

Gene networks are conserved across reproductive development between the fern *Ceratopteris richardii* and the flowering plant *Arabidopsis thaliana*

Andrew R.G. Plackett^{1*}, Marco Catoni¹

¹University of Birmingham, School of Biosciences, Edgbaston Campus, Birmingham, B15 2TT, UK.

*Corresponding author. A.R.G.Plackett@bham.ac.uk

ABSTRACT

How plants first invented seeds is a longstanding unresolved evolutionary question. Seed-bearing plants arose from within seedless vascular plants, with ferns as their closest relatives, but how the seed developmental programme first originated has remained intractable through comparative morphology or the fossil record. To investigate this question at the level of gene network evolution, we established a transcriptional expression atlas across sporophyte and gametophyte reproductive development of the fern *Ceratopteris richardii* and compared reproductive-associated genes to those of *Arabidopsis thaliana*. Conservation was detected between distinct organs. As in flowering, the hormone gibberellin promoted sporing of the fern sporophyte shoot. Sporophyll genes were conserved with floral development. Gametophyte post-fertilization egg chamber (archegonium) genes were conserved with both pre- and post-fertilization seed development, as were two post-fertilization signals, auxin and sugar. We conclude that the seed may first have arisen through the mis-expression of archegonium developmental programmes within the developing sporophyte sporangium.

INTRODUCTION

The evolution of the seed entailed dramatic changes to plant reproduction, arising from a single origin within the seedless vascular plants^{1,2} and with monilophytes (horsetails and ferns) the closest surviving outgroup^{2,3,4}. Extant and extinct seeds ('ovules' before fertilization) share a common body-plan: a diploid nucellus apically specialised for pollen reception, containing a multicellular haploid egg-producing structure, with a subtending chalaza from which one or more integuments arise⁵. Ovule tissues undergo further development after fertilization to form the seed coat and (in angiosperms) specialised storage tissues (the endosperm)^{6,7,8}. But how this developmentally-complex organ first originated is a longstanding unresolved question.

Seed evolution has been deduced by comparing to ancestral seedless reproduction still found in extant ferns, in which separate free-living diploid and haploid 'generations' both form reproductive organs^{9,10}. The diploid sporophyte forms reproductively-modified fronds (sporophylls) within which meiosis generates haploid spores. The thalloid gametophytes that germinate from these produce gametangia through mitosis: antheridia (containing free-swimming sperm) and archegonia (egg chambers) which when fertilized undergo post-fertilization modifications to support embryo development¹¹. Early ovule evolution is proposed to have involved sporophyll modifications to encapsulate a reduced gametophyte (endospory)^{12,13}, with the archegonium still present in gymnosperms but lost in angiosperms⁷. A further crucial innovation was heterospory, genetically-predetermined male and female sporangia (pollen sac and ovule nucellus) arising from ancestral homosporous sporangia^{14,15,16,17}. Many ferns remain homosporous, including the genetic model *Ceratopteris richardii*¹⁸. 'Male' (micro)sporogenesis resembles ancestral development whereas 'female' (mega)sporogenesis has apparently undergone significant modification¹⁹, now typically resulting in a single viable megaspore (monospory)^{20,21}. The genetic changes underlying these evolutionary transitions remain unknown.

A complete molecular hypothesis for ovule evolution has not been proposed, but a role for the MADS-box gene family, which specifies angiosperm ovule identity^{22,23} has long been suggested^{24,25}. Recent analysis of individual fern genes supports the partial conservation of reproductive gene networks^{24,26,27,28}, but to what extent is unclear. Although fern genomic information is increasingly available^{29,30,31,32,33}, as are transcriptomes at either whole-lifecycle resolution^{34,35,36} or focussed on single reproductive organs^{30,37}, a detailed characterisation of gene networks across fern reproductive development to compare to seed plant reproduction is lacking. We thus undertook mRNA-seq analysis to identify genes associated with reproductive processes in both the *Ceratopteris* sporophyte and gametophyte. We identified and experimentally validated a conserved role for the phytohormone gibberellin (GA) in the

sporophyte shoot reproductive transition and two post-fertilization regulatory mechanisms between the seed and the fern archegonium. GO term enrichment and whole-network comparison suggest that ovule/seed gene networks are conserved specifically with the post-fertilization fern archegonium. This raises a new hypothesis that the ovule arose by mis-expressing post-fertilization archegonium gene networks within the sporophyte.

RESULTS

*Identifying reproductive gene networks in the *Ceratopteris* sporophyte*

To establish a tractable experimental system, the *Ceratopteris* sporophyte reproductive transition was first characterised. The shoot apex switched to reproductive development ('sporing') consistently between replicates, chronologically (100% by 18 weeks) and developmentally (20.6 ± 0.07 fronds) (**Extended Data Fig. 1; Supplementary Data 1**). Transcriptional changes associated with reproductive development were identified by comparing gene expression across frond development between the sporing phase (130 days old) (**Fig. 1a-f**) and the vegetative phase immediately prior to sporing (60 days old) (**Fig. 1g-l**). Further 25-day-old apex and primordia samples were included as a control for the sporing transition (**Supplementary Fig. 1**). Frond development was divided into sequential stages common to vegetative fronds and sporophylls: the shoot apical region ('Apx') (**Fig. 1b,h**); frond primordium ('Prm') (**Fig. 1c,i; Supplementary Fig. 1**); fiddlehead ('Fid') (**Fig. 1d,j**); expanding frond ('Exp') (**Fig. 1e,k**) and mature frond ('Mat') (**Fig. 1f,l**). Sporangia were visible by the start of sporophyll expansion (**Extended Data Fig. 2a-l**), with sporangium sizes suggesting overlapping development throughout (**Extended Data Fig. 2m,n**). Post-meiotic sporangia were only observed in mature sporophyll samples (**Extended Data Fig. 2e-h**).

Principal Component Analysis (PCA) and mRNA-seq expression correlation confirmed that biological replicates clustered by sample (**Extended Data Fig. 3a, Supplementary Fig. 2**). Frond developmental stages separated progressively along PC1 (excluding Apx and Prm, which overlapped except for 25d Prm) while vegetative and reproductive phases separated across PC2. Differential expression (DE) analysis identified 2098 transcripts with increased abundance ($P_{\text{adj}} < 0.05$, $\log_2\text{foldchange} > 1$) in sporophylls vs vegetative controls in at least one frond stage (**Fig. 1m; Supplementary Data 2**). Among these we categorised 295 as present (FPKM > 1) only in the sporophyll ('sporophyll-specific') and 1437 present in the sporophyll and at least one vegetative stage ('sporophyll-enriched') (**Fig. 1n**), representing potential direct regulators of reproduction and genes either with dual vegetative and reproductive function or frond developmental genes whose expression is modified in sporophylls. 366 remaining candidates were expressed at FPKM < 1 in all sporophyll samples.

Gibberellin promotes the Ceratopteris sporing transition.

Because Apx and Prm samples overlapped under PCA, Apx reproductive candidates were analysed separately (**Fig. 2a**). Fifteen significantly-enriched reproductive GO terms ($p < 0.05$) were identified and all associated with one MADS-box gene, *CMADS1*²⁴ (**Supplementary Data 3**). Genes promoting a shoot reproductive transition were predicted to have increased expression immediately prior to it, and to be preferentially expressed in the shoot apex vs primordia. Expression data was used to rank thirteen Apx candidates with progressive significant increases in abundance ($P_{adj} < 0.05$) between 25d, 60d and 130d (**Fig. 2b**). These included *CMADS1*, but this was present in all ages and was not preferential to Apx samples ($P_{adj} > 0.05$). Apx candidates contained two further MADS-box homologs which showed increased abundance at 130d only (**Supplementary Data 2**).

Candidates whose expression profile matched that predicted for sporing regulators included a GA 20-oxidase GA biosynthetic enzyme (**Fig. 2b**). Two further GA-related orthologs were identified: *RGA TARGET 1 (RGAT1)*, a direct target of GA signalling in *Arabidopsis* pollen development³⁸, and a *GA 3-OXIDASE* biosynthesis gene significantly decreased in abundance ($P_{adj} < 0.05$) at 130d (**Supplementary Data 2**), resembling a known homeostatic response to GA signalling³⁹. These results suggested a role for GA in promoting *Ceratopteris* sporing. We verified this through exogenous application of GA₃, which significantly accelerated the sporing transition ($p < 0.05$) (**Fig. 2c-h; Supplementary Data 1**). Our results thus support a conserved GA function between ferns and angiosperms promoting the sporophyte shoot reproductive transition.

Sporophyll reproductive candidates include meiotic and floral development gene orthologs

Hierarchical clustering by normalised expression successfully identified differential expression of candidate genes across sporophyll development, with seven clusters exhibiting maximum abundance in different stages (**Fig. 3a**). 54.7% of candidates possessed GO term annotations, with reproduction (11.5%) and gene regulatory GO terms (6.9%) both significantly enriched ($p < 0.05$) (**Supplementary Fig. 3**). Reproductive GO terms related to meiosis or to reproductive shoot/organ development. Most meiosis-associated GO terms were significantly enriched in mature sporophylls (cluster 3) (**Fig. 3b**) consistent with our developmental characterisation (**Extended Data Fig. 2**), and in the sporophyll-specific repression category (**Supplementary Fig. 4**). Through these we identified orthologs of *ZIP4*, *DISRUPTED MEIOTIC CDNA1 (DMC1)*, *MSH4*, *MSH5* and a *SHUGOSHIN (SGO)*-like homolog (**Supplementary Data 3**), which are all conserved regulators of meiosis^{40,41,42,43,44}.

Reproductive shoot GO terms were significantly enriched in sporophyll primordia (clusters 1 and 7) and expanding sporophylls (cluster 4) (**Fig. 3b**) and sporophyll-enriched candidates (**Supplementary Fig. 4**). All related to three MADS-box orthologs including *CMADS1* (**Supplementary Data 3**). Post-translational regulation GO terms were significantly enriched in cluster 1 (**Supplementary Fig. 5**) including an ortholog of *MERISTEMATIC RECEPTOR-LIKE KINASE (MRLK)*, which targets floral MADS-box protein *AGL24*^{45,46,47}, and *CrEMS1-2*²⁸, an uncharacterised ortholog of *EXCESS MICROSPOROCTES1 (EXS/EMS1)*^{48,49} (**Supplementary Data 3**). These findings suggest conservation between the *Ceratopteris* sporophyll and *Arabidopsis* floral gene networks.

Identifying reproductive gene networks in the *Ceratopteris* gametophyte

The *Ceratopteris* hermaphrodite gametophyte (from hereon 'gametophyte') develops archegonia from a 'notch' meristem^{50,51}. To capture gene expression associated with female reproductive development, we confirmed that fertilization and embryo development could be synchronised experimentally by restricting water availability during development and then flooding once, with the majority of gametophytes bearing visible embryos 4 days after flooding (DAF) (**Extended Data Fig. 4a-i; Supplementary Fig. 6, Supplementary Data 4**). The notch meristem also exhibited fertilization-dependent shutdown by 2 DAF (**Extended Data Fig. 4j**). Genes relating to archegonium formation, fertilization and post-fertilization development were thus identified by comparing between unflooded gametophytes and 1, 2 and 4 DAF, dissecting out the notch region ('Ntc') (**Fig. 4a-d**) and using the surrounding thallus ('Thl') as a vegetative control (**Fig. 4e-h**), unavoidably including antheridia at the thallus margin (**Extended Data Fig. 4a,c,e**). The 4 DAF notch was subdivided into embryo-bearing archegonia (Aeb) and embryo-less archegonia (Arc) (**Fig. 4d**). To discriminate embryo from archegonium gene networks, we isolated 4 DAF embryos as an additional control ('Dem') (**Supplementary Fig. 1**).

PCA and correlation analysis found that biological replicates clustered more closely within each sample than between (**Extended Data Fig. 3b, Supplementary Fig. 7**) except for 1 DAF and 2 DAF (which overlapped within Ntc and Thl). PC1 separated Ntc from Thl and DAF separated progressively along PC2 (except for Dem). 7817 reproductive candidate genes were identified with increased abundance in Ntc compared to Thl ($P_{adj} < 0.05$, $\log_2\text{foldchange} > 1$) within at least one developmental stage (**Fig. 4i; Supplementary Data 5**). Similar to the sporophyll, candidate genes were categorised as 'notch-specific' (1435) and 'notch-enriched' (5057) (**Fig. 4j**), representing potential direct reproductive regulators and downstream developmental gene targets. 1325 remaining candidate genes were expressed at FPKM<1 in the notch.

Notch candidates include ovule gene orthologs and regulators of post-fertilization development

52.0% of notch candidates possessed GO term annotation (**Supplementary Fig. 3**). Cell division and the cell cycle were the second and fourth most-prevalent categories of significantly-enriched GO terms ($p < 0.05$), reflecting the notch's meristematic status. Fewer reproductive GO terms were enriched than in the sporophyll (1.8%) but a greater proportion related to gene regulation (13.7%), including additional post-transcriptional and epigenetic categories.

