanding initiation and enforcement of stomatal cell fate
953   transitions. *Current Opinion in Plant Biology* **79**, 102449

954   Somers DE, Schultz TF, Milnamow M, Kay SA. 2000. ZEITLUPE encodes a novel clock-associated PAS protein from
955   *Arabidopsis*. *Cell* **101**, 319–329.

956   Steed G, Ramirez DC, Hannah MA, Webb AAR. 2021. Chronoculture, harnessing the circadian clock to improve crop
957   yield and sustainability. *Science* **372**, eabc9141.

958   Swift J, Greenham K, Ecker JR, Coruzzi GM, Robertson McClung C. 2022. The biology of time: dynamic responses of
959   cell types to developmental, circadian and environmental cues. *The Plant Journal* **109**, 764–778.

960   van der Honing HS, Kieft H, Emons AM, Ketelaar T. 2012. *Arabidopsis* VILLIN2 and VILLIN3 are required for the
961   generation of thick actin filament bundles and for directional organ growth. *Plant Physiology* **158**, 1426–1438.

962   Wang X, Niu Y, Zheng Y. 2021. Multiple functions of MYB transcription factors in abiotic stress responses.
963   *International Journal of Molecular Sciences* **22**, 6125.

964   Ware A, Jones DH, Flis P, Chrysanthou E, Smith KE, Kümpers BMC, Yant L, Atkinson JA, Wells DM, Bhosale R,
965   Bishopp A. 2023. Loss of ancestral function in duckweed roots is accompanied by progressive anatomical
966   reduction and a re-distribution of nutrient transporters. *Current Biology* **33**, 1795-1802.

967   Watanabe E, Isoda M, Muranaka T, Ito S, Oyama T. 2021. Detection of uncoupled circadian rhythms in individual
968   cells of *Lemna minor* using a dual-color bioluminescence monitoring system. *Plant & Cell Physiology* **62**, 815–
969   826.

970   Wolf S. 2022. Cell wall signaling in plant development and defense. *Annual Review of Plant Biology* **73**, 323–353.

971   Xiong C, Xie Q, Yang Q, Sun P, Gao S, Li H, Zhang J, Wang T, Ye Z, Yang C. 2020. WOOLLY, interacting with MYB
972   transcription factor MYB31, regulates cuticular wax biosynthesis by modulating CER6 expression in tomato. *The*
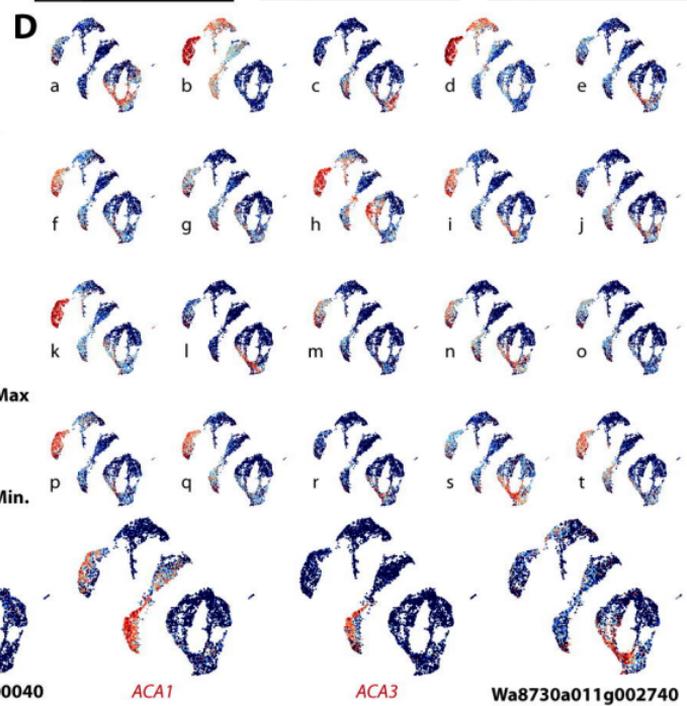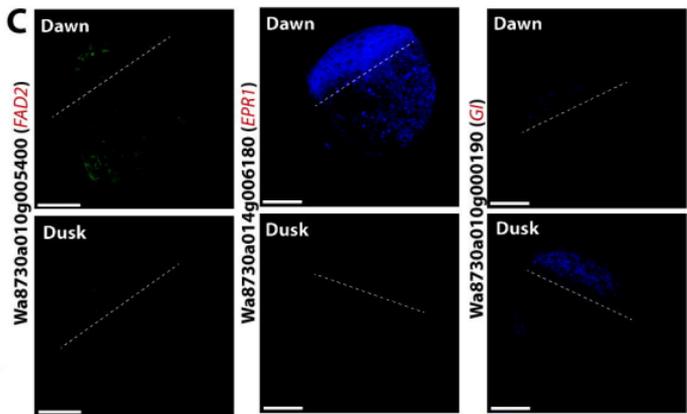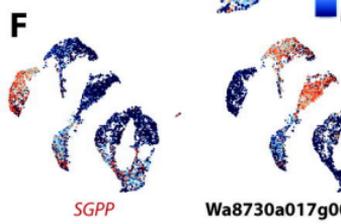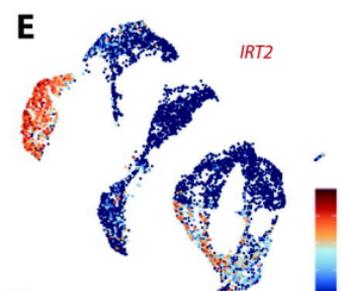973   *Plant Journal* **103**, 323–337.