Hierarchical clustering of candidates by normalised expression successfully identified six clusters with maximum abundance at different notch developmental stages (**Fig. 5a**). Clusters 3 and 4 were significantly enriched in notch-specific candidates ($P_{adj} < 0.05$) and 1, 2 and 5 in notch-enriched candidates (**Supplementary Data 6**). Reproductive GO terms were significantly enriched ($p < 0.05$) predominantly in clusters 2 and 3, associated with the Aeb (**Fig. 5b**). A subset were significantly enriched in transcripts with significantly greater abundance ($P_{adj} < 0.05$, $\log_2\text{fold} > 1$) in Dem vs Aeb (**Fig. 5c**), including *CMADS1* (**Supplementary Data 3**). Surprisingly, many enriched reproductive terms were meiosis-related (**Fig. 5b**), identifying the same *ZIP4* and *SGO* transcripts as in the sporophyll and orthologs of *BRCA2* and *RAD51*^{52,53} (**Supplementary Data 3**). Clusters 2 and 3 were also specifically significantly enriched in post-transcriptional and epigenetic GO terms (**Extended Data Fig. 5a,b**). The associated candidate genes included orthologs of *Arabidopsis* genes from across ovule and seed development (**Extended Data Fig. 6**) and two previously-published *Ceratopteris* embryo-expressed genes (*CrANT* and *CrLFY1*)^{54,55} (**Supplementary Fig. 8**). Reproductive GO terms were not significantly enriched ($p > 0.05$) in pre-fertilization archegonia (cluster 4) (**Fig. 5b**), but enriched gene regulatory GO-terms identified orthologs of the meristematic genes *CLAVATA1* (*CLV1*)⁵⁶ and *LATERAL SUPPRESSOR* (*LAS*)⁵⁷ (**Supplementary Data 3**). Interestingly, pollen recognition GO terms linked to homologs of pollen self-recognition genes⁵⁸ were associated with archegonia, being significantly enriched in 4 DAF embryo-less archegonia (cluster 5) (**Fig. 5b**), notch-specific candidates (**Supplementary Fig. 9**) and in genes with increased abundance in Aeb vs Dem (**Fig. 5c**). These results suggest that angiosperm reproductive networks, including seed development, are conserved with the fern archegonium.

Further evidence for this hypothesis emerged, with two angiosperm post-fertilization seed mechanisms identified and validated in *Ceratopteris*. Hormone response GO terms significantly enriched prior to and during fertilization (clusters 4 and 6) and in candidates with significantly greater abundance in Aeb vs Dem (**Extended Data Fig. 7a-c**) were found to relate

mostly to auxin (**Supplementary Data 3**), a critical regulator of post-fertilization seed development that can drive parthenocarpic seed and fruit development^{59,60,61}. To test its role in *Ceratopteris* fertilization-dependent development, exogenous auxin was applied to gametophytes at flooding. Auxin-treated gametophytes successfully bore embryos (**Extended Data Fig. 7d**), while embryo-less archegonia of the same gametophytes exhibited ectopic cell divisions not detectable in mock-treated controls (**Extended Data Fig. 7e,f**). Carbohydrate metabolism-related GO terms exhibited similar archegonium-specific enrichment (**Extended Data Fig. 8a-c**). In angiosperms, sugar signalling inhibits post-fertilization seed abortion^{62,63} and promotes apogamy in *Ceratopteris*⁶⁴. To test the effect of carbohydrate status on *Ceratopteris* sexual reproduction, gametophytes were flooded with a 2% sucrose solution vs water. The proportion of gametophytes bearing embryos was not affected by sucrose treatment ($p = 0.5393$) but the mean number of embryos per embryo-bearing gametophyte significantly increased ($p < 0.0001$; **Extended data Fig. 8d-g**). These results suggest post-fertilization regulatory mechanisms are conserved between the fern archegonium and angiosperm seed.

***Ceratopteris* sporophyll and notch candidate gene networks are conserved with distinct *Arabidopsis* reproductive organs**

Ceratopteris reproductive candidates partially overlap between the sporophyll and notch (44.5% and 11.9%, respectively) (**Fig. 6a**), suggesting distinct networks. We quantified conservation between these two organs and *Arabidopsis* reproductive gene networks using gene orthology, identifying all *Arabidopsis* gene orthologs of these populations (**Fig. 6b,c**; **Supplementary Fig. 10**; **Supplementary Data 7**) and testing for enrichment of these within published *Arabidopsis* reproductive datasets^{65,66,67,68,69,70,71,72,73,74,75,76,77} compared to vegetative-specific genes⁷³.

Sporophyll orthologs were significantly enriched ($P_{adj} < 0.05$) only in stamen datasets (**Fig. 6d-i**), including *SPOROCTELESS* (*SPL*)-dependent genes relating to sporogenesis^{78,79}. This persisted in sporophyll-enriched orthologs only (**Supplementary Data 8**). Notch orthologs were significantly enriched in stamen, pollen, carpel and ovule datasets (**Fig. 6d-i**), with orthologs of notch-enriched candidates showing additional enrichment in post-fertilization seeds (**Extended Data Fig. 9a-c**). Notch orthologs were significantly enriched in all ovule tissue datasets tested except the integument (**Fig. 6h**). Common orthologs were not significantly enriched ($P_{adj} > 0.05$) in any dataset (**Supplementary Data 8**), supporting separate conservation of gene networks between the sporophyll and the stamen, and between the notch and the stamen, pollen and ovule/seed.

Notch-enriched orthologs significantly-enriched within the ovule formed two highly-correlated groups, between sporophytic and gametophytic tissues (**Extended Data Fig. 9d; Supplementary Fig. 11**). However, most tissue-specific and commonly-expressed subsets within and between these two groups remained significantly enriched in notch-enriched orthologs (**Extended Data Fig. 9e-g**). 43.7% of orthologs significantly-enriched in the seed were not present in the ovule (**Supplementary Fig. 11**). Notch-enriched orthologs were more equally correlated between seed tissues and commonly-expressed subsets remained significantly enriched (**Extended Data Fig. 9h,i**). However, we found that seed-specific genes were not significantly enriched in notch-enriched orthologs ($P_{\text{adj}} > 0.05$) (**Extended Data Fig. 9j**), suggesting a common evolutionary origin for ovule and seed gene networks from the notch.

To determine if specific stages of *Ceratopteris* sporophyll and notch development were conserved with *Arabidopsis* reproductive organs, reciprocal enrichment testing was undertaken using all *Ceratopteris* reproductive genes orthologous to the conserved genes significantly-enriched within each *Arabidopsis* organ (**Extended Data Fig. 10a,b; Supplementary Data 9**) within sporophyll and notch expression clusters. Sporophyll orthologs of conserved stamen genes were not significantly enriched in any cluster compared to all sporophyll candidates ($P_{\text{adj}} > 0.05$; **Extended Data Fig. 10c,d**). Notch orthologs of conserved stamen genes were significantly enriched ($P_{\text{adj}} < 0.05$) (**Extended Data Fig. 10e**), which could be ascribed to conserved pollen genes. Conserved ovule genes were significantly enriched in the 4 DAF Aeb (clusters 2 and 3), which could be ascribed to genes conserved with the nucellus, MMC and embryo sac central cell (**Extended Data Fig. 10f-i**). To test the developmental origin of post-fertilization seed gene networks, enrichment testing was repeated using orthologs of notch-enriched genes conserved in *Arabidopsis* (**Supplementary Fig. 12**). Conserved embryo, endosperm and seed coat orthologs were significantly enriched in cluster 3, further supporting the conservation of gene networks between the ovule/seed and post-fertilization archegonium.

DISCUSSION

Our mRNA-seq analysis identified gene conservation at points throughout *Ceratopteris* and *Arabidopsis* reproductive development, supporting the hypothesis that seed plant reproductive innovations arose by repurposing ancestral reproductive gene networks. This allows us to provisionally assign development origins to these networks.

GA signalling was identified and validated as promoting the *Ceratopteris* sporing transition. GA signalling promotes the *Arabidopsis* floral transition by indirectly upregulating floral

meristem identity genes⁸⁰ including MADS-box genes⁸¹. MADS-box gene clades regulating floral meristem and floral organ identity are angiosperm or seed-plant specific^{4,82} but *Ceratopteris* homologs were identified with expression patterns consistent with downstream targets of the sporing transition. This suggests that the GA module of angiosperm shoots was co-opted from an ancestral mechanism in the sporing transition.

Angiosperm MADS-box proteins specify different reproductive organ identities during flower development, with similar functions proposed in gymnosperms²⁵. We further identified MADS-box genes (including *CMADS1*) in early sporophyll development, together with an upstream floral regulator of MADS-box activity (*MRLK*). These were sporophyll-enriched rather than specific, and previously-published *CMADS1 in situ* localisation data²⁴ supports dual vegetative/reproductive expression. Sporophyll candidates were also significantly enriched in stamen genes, consistent with the *Adiantum capillus-veneris* homosporangium³⁰. Floral-specific networks may thus have arisen from an ancestral sporophyll programme.

We identified orthologs of genes regulating male and female meiosis within sporophyll candidates^{40,41,42,43} but two of these and further meiosis genes (*BRCA2*, *RAD51*)^{52,83} with highly-conserved functions^{84,85,86,87} were also identified in the post-meiotic gametophyte notch. *AtBRCA2A* and *-B* have a dual somatic function⁸³, likewise paralogs of *AtRAD51* (*AtRAD51C*)⁸⁸. Their presence in the actively-dividing notch could reflect a similar role in *Ceratopteris*. Notch candidates also included orthologs of female-specific meiosis regulators⁸⁹: *AtDRM1* and *-2* prevent the formation of ectopic MMCs⁹⁰. These results suggest that homosporous meiosis genetically resembles seed plant microsporogenesis, consistent with their morphological similarity¹⁹, while female meiosis could have diverged by co-opting notch gene networks already utilising meiosis genes.

Contrary to expectations, *Arabidopsis* ovule/seed genes were conserved specifically with the *Ceratopteris* post-fertilization archegonium. This raises the intriguing hypothesis that the diploid ovule/seed body-plan arose by co-opting gene networks from archegonium development. Although counterintuitive, this accords with the striking functional and developmental similarities between the post-fertilization fern archegonium and the seed¹¹, together with the conservation of carbohydrate status and auxin regulating archegonium and seed post-fertilization development. The feasibility of mis-expressing developmental programmes between gametophyte and sporophyte generations has been demonstrated in *Physcomitrium patens*, where loss of *CURLY LEAF* caused activation of sporophyte development in the gametophyte⁹¹. It is therefore possible that mis-activation of the post-

fertilization archegonium gene network during sporophyll development contributed to the formation of the ovule.

ONLINE METHODS

Plant material and growth conditions

Ceratopteris richardii lab strain Hn-n⁹² was used for all experiments. Spore sterilization, plating and gametophyte growth on C-fern 1% agar medium (pH6.0) under tissue culture conditions were performed as previously described. Plates were kept humid by adding 1.5 ml sterile water at 3 and 7 days after sowing, with no water added beyond this time to prevent premature sperm release and fertilization. Fertilization was triggered at 14 days by flooding plates with 7.5 ml sterile water, with embryo development assumed to have begun on the day of flooding. Individual sporophytes transferred individually to fresh C-fern agar media once visible (9-11 days after flooding [DAF]). Individual sporophytes were transferred to sterile GA-7 Magenta vessels containing 100 ml C-fern 1% agar medium (pH6.0) when 2-3 fronds and elongating roots were visible (14-16 DAF) and wetted with 5 ml sterile water. Sporophytes were subsequently maintained in Magenta vessels until 130 days old (**Supplementary Fig. 13**), with sterile water re-applied whenever no surface water was visible (typically every 2-3 weeks). At 100 and 120 days 5 ml of sterile C-fern liquid media (pH6.0) was applied instead of water to maintain healthy growth. Sporophytes grown for harvesting 25 day-old samples were transferred to 90 mm plates at a density of 30-40 per plate and grown there until harvested. All plant material was grown in an MLR-352H-PE controlled environment cabinet (PHC Europe, Breda, Netherlands) at 28°C, constant 50% relative humidity and a 16hr light/ 8hr dark photoperiod at a fluence of 150 $\mu\text{mol m}^{-2} \text{s}^{-1}$.

Embryo development synchronisation

To confirm the efficacy of restricted watering at synchronising embryo development, populations of gametophytes were grown at a density of 200-300 spores per plate under tissue culture conditions, with parallel growth regimes applied either restricted watering (as described above) or 'well-watered' conditions with sterile water applied at 9 and 11 DAS to maintain a constant layer of surface water, at ratio of 2 restricted water:1 well-watered. At 14 DAS, all well-watered plates and half of the restricted water plates were flooded with sterile water as described above ('+Flooding'), with the remaining restricted water plates left dry ('-Flooding'). One plate per treatment was sacrificed for analysis at 14 DAS (Restricted-water, Well-watered), 4 DAF and 7 DAF (Restricted water -Flooding, Restricted water +Flooding, Well-watered +Flooding). From each plate 50 hermaphrodite gametophytes were selected at random and scored for the presence of an embryo using an ultraZOOM-2 stereomicroscope

(GT Vision Ltd, Newmarket, UK) at 20x magnification. The experiment was replicated on four different occasions.

To characterise the timing of post-fertilization embryo development and archegonium production in response to flooding, populations of gametophytes were grown at a density of 200-300 spores per plate under restricted water conditions (as described above) and flooded at 14 DAS. At 0 (unflooded), 1, 2, 3, 4, and 7 DAF single plates were sacrificed for analysis. From each plate 50 hermaphrodite gametophytes were selected at random and scored for the presence of an embryo (as described above). At 1-7 DAF gametophyte selection was restricted to those with evidence of fertilization in response to flooding (brown archegonia). The experiment was replicated on five different occasions. During three of these replicates the number of archegonia per gametophyte were counted alongside scoring for the presence of an embryo. Numbers of embryo-bearing and embryo-less flooded gametophytes recorded per replicate varied due to availability within the population.

Sporing transition characterisation

Sporophytes were grown in Magenta vessels (as described above) in populations of 25. Eleven replicate populations were grown, with populations initiated at intervals over 13 months. From 4 weeks old onwards, the number of sporophytes with visible sporophylls within each population was recorded once per week ± 1 days until 19 weeks old, by which time 100% of sporophytes had begun to spore. For nine of these replicates the number of fronds per sporophyte was also recorded weekly (**Supplementary Data 1**).

Sporangium development characterisation

Single sporophylls were collected from 130d sporophytes, secured dorsal-side down to microscope slides using double-sided tape and pinnae dissected open under a stereomicroscope using hypodermic needles as micro-scalpels to expose developing sporangia on the ventral surface. Sections of pinnae were recorded photographically to capture sporangium planar area. Sporangia were recorded at two locations per pinna (base and tip) and from two pinnae per frond, positioned basally and distally. Four sporophytes were sampled per frond developmental stage. Sporangia areas were measured from scaled micrographs using the FIJI image analysis software package⁹⁴.

mRNA-seq sample preparation

Tissue samples for mRNA-seq were harvested by dissection and immediately frozen in liquid nitrogen. Sequential stages of sporophyte frond development were harvested within 2 hours per population grown, with replicate populations harvested between 5 and 10 hours after dawn

(ZT5-ZT10). Frond developmental stages were determined by eye based on frond morphology. Dissection of Apx and Prm samples was performed under a stereomicroscope using watchmaker's forceps. Tissues were frozen immediately after dissection by collection into tubes floating on liquid nitrogen. All biological replicates comprised tissue from at least 4 independent sporophytes (**Supplementary Data 10**), harvested from one (all Mat, Exp, Fid, 60d Apx and Prm), 2-3 (130d Apx and Prm) or 3-4 independently-harvested populations (25d Apx and Prm), respectively. Exp samples were designed to encompass 'early' (unfurling), 'mid' (unfurled to 50% final size) and 'late' frond expansion (50-75% final size) (**Extended Data Figure 2**) within each biological replicate. Samples across frond development comprised tissue from the same harvested population(s) per biological replicate.

Gametophyte stages were necessarily harvested on separate days, individual populations harvested within 2-4 hours with replicate populations harvested between 5 and 10 hours after dawn (ZT5-ZT10). Dissection of the notch was performed under a stereomicroscope using hypodermic needles as micro-scalpels, the developmental status of each individual confirmed by eye. Each biological replicate comprised notch or thallus tissue from 150-225 individual gametophytes (**Supplementary Data 10**) pooled from 1 (0 DAF, 1 DAF, 2DAF), 2-3 (4 DAF) or 3-4 independently-harvested populations (Dem, 761-788 embryos per sample) respectively. Notch and thallus from the same individuals remained paired when preparing biological replicates for sequencing.

RNA was extracted using the Qiagen RNeasy plant mini kit with on-column DNase treatment (QIAGEN, Hilden, Germany). RNA quality was determined using a 2100 Bioanalyser RNA 6000 nano assay (Agilent Technologies, Santa Clara, USA). Individual sample RIN values are supplied in **Supplementary Data 10**. Library preparation, 150bp paired-end Illumina sequencing was performed separately for sporophyte and gametophyte samples by Novogene (Novogene UK, Cambridge UK) to a depth of ≥ 20 million reads per sample. QC data is provided in **Supplementary Data 10**.

Imaging

Whole sporophytes and gametophyte plates were photographed using a Cybershot DSC-HXTV digital camera (Sony Corporation, Tokyo, Japan). Individual gametophytes and dissected tissue samples were photographed using an ultraZOOM-2 stereomicroscope with an attached GXCAM-U3PRO-X digital camera and GXCam software (GT Vision Ltd, Newmarket, UK). Photographs were minimally processed by adjusting brightness and contrast and sharpness for whole images.

Data analysis

Statistical analyses of characterization experiments were conducted in R version 4.2.2⁹⁵ using the car package⁹⁶. Pairwise comparisons were performed using two-tailed students T-tests or Mann-Whitney nonparametric tests (indicated within each figure), following analysis of the data structure using Shapiro-Wilk for Normal distribution⁹⁷ and Levene's Test for equal variance⁹⁸. Where necessary data was mathematically transformed to meet the assumptions of the statistical tests (indicated within each figure). A post-hoc correction for multiple comparisons was applied to p-values using the Benjamini-Hochberg method⁹⁹.

Transcript assembly against the *Ceratopteris richardii* genome v2.1³² using HISAT2¹⁰⁰ was performed by Novogene, as was annotation and quantification of transcripts and pairwise DE analysis (**Supplementary Data 2; Supplementary Data 5**). Orthology of *Ceratopteris* genes to *Arabidopsis* was identified using Orthofinder¹⁰¹ with default settings against the following proteomes: *Arabidopsis thaliana* (TAIR10)¹⁰², *Amborella trichopoda* (v1.0)¹⁰³, *Glycine max* (Wm82.a4.v1)¹⁰⁴, *Manihot esculenta* (v8.1) (https://phytozome-next.jgi.doe.gov/info/Mesculenta_v8_1), *Micromonas pusilla* CCMP1545 (v3.0)¹⁰⁵, *Ostreococcus lucimarinus* (v2.0)¹⁰⁶, *Oryza sativa* (v7.0)¹⁰⁷, *Physcomitrium patens* (v3.3)¹⁰⁸, *Quercus robur* (PM1N)¹⁰⁹, *Solanum lycopersicum* (ITAG4.0)¹¹⁰, *Selaginella moellendorffii* (v1.0)¹¹¹, *Spirodela polyrhiza* (v2)¹¹² and *Volvox carteri* (v2.1)¹¹³. Additional GO term annotations were assigned to *Ceratopteris* genes from gene orthologs in the model species *Arabidopsis*, *Oryza*, *Selaginella*, *Physcomitrium* and *Marchantia*. Higher-order Venn comparisons were performed in InteractiVenn¹¹⁴). Further analysis of transcriptome data was performed in R, with cluster analysis using the pheatmap package¹¹⁵ and GO term enrichment testing using the TopGO package¹¹⁶, employing Fisher's Exact Test. *Arabidopsis* reproductive gene datasets were obtained from past publications and filtered against published vegetative-specific genes (**Supplementary Data 8**). Enrichment testing of *Arabidopsis* orthologs of *Ceratopteris* genes against these datasets was performed in R using Fisher's Exact Test (one tailed, greater than). The same methodology was used for reciprocal enrichment testing of *Ceratopteris* orthologs of conserved *Arabidopsis* genes within candidate expression clusters. All other plots were prepared in R using the ggplot2¹¹⁷, ggvenn¹¹⁸ and ggcorrplot¹¹⁹ packages.

DATA AVAILABILITY

Raw sequencing read data is available via the NCBI Gene Expression Omnibus (GEO) data repository (reference GSE291236). All other raw and processed data from the experiments presented here are provided in online supplemental information.

ACKNOWLEDGEMENTS

This research was funded by Royal Society University Research Fellowship URF\R1\191326. We are grateful to Dr Juliet Coates and Prof Daniel Gibbs for helpful discussions about mRNA-seq experimental design, Drs Megan McDonald and Peter Keane for helpful discussions about data analysis and Profs Daniel Gibbs, Christine Foyer and Eugenio Sanchez-Moran for helpful discussions about the manuscript.

AUTHOR CONTRIBUTIONS

ARGP conceived and planned experiments, performed developmental characterisation, tissue harvesting, RNA extraction, GO term analysis and enrichment testing. MC performed gene orthology assignment. ARGP and MC wrote the manuscript.

FIGURES

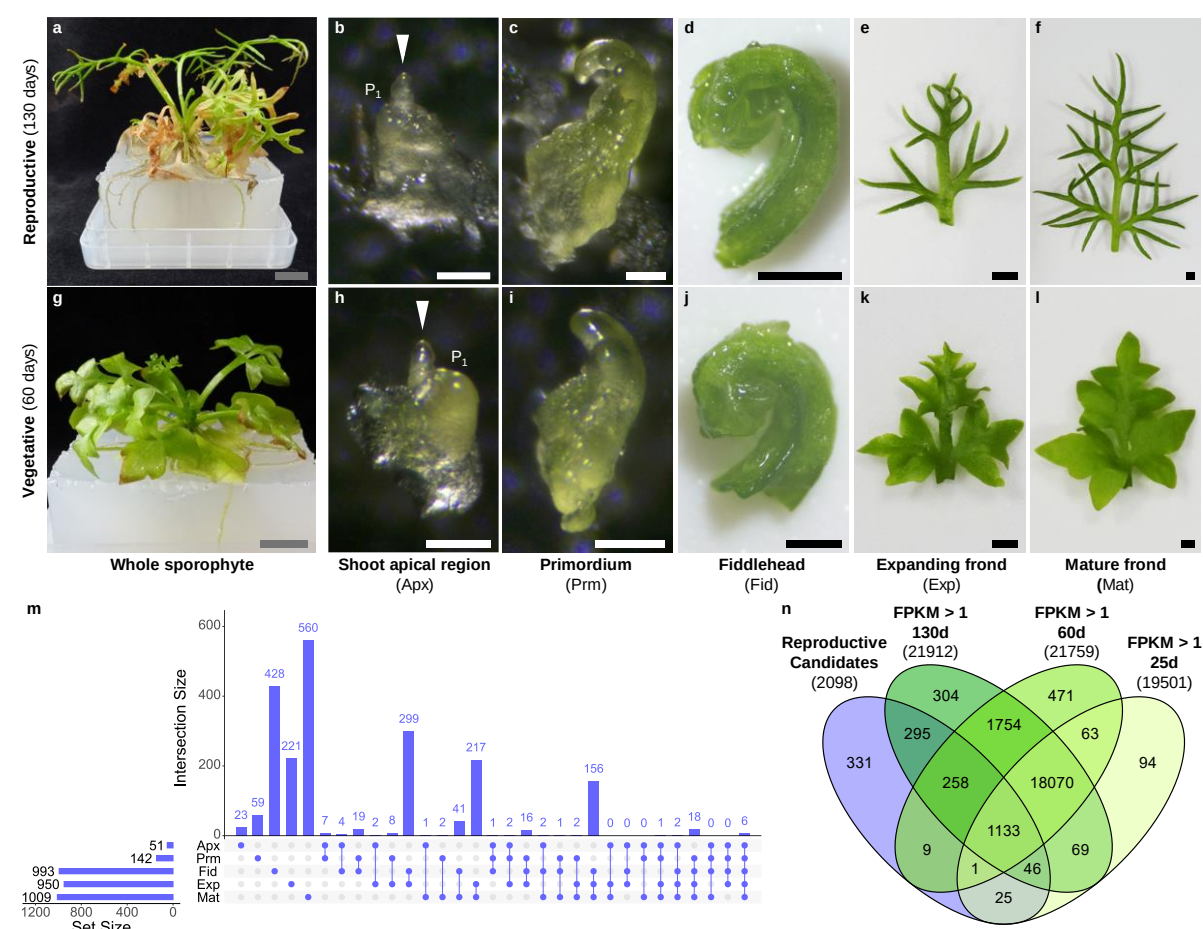


Fig. 1. Reproduction-associated genes were identified across *Ceratopteris* sporophyll development.

a-l, The spring *Ceratopteris* shoot (**a**) was dissected into sequential samples across frond development: the shoot apical region (Apex) (**b**) comprising the shoot apex (white arrowhead) and early primordia (P_1), the next-most recent primordium (Prm) (**c**) undergoing morphogenesis, the frond fiddlehead (Fid) (**d**) emerged from the shoot apical region, expanding fronds (Exp) (**e**) where the fiddlehead had terminated, from unfurling to approximately 75% of maturity, and mature non-senescent fronds (Mat) (**f**) where growth had ceased compared to prior fronds. Candidate reproductive genes were identified through significantly increased abundance ($P_{adj} < 0.05$, log2fold expression change > 1) in sporophyll samples vs corresponding frond samples in the vegetative phase immediately preceding the sporing transition (**g-l**). Scale bars = 10 mm (grey), 200 μ m (white) and 1 mm (black).

m,n Comparing reproductive candidate reproductive genes between frond developmental stages identified 2098 candidates (**m**). These were categorised based on transcript abundance (**n**): candidates at FPKM >1 at 130d but not in vegetative phases as 'sporophyll-specific'; at FPKM >1 in both 130d and at least one vegetative phase as 'sporophyll-enriched'; and at FPKM <1 at 130d as 'sporophyll-FPKM<1'.