974   Yoshida T, Nishimura N, Kitahata N, Kuromori T, Ito T, Asami T, Shinozaki K, Hirayama T. 2006. ABA-hypersensitive

975   germination3 encodes a protein phosphatase 2C (AtPP2CA) that strongly regulates abscisic acid signaling

976   during germination among *Arabidopsis* protein phosphatase 2Cs. *Plant Physiology* **140**, 115–126.

977   Zhang TQ, Chen Y, Wang JW. 2021. A single-cell analysis of the *Arabidopsis* vegetative shoot apex. *Developmental*

978   *Cell* **56**, 1056–1074.

979   Zhao K, Rhee SY. 2022. Omics-guided metabolic pathway discovery in plants: Resources, approaches, and

980   opportunities. *Current Opinion in Plant Biology* **67**, 102222.

981   Zhu J, Lolle S, Tang A, Guel B, Kvitko B, Cole B, Coaker G. 2023. Single-cell profiling of *Arabidopsis* leaves to

982   *Pseudomonas syringae* infection. *Cell Reports* **42**, 112676.

983

A Labels: Daughter Frond, GD, D2, Mother Frond, 500 μm; Stomata, Epidermis, Palisade Parenchyma, Spongy Parenchyma, M, D, Stipe, Water; M, D

B R = 0.83

Bulk RNA-seq Expression Log$_2$ ((mean RPM)+1)

Pseudo-bulk scRNA-seq Expression Log$_2$ ((mean RPM)+1)

C CC, CA, CD, CB; C8, C0, C7, C6, C3, C1, C2, C4, C5

CA - Above Water Epi.
CB - Above Water Paren.
CC - Below Water Epi.
CD - Below Water Paren.

D *CER3*, *CASP5*, *SBH2*, *YAB2*, *GASA3*, *CYP94B1*

≥1, 0

A CC

CA

Above
Water
Signature

CD

CB

Below
Water
Signature

B

C

1 § §

Exterior § Interior +

+

Wa8730a008g003780
(*ZIP10*)

Wa8730a002g009950
(*KCS12*)

Wa8730a011g003030
(*IRT1*)

Wa8730a010g000100
(*PRR7*)

D

2

3

4

5

• Wa8730a016g000410
• Wa8730a005g008250 (*RID1*)
• Wa8730a012g001000 (*ABCG29*)

• Wa8730a002g009150 (*4CL3*)
• Wa8730a010g005400 (*FAD2*)
• Wa8730a003g003640 (*WNK2*)

• Wa8730a011g000940 (*4CL2*)
• Wa8730a005g006050 (*MNB1*)
• Wa8730a010g000190

• Wa8730a010g002840
• Wa8730a012g001940
• Wa8730a014g006180 (*ERP1*)

A

CC
CA
KCS
CD
CB

MYB31

ABCG29

≥1
0

VLN2

SWEET2

B *FAMA*

C

D Daughter Frond Enriched

0.2
-0.0
-0.2

E S Phase

0.2
-0.0
-0.2

M Phase

0.1
-0.0
-0.1

G1/G0

0.1
-0.0
-0.1

A

C0
C6
C1
C2
C3
C4
C5
C7
CC
CA
CB
CD

B

- Dawn 1
- Dawn 2
- Dusk 1
- Dusk 2

CA
CC
CD
CB

C

CA   R = 0.79

CB   R = 0.88

CC   R = 0.88

CD   R = 0.87

Dusk

Dawn

D

CA
Dusk 41 120 98 Dawn

CB
Dusk 22 28 30 Dawn

CC
Dusk 39 73 23 Dawn

CD
Dusk 28 152 68 Dawn

Streamlined spatial and environmental expression signatures characterize the minimalist duckweed *Wolffia australiana*

Tom Denyer, Pin-Jou Wu, Kelly Colt, et al.

Supplemental Material	http://genome.cshlp.org/content/suppl/2024/08/01/gr.279091.124.DC1
P<P	Published online July 1, 2024 in advance of the print journal.
Accepted Manuscript	Peer-reviewed and accepted for publication but not copyedited or typeset; accepted manuscript is likely to differ from the final, published version.
Open Access	Freely available online through the *Genome Research* Open Access option.
Creative Commons License	This manuscript is Open Access.This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International license), as described at http://creativecommons.org/licenses/by-nc/4.0/.
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

To subscribe to *Genome Research* go to:
https://genome.cshlp.org/subscriptions