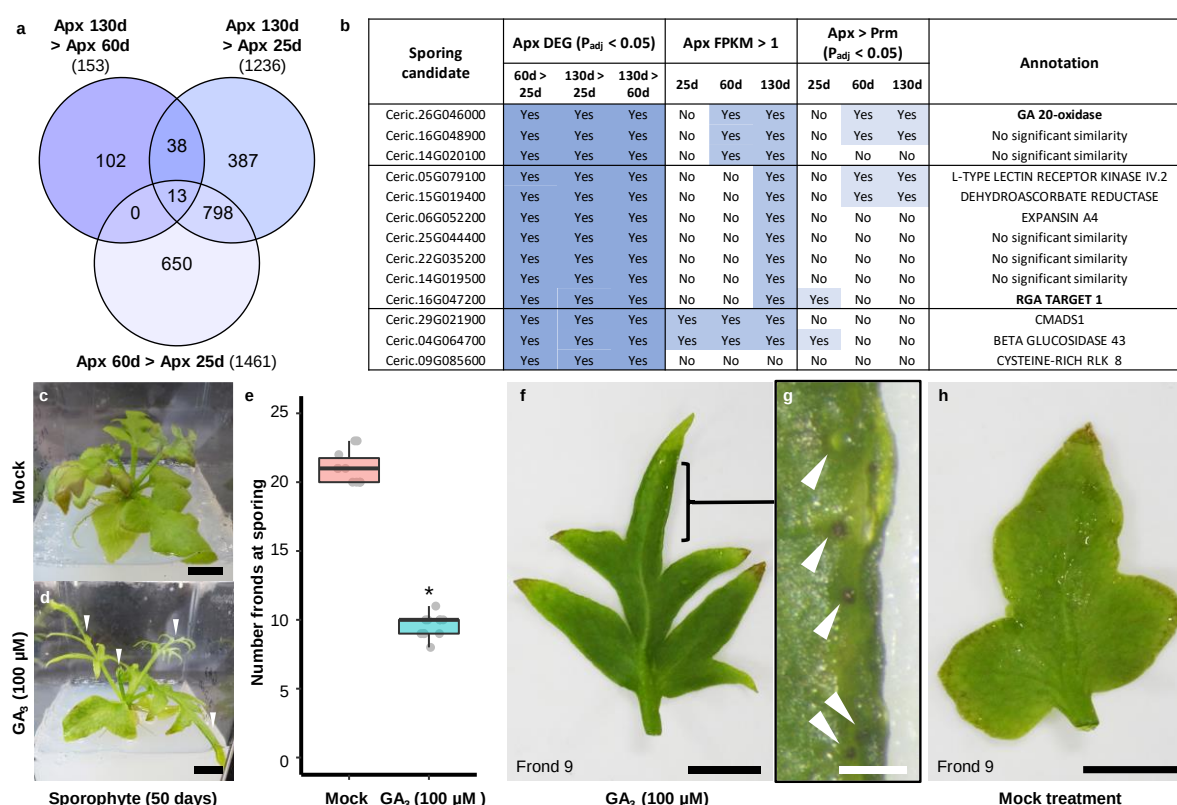


Fig. 2. Gibberellin promotes the *Ceratopteris* sporophyte shoot sporing transition.

a, Differential gene expression analysis identified 51 candidates with significantly greater expression in the 130d Apx vs 60d and 25d, with 13 showing a progressive increase immediately prior to the sporing transition (**a**). These 13 candidates were ranked according to their similarity to the predicted expression characteristics of a shoot apex sporing regulator (**b**), with the strongest candidate being a GA 20-oxidase gibberellin biosynthesis gene.

c-h, Under exogenous GA treatment the sporing transition of the *Ceratopteris* sporophyte is accelerated both chronologically (**c,d**) and developmentally (**e**). The first spore-bearing fronds under GA treatment (**f**) exhibit margin rolling and sporangium production (**g**) compared to mock-treated fronds at the same position (**h**). White arrowheads denote visible sporophylls (**d**) and sporangia within the rolled frond margin (**g**). Asterisk denotes a significant difference ($p < 0.05$) between mock and GA treatment. $n = 11$. Scale bars = 10 mm (black), 500 μ m (white). Pairwise comparisons were performed using two-tailed Mann-Whitney tests (**Supplementary Data 1**).

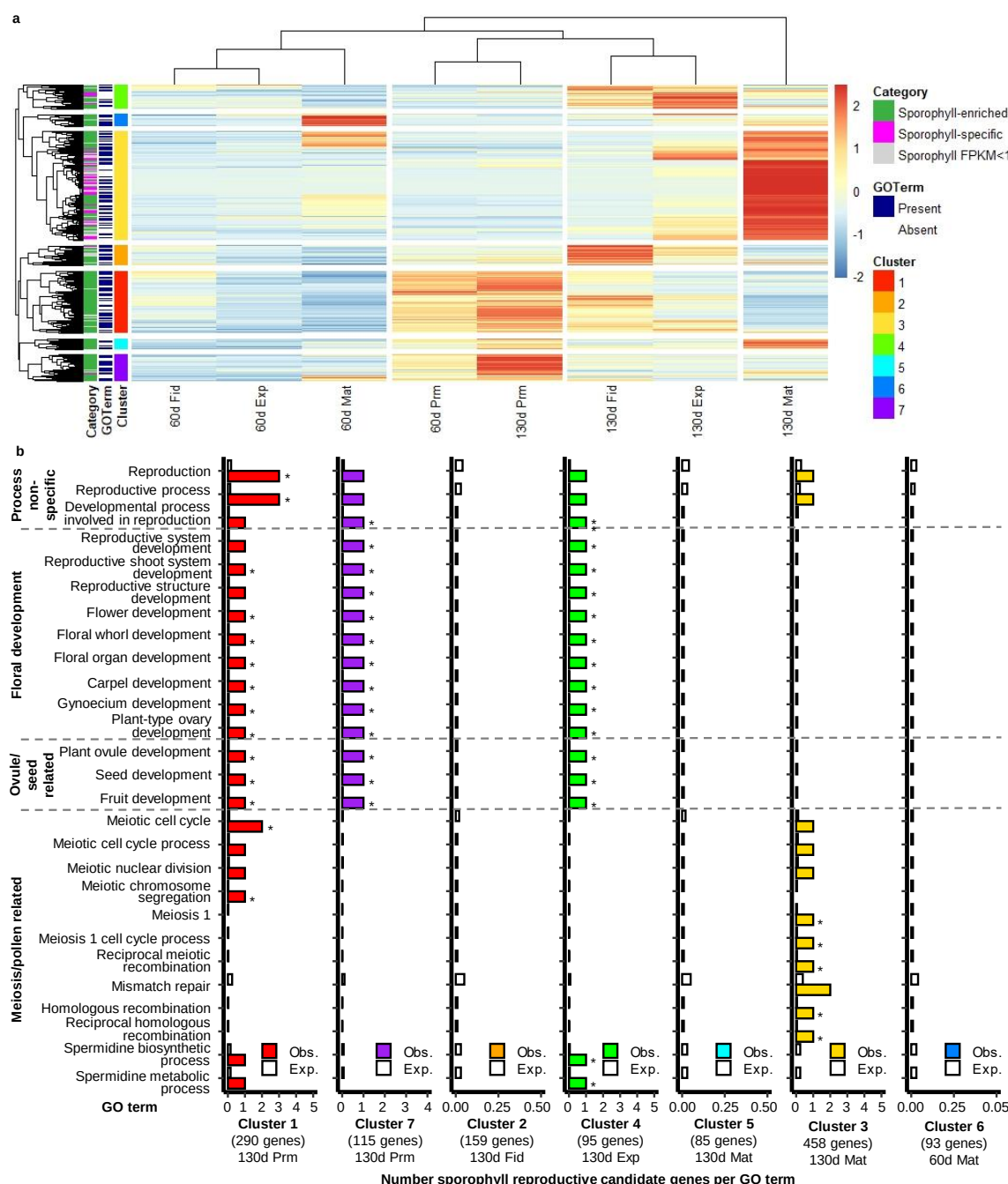


Fig. 3. *Ceratopteris* sporophyll candidate reproductive genes include orthologs conserved with *Arabidopsis* floral development and meiosis.

a, Clustering sporophyll reproductive candidate genes based on normalised expression identified seven discrete clusters of genes with maximum expression in specific stages of frond development.

b, Reproductive GO term enrichment ($p < 0.05$) across the sporophyll clusters identified in (a) in order of their peak in abundance during sporophyll development, showing the frequency of sporophyll reproductive candidate genes with enriched reproductive GO terms against the frequency expected based on the whole genome. All reproductive GO terms in which significant enrichment was detected are shown except spermine-related terms, which follow

the same pattern as spermidine (**Supplementary Data 3**). Enrichment of reproductive GO terms within different expression categories are shown in **Supplementary Fig. 4**.

Obs., observed frequency, Exp., expected frequency. Asterisks denote a significant increase ($p < 0.05$) in the frequency of genes with that GO term compared to the expected frequency.

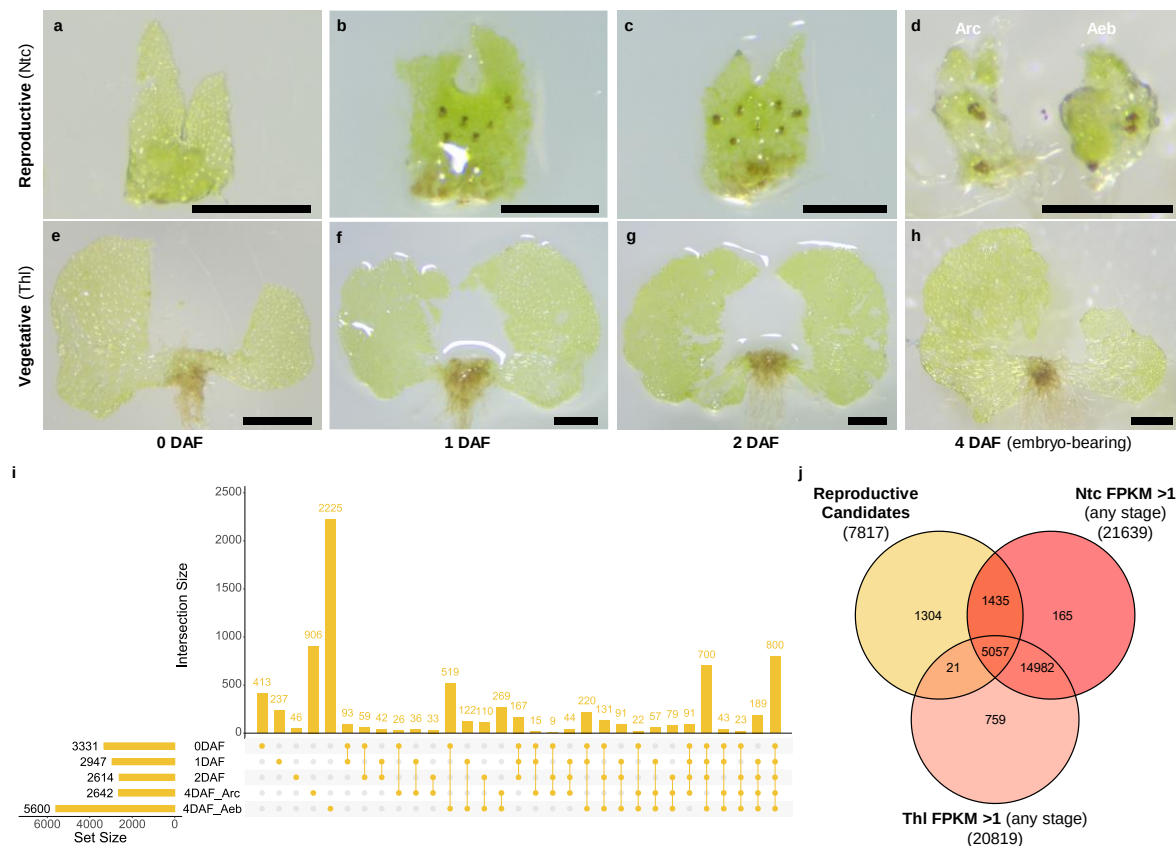
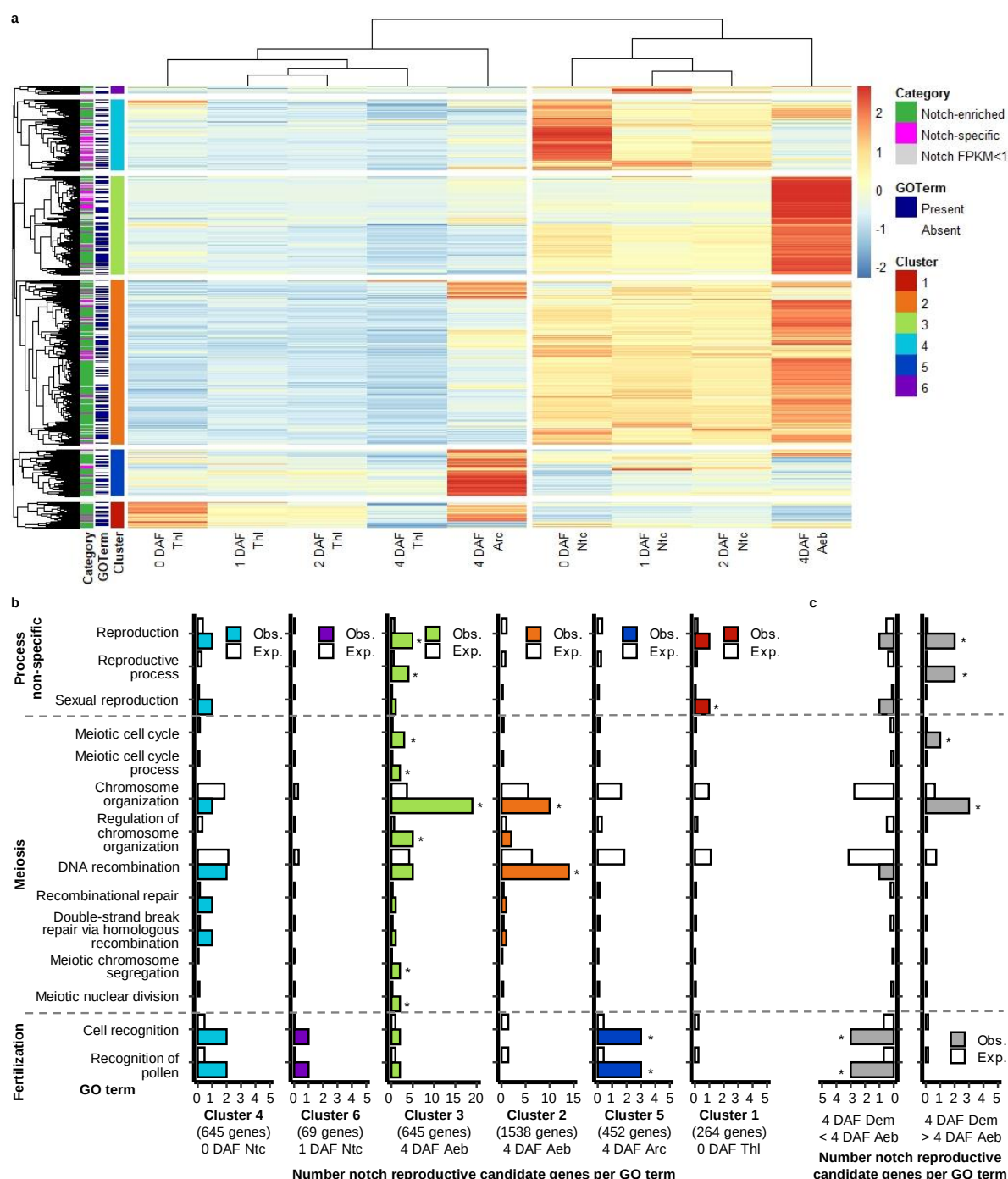


Fig. 4. Genes associated with archegonium initiation, fertilization and post-fertilization development were identified in the *Ceratopteris* gametophyte notch.

a-h, Reproduction-associated genes were identified by dissecting gametophytes into the notch region (Ntc), comprising the notch meristem and developing archegonia (**a-d**), and the surrounding thallus (Thl), containing vegetative tissue and antheridia (**e-h**), before fertilization (0 DAF; **a,e**), shortly after flooding (1 DAF and 2 DAF; **b,c,f,g**) and once an embryo was visible (4 DAF; **d,h**). 4 DAF Ntc samples were subdivided into embryo-bearing archegonia (Aeb) and embryo-less archegonia (Arc). Scale bars = 500 μ m.

i,j, 7187 nonredundant candidates were identified with significantly increased abundance ($P_{adj} < 0.05$, $\log_2$ fold expression change > 1) in Ntc vs Thl samples within at least one gametophyte developmental stage (**i**). These were categorised based on transcript abundance (**j**): candidates expressed at FPKM > 1 in at least one Ntc sample but not in any Thl sample were categorised as 'notch-specific'; those expressed at FPKM > 1 in both Ntc and Thl as 'notch-enriched'; and those with abundance of FPKM < 1 in all Ntc samples as 'notch-FPKM < 1 '.



(Supplementary Data 3). Enrichment of reproductive GO terms within different expression categories are shown in **Supplementary Fig. 9**.

Obs., observed frequency, Exp., expected frequency. Asterisks denote a significant increase ($P < 0.05$) in the frequency of genes with that GO term compared to the expected frequency.

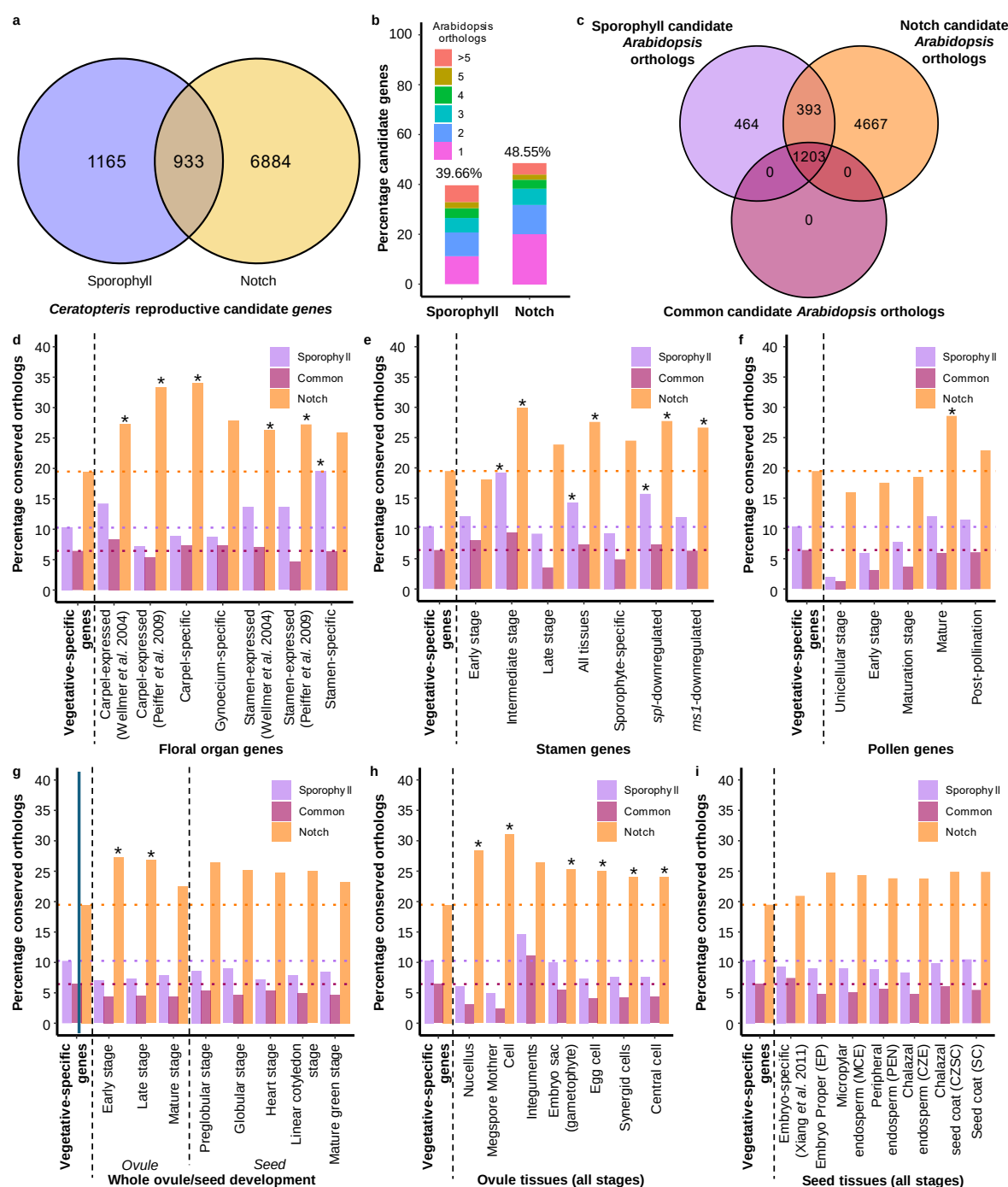
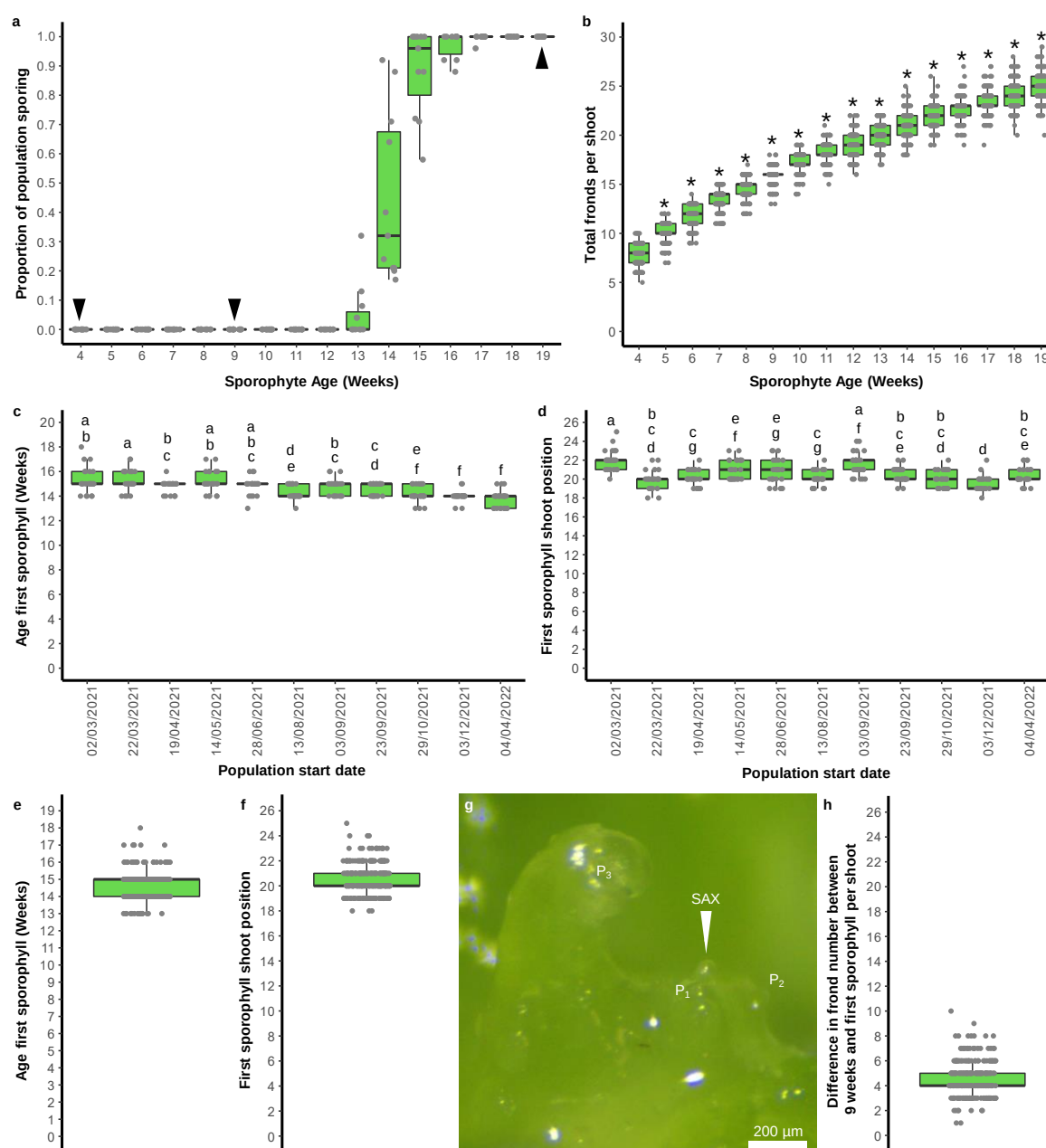


Fig. 6. *Ceratopteris* candidate sporophyll and notch gene networks show separate conservation with distinct *Arabidopsis* reproductive organ networks.

a-c, Sporophyll and notch reproductive candidates are partially overlapping (**a**). All direct *Arabidopsis* orthologs of these candidates were identified (**b**), of which 83.1% and 90.1% respectively possessed 5 *Arabidopsis* orthologs or less. Populations of nonredundant *Arabidopsis* orthologs for sporophyll, notch and common candidate genes were thus established (**c**) to enable enrichment testing against *Arabidopsis* reproductive gene datasets. **d-i**, Enrichment testing of *Arabidopsis* gene orthologs for the sporophyll, notch and common orthologs within published datasets of *Arabidopsis* reproductive genes expressed in floral

organs^{67,71} (**d**), stamen development⁶⁹ (**e**), pollen development^{65,66} (**f**), whole ovule/seed development^{68,73,77} (**g**), ovule tissues^{70,72,74,75} (**h**) and seed tissues^{76,77} (**i**) vs published vegetative-specific genes⁷³. Barplots show the overlap of candidate orthologs against each target as a proportion of that dataset. Asterisks denote a significant increase ($P_{\text{adj}} < 0.05$) in the proportion of overlapping genes between candidate orthologs and the target reproductive organ compared to the overlap with vegetative-specific genes. Dotted horizontal lines track the baseline overlap of each candidate ortholog dataset with the vegetative control. p-values for individual comparisons are provided (**Supplementary Data 8**).

EXTENDED DATA

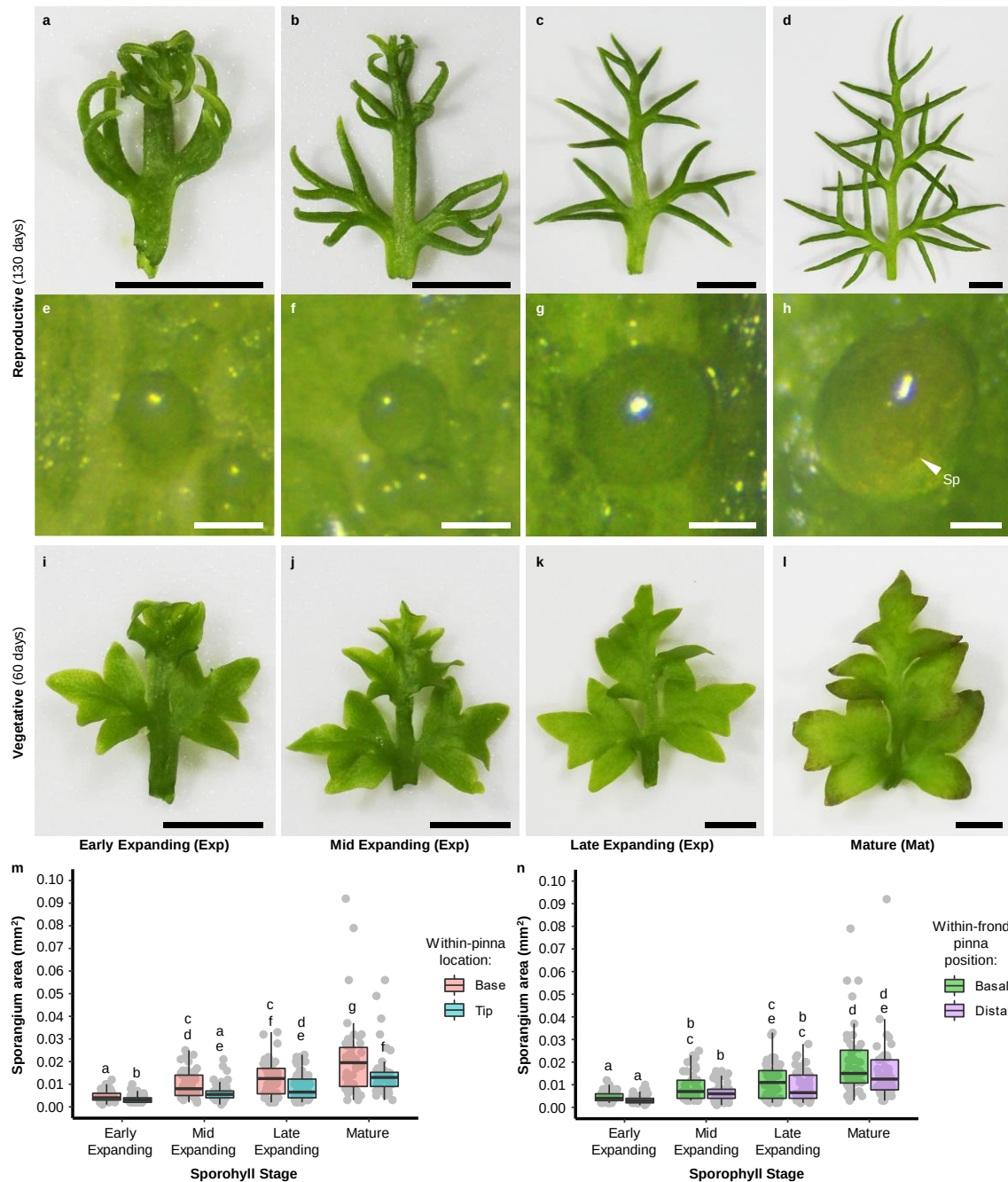


Extended Data Fig. 1. Characterising the timing and stability of sporing in the *Ceratopteris* sporophyte shoot.

a,b, Plots show the proportion of sporophytes sporing over time (**a**) and the total number of fronds per sporophyte over the same period (**b**), confirming that shoot development remained active throughout. Asterisks in (**b**) denote a significant difference ($P_{\text{adj}} < 0.05$) between the age indicated and the preceding week.

c-f, Earliest sporing time by chronological age and shoot position, within each population grown (**c,d**) and across all sporophytes (**e,f**). Black arrowheads in (**a**) indicate shoot ages sampled for mRNA-seq. Letters in (**c,d**) denote statistical similarity, with different letters indicating a significant difference ($P_{\text{adj}} < 0.05$) between experimental populations.

g,h The 9 week-old shoot apical region contains the shoot apex (SAX) and sequentially-developing primordia (P_1 - P_3) not visible to the naked eye and so not included in frond counts. Comparing the number of visible fronds at 9 weeks old and the first sporophyll shoot position (**h**) found a mean difference significantly greater than 3 ($p < 0.0001$, one-tailed Mann-Whitney test), indicating that 9 week-old shoots have not transitioned. Boxplots denote the data median and interquartile range (box) ± 1.5 x the interquartile range (whiskers). $n = 11$ (**a**), 225 (**b**) 25 (**c,d**), 271 (**e,f,h**). Pairwise comparisons were performed using two-tailed Mann-Whitney tests on square-root (**b**) and log₁₀-transformed data (**c,d**), respectively (**Supplementary Data 1**).

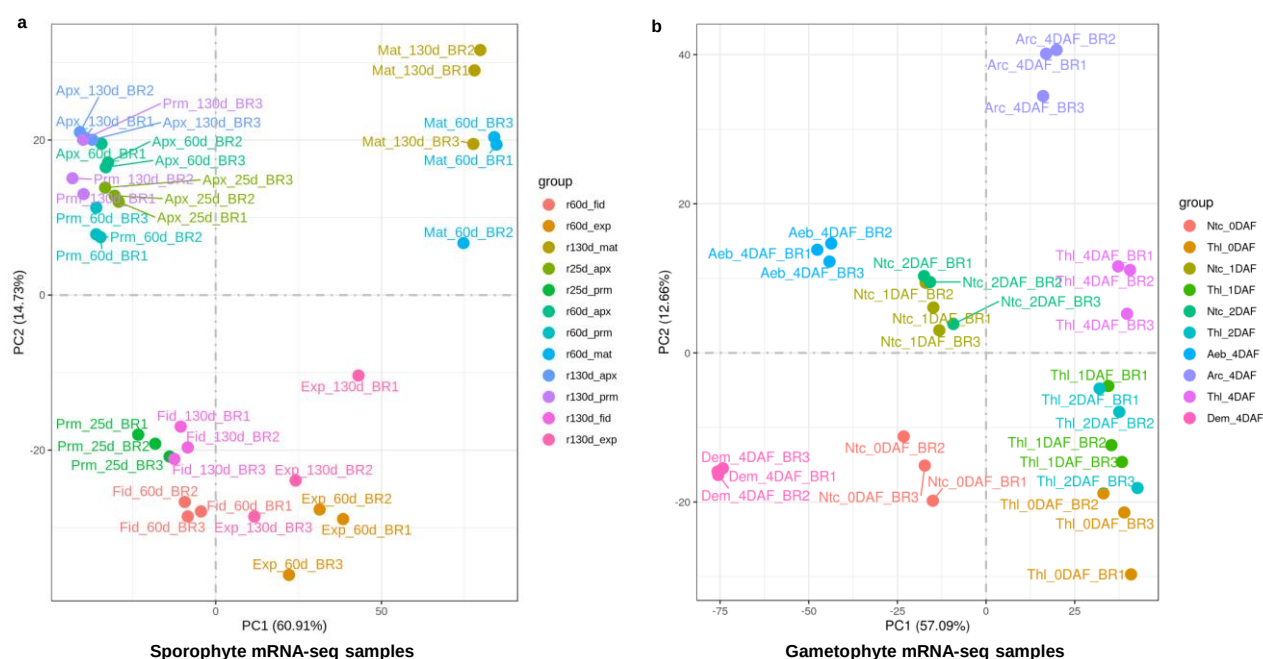


Extended Data Fig. 2. Characterisation of *Ceratopteris* sporangium development during sporophyll expansion.

a-h, Sub-stages within 'expanding sporophyll' (Exp) and representative sporangia, based on frond size: 'early expanding' (**a,e**, where fronds are still unfurling); 'mid expanding' (**b,f**, from newly-unfurled fronds to approximately 50% of final size); 'late expanding' (**c,g**, from approximately 50% to 75% of final size) and 'mature' (**d,h**) where fronds are of similar size to those immediately preceding it. Post-meiotic sporangia containing individual spores (Sp) were visible in mature frond samples only.

i-l, Expanding vegetative fronds from 60 day-old sporophytes, following the sub-stages established for sporophylls. Scale bars = 5 mm (black) and 100 μ m (white).

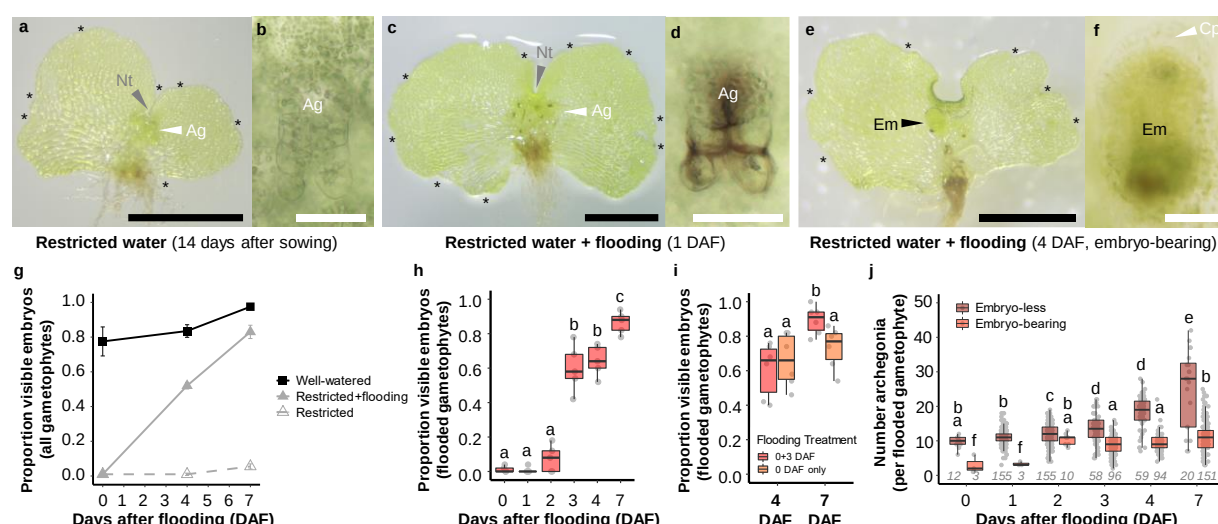
m,n, Sporangium planar area measured across sub-stages of frond expansion, comparing sizes within individual pinnae (**m**) and between pinnae within an individual frond (**n**). $n = 24$ (4 sporangia per pinna location, measuring 6 independent sporophytes). Boxplots denote the data median and interquartile range (box) $\pm 1.5x$ the interquartile range (whiskers). Letters denote statistical similarity, with different letters indicating a significant difference ($P_{adj} < 0.05$) between groups. Sporangium sizes were found to overlap between adjacent frond substages. Pairwise comparisons were performed using two-tailed Mann-Whitney tests on log10-transformed data (**Supplementary Data 1**).



Extended Data Fig. 3. Principal Component Analysis of *Ceratopteris* mRNA-seq data.

a, Biological replicates within sporophyte samples typically cluster more closely together than with replicates of other samples, the exceptions being primordia samples from 25d sporophytes (Prm_25d) which cluster more closely with fiddleheads. PC1 separates frond developmental stages progressively in order of developmental stage. PC2 separates vegetative and reproductive samples within each frond stage.

b, Biological replicates within gametophyte samples typically cluster more closely together than with replicates of other samples, with the exception of 1DAF vs 2DAF within notch and thallus samples. PC1 separates notch and embryo-bearing samples from thallus samples, with embryo-less 4DAF notch tissues (Arc_4DAF) falling between notch and thallus samples on this axis. PC2 separates samples progressively by age (with 1DAF and 2DAF overlapping), with the exception of dissected embryos (Dem_4DAF).



Extended Data Fig. 4. Restricted watering synchronises *Ceratopteris* fertilization.

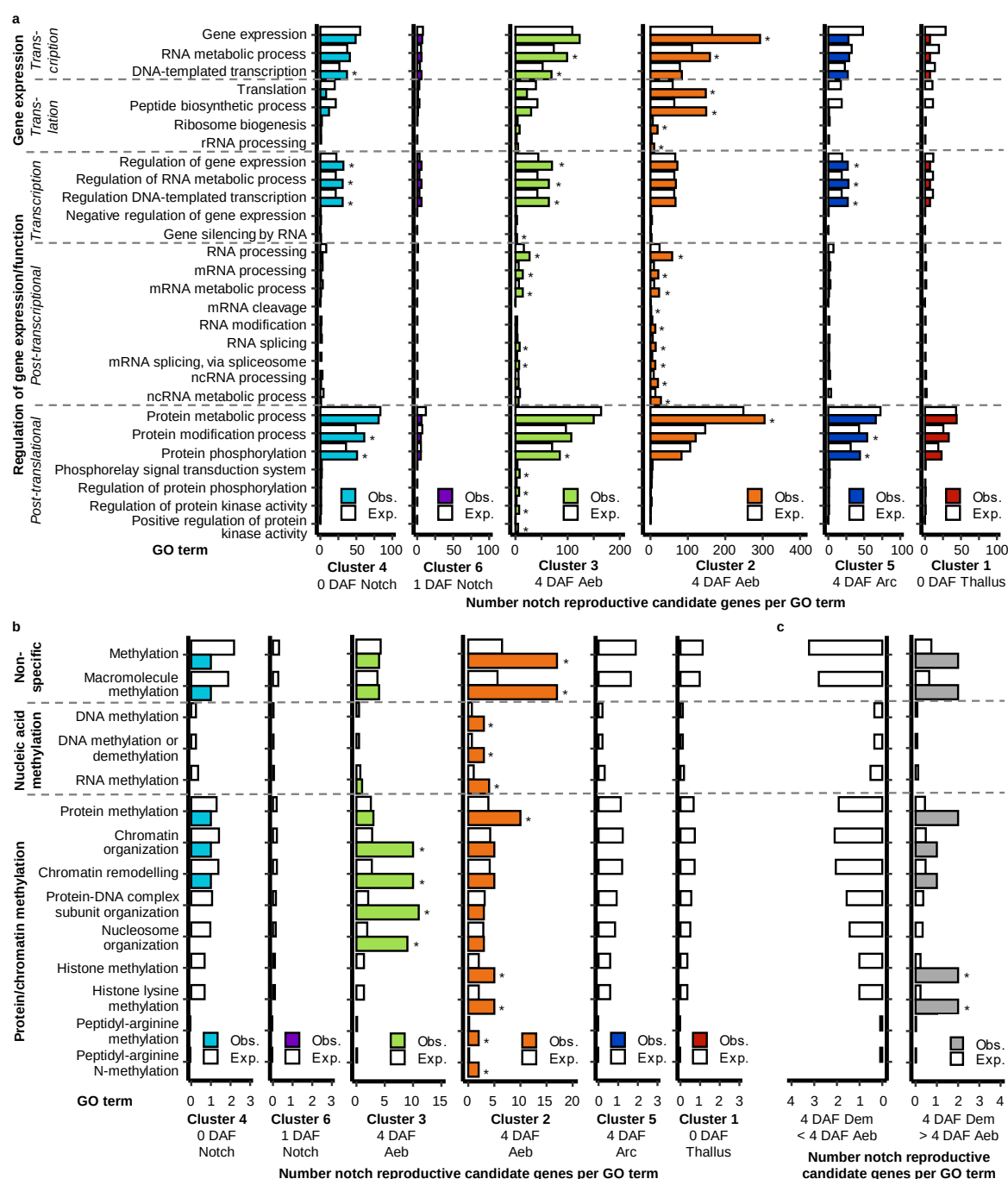
a-f, When grown with restricted water, mature unfertilized gametophytes (**a**) have green archegonia (Ag) (**b**) arising sequentially from the notch meristem (Nt), while antheridia (asterisks) develop at the thallus margin. By 1 day after flooding (DAF) (**c**) all archegonia turn brown (**d**). Typically a single embryo (Em) develops per gametophyte (**e**), the enclosing archegonium expanding into a calyptra (Cp) (**f**). Scale bars = 1 mm (black), 50 μ m (white).

g, Proportion of gametophytes (mean $\pm$ SE) with visible embryos under well-watered conditions or if watering was stopped at 7 days after spore sowing ('Restricted'). By 14 days after sowing (0 DAF) $77.5 \pm 12.5\%$ of well-watered controls bore embryos at various developmental stages (**Supplementary Fig. 6**) compared to $1.0 \pm 1.0\%$ under restricted water. Fertilization could be triggered in restricted-water populations by flooding, with $83.0 \pm 5.74\%$ bearing visible embryos by 7 DAF compared to $5.5 \pm 2.2\%$ without flooding. $n = 4$.

h-i, Proportion of flooded gametophytes with visible embryos per DAF. Most embryos were first visible by 3-4 DAF (**h**). If re-flooded subsequently ('0+3 DAF') this proportion increased between 4 and 7 DAF ($p < 0.05$) (**h,i**), suggesting additional fertilization events, but if flooded only once ('0 DAF only') the proportion did not increase beyond 4 DAF. $n = 5$.

i, Number of archegonia per flooded gametophyte per DAF, comparing between gametophytes with visible embryos ('embryo-bearing') and those without ('embryo-less'). Archegonium numbers stopped increasing in embryo-bearing gametophytes after 2 DAF but continued in embryo-less individuals. n shown in grey italic for each embryo status vs DAF combination.

Boxplots show the median and interquartile range (box) ± 1.5 x interquartile range (whiskers). Letters denote statistical similarity, with different letters indicating a significant difference ($P_{adj} < 0.05$) between groups. Pairwise comparisons were performed using two-tailed Mann-Whitney tests on logit (**h,i**) or log10-transformed (**j**) data (**Supplementary Data 4**).



Extended Data Fig. 5. Post-transcriptional gene regulation and epigenetic GO terms are enriched specifically in the *Ceratopteris* embryo-bearing archegonium.

a-c, Enrichment of GO terms relating to gene expression (**a**) or epigenetic regulation (**b**) within each expression cluster following the stages of notch development (see **Fig. 5**), and epigenetic GO terms enriched within genes showing significantly increased or decreased abundance in 4 DAF embryo-bearing archegonia vs 4 DAF dissected embryos (**c**). Bar plots show the observed ('Obs.') and expected ('Exp.') gene frequencies per GO term. Asterisks denote a significant increase ($p < 0.05$) in genes with that GO term compared to that expected based on the whole genome. Individual p-values are provided (**Supplementary Dataset 3**).

744

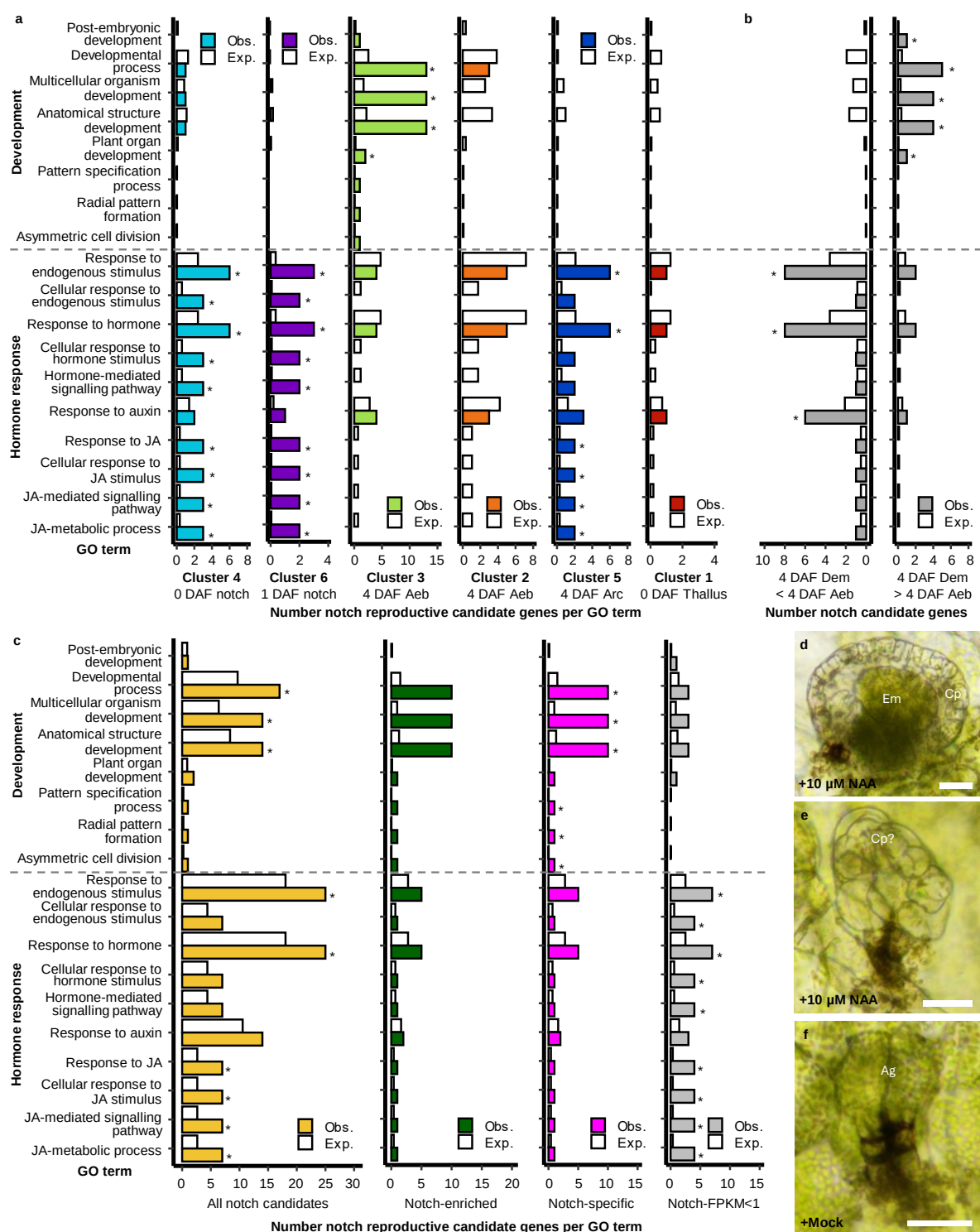
<i>Ceratopteris</i> transcript	Enriched GO term	Expression categories with GO term enriched	<i>Arabidopsis</i> orthologs	<i>Arabidopsis</i> ovule function	<i>Arabidopsis</i> molecular function
Ceric.26G015100.1	Regulation of DNA-templated transcription	Notch all; Notch-enriched; Notch Cluster 2; DEM > AEB	<i>BLH4, SAW2, BLH2, SAW1, BEL1</i>	Ovule identity ^{120,121}	Transcription factor
Ceric.19G034000.1	RNA methylation	Notch-enriched; Notch Cluster 2	<i>CRM2, HEN1</i>	Integument development ¹²²	RNA-dependent DNA Methylation (RdDM)
Ceric.12Z77300.1	RNA processing	Notch all; Notch-specific; Notch cluster 3	<i>ASU1, ATDCL1, CAF, DCL1, EMB60, EMB76, SIN1, SUS1</i>	Integument development ^{123,124}	microRNA regulation
Ceric.02G091600.1	DNA methylation	Notch-enriched; Notch Cluster 2	<i>DMT7, DRM2, DRM1</i>	MMC specification ⁹⁰	RNA-dependent DNA Methylation (RdDM)
Ceric.21G012500.1	Regulation of gene expression	Notch all; Notch-enriched; Notch cluster 3	<i>EDA32</i>	Female gametophyte development ¹²⁵	Endonuclease (putative)
Ceric.21G012700.1	Regulation of gene expression	Notch all; Notch-enriched; Notch cluster 3	<i>EDA32</i>	Female gametophyte development ¹²⁵	Endonuclease (putative)
Ceric.28G000900.1	Protein phosphorylation	Notch all; Notch-enriched; Notch cluster 3	<i>FU</i>	Female gametophyte development ¹²⁶	Ser/Thr protein kinase
Ceric.14G005300.1	RNA processing	Notch all; Notch-enriched; Notch Cluster 2	<i>DCL3</i>	Female gametophyte development ¹²⁷	RNA-dependent DNA Methylation (RdDM)
Ceric.22G046500.1	RNA processing	Notch all; Notch-enriched; Notch cluster 3	<i>DCL3</i>	Female gametophyte development ¹²⁷	RNA-dependent DNA Methylation (RdDM)
Ceric.09G088700.1	RNA splicing	Notch all; Notch-enriched; Notch Cluster 2	<i>EMB2776, LIS</i>	Egg cell specification ¹²⁸	mRNA splicing
Ceric.15G036100.1	rRNA processing	Notch all; Notch-enriched; Notch Cluster 2	<i>TOZ</i>	Embryogenesis (cell division plane) ¹²⁹	rRNA processing
Ceric.36G039200.1	Histone lysine methylation	Notch all; Notch-enriched; Notch Cluster 2; DEM > AEB	<i>KYP, SDG33, SUVH4</i>	Embryogenesis (imprinting) ¹³⁰	DNA methylation
Ceric.36G048900.1	Histone lysine methylation	Notch all; Notch-enriched; Notch Cluster 2	<i>KYP, SDG33, SUVH4</i>	Embryogenesis (imprinting) ¹³⁰	DNA methylation
Ceric.15G054600.1	Regulation of DNA-templated transcription	Notch all; Notch-specific; Notch cluster 3	<i>PLT2, PLT1, AIL6, PLT3, BBM, AIL5, CHO1, EMK, AIL7</i>	Embryogenesis; endosperm proliferation ¹³¹	Transcription factor
Ceric.32G008600.1	Regulation of DNA-templated transcription	Notch all; Notch-specific; Notch cluster 3	<i>PLT2, PLT1, AIL6, PLT3, BBM, AIL5, CHO1, EMK, AIL7</i>	Embryogenesis; endosperm proliferation ¹³¹	Transcription factor
Ceric.12G044400.1	Chromosome organization	Notch all; Notch-enriched; Notch cluster 3	<i>ATSMC2, ATCAP-E1, ATSMC4, SMC2, TTN3</i>	Endosperm proliferation/cellularisation ¹³²	DNA Condensin
Ceric.05G021600.1	Chromosome organization	Notch all; Notch-enriched; Notch Cluster 2	<i>SMC3, TTN7</i>	Endosperm proliferation/cellularisation ¹³²	DNA Cohesin

745

746 **Extended Data Fig. 6. *Ceratopteris* notch reproductive candidate genes enriched in the**
747 **embryo-bearing archegonium are conserved with *Arabidopsis* ovule/seed genes.**

748 List of notch reproductive candidate genes associated with GO terms significantly-enriched
749 within clusters 2 and 3 also orthologous to *Arabidopsis* genes with functions in ovule/seed
750 development. Genes are listed in order of the *Arabidopsis* ortholog's function during ovule
751 development, separating pre-fertilization events (dark green) from post-fertilization (blue). The
752 two clusters of *Ceratopteris* genes do not correspond to specific stages of ovule/seed
753 development. The enriched GO terms shown represent the most specific term associated with
754 each gene. Full lists of associated GO terms are provided (**Supplementary Data 3**).

755

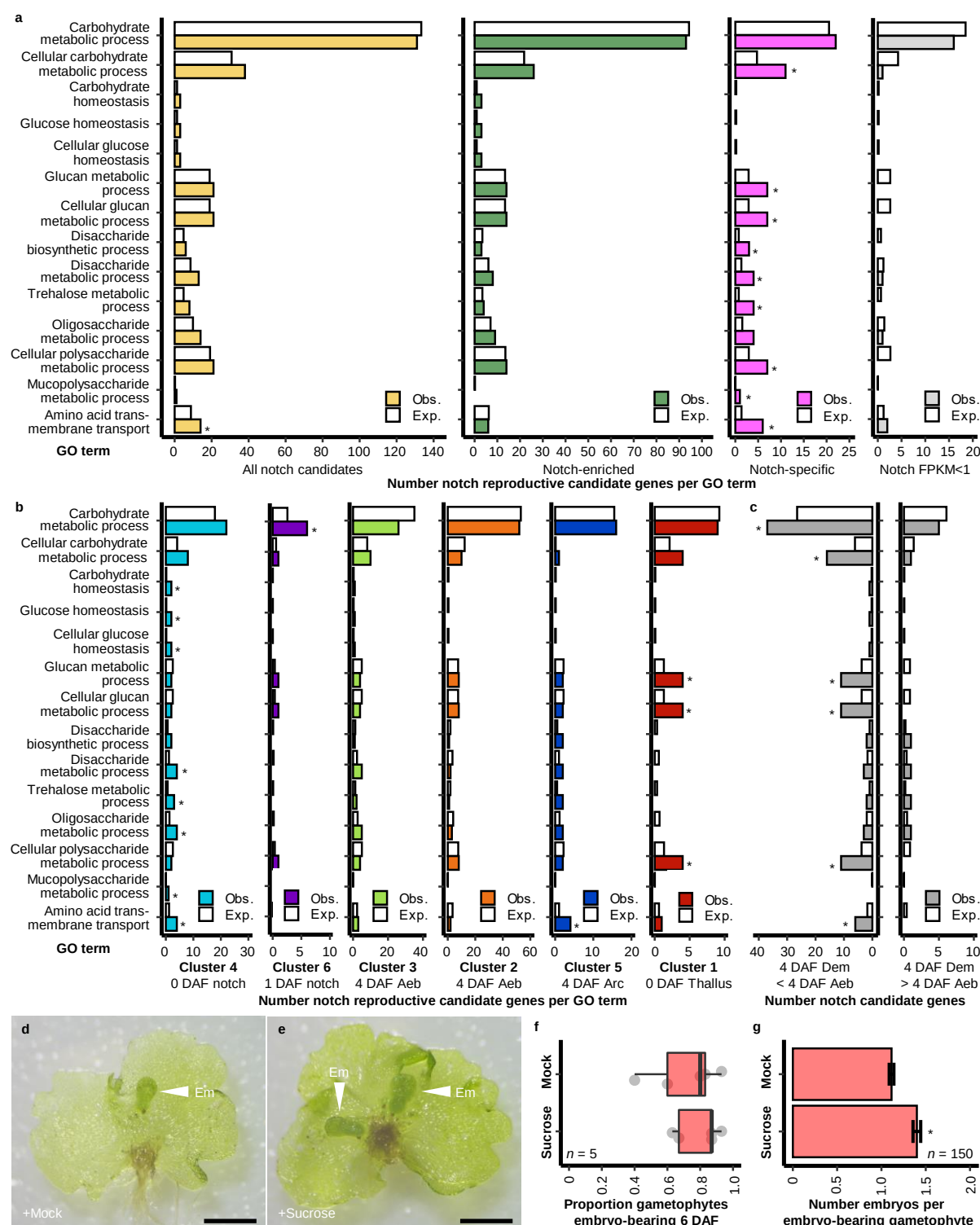


Extended Data Fig. 7. Auxin signalling promotes *Ceratopteris* fertilization-dependent archegonium development.

a-c, Enrichment of GO terms relating to development or hormone response within each notch expression cluster following the stages of notch development (**a**, see **Fig. 5**), within genes showing significantly increased or decreased abundance in 4 DAF embryo-bearing archegonia vs 4 DAF dissected embryos (**b**) and within each notch expression category (**c**). Bar plots show the observed ('Obs.') and expected ('Exp.') gene frequencies per GO term.

Asterisks denote a significant increase ($p < 0.05$) in genes with that GO term compared to that expected based on the whole genome. Individual p-values are provided (**Supplementary Data 3**).

d-f, Brightfield microscopy of archegonia from embryo-bearing gametophytes after flooding with 10 μ M NAA (**d,e**) or mock solution (**f**). At 4 DAF embryos were visible inside expanding calyptras under NAA treatment (**d**). Embryo-less archegonia of the same gametophytes exhibited ectopic cell division in response to NAA (**e**) compared to mock-treated controls (**f**). Scale bars = 50 μ m. Ag, archegonium, Cp, calyptra, Em, embryo.



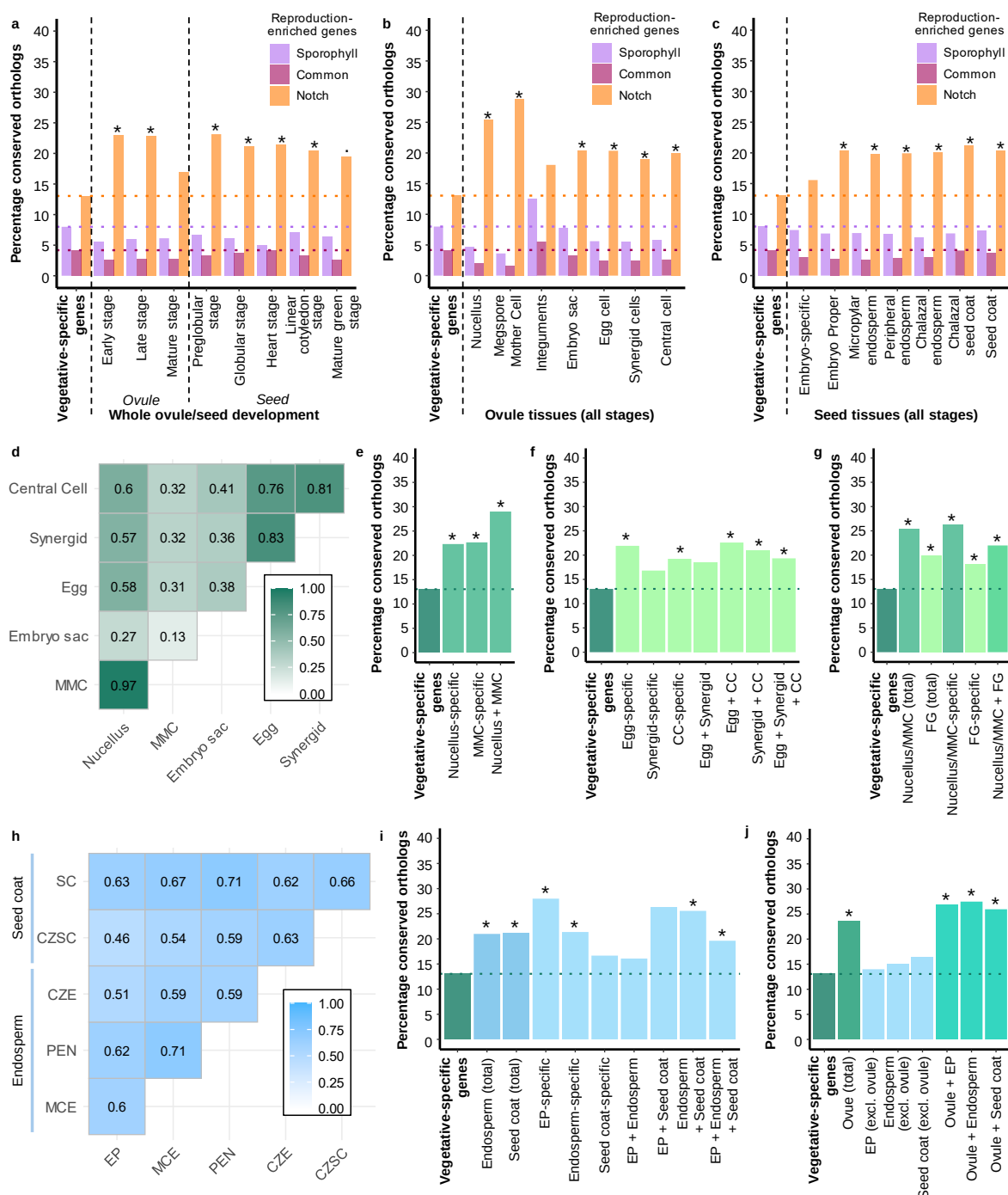
Extended Data Fig. 8. Carbohydrate status plays a role in regulating *Ceratopteris* embryo number.

a-c, Enrichment of GO terms relating to nutrients (carbohydrate metabolism and amino acid transport) within each notch candidate expression category (**a**), notch expression clusters following the stages of notch development (**b**, see **Fig. 5**), and within genes showing significantly increased or decreased abundance in 4 DAF embryo-bearing archegonia vs 4

DAF dissected embryos (**c**). Bar plots show the observed ('Obs.') and expected ('Exp.') gene frequencies per GO term. Asterisks denote a significant increase ($p < 0.05$) in genes with that GO term compared to that expected based on the whole genome. Individual p-values are provided (**Supplementary Data 3**).

d-g, Gametophytes typically carry one visible embryo by 6 DAF (**d**), but those flooded with 2% sucrose solution were frequently observed carrying two (**e**). The proportion of gametophytes bearing at least one embryo 6 DAF was similar between mock and sucrose treatments (**f**) whereas the mean number of visible embryos per embryo-bearing gametophyte (**g**) was significantly increased ($p < 0.0001$).

Scale bars = 1 mm. Boxplots denote the median and interquartile range (box) ± 1.5 the interquartile range (whiskers). Bar plots denote the mean $\pm$ SE. Asterisks denote significant difference ($p < 0.05$) between mock and sucrose treatment. n shown per experiment. Pairwise comparisons were performed using a two-tailed T-test on logit-transformed data (**f**) or a two-tailed Mann-Whitney test (**g**) (**Supplementary Data 4**).



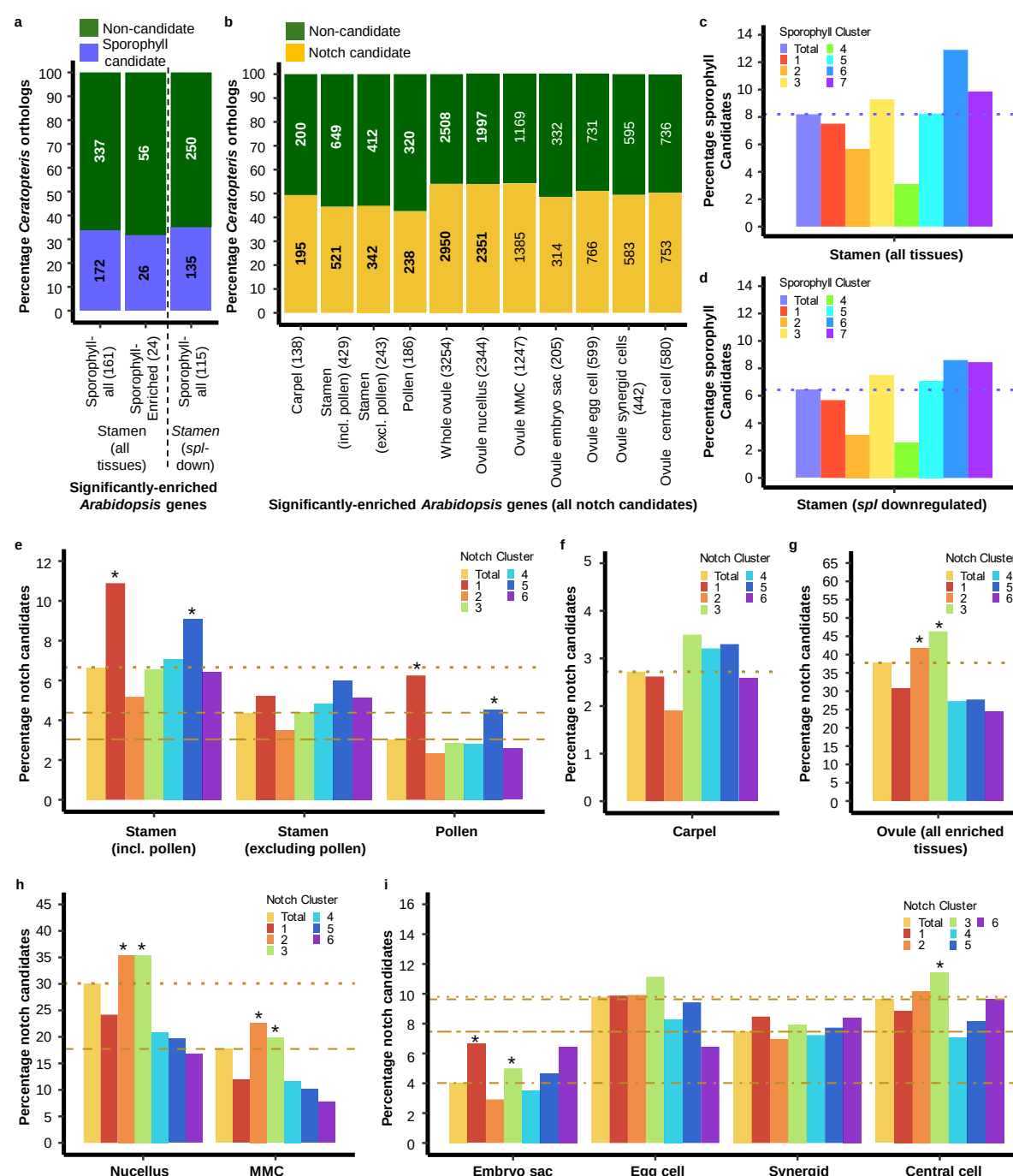
Extended Data Fig. 9. Notch-enriched candidate orthologs overlap with *Arabidopsis* ovule and post-fertilization seed gene networks, but not seed-specific gene networks.

a-c, Enrichment testing of *Arabidopsis* gene orthologs for sporophyll-enriched, notch-enriched and common orthologs in *Arabidopsis* whole ovule/seed (**a**), ovule tissue (**b**) and seed tissue datasets (**c**) against published vegetative-specific genes. Dotted horizontal lines track the overlap of each candidate ortholog dataset with the vegetative control.

d-j, Overlap of notch-enriched orthologs significantly-enriched between different tissues within the ovule (**d-g**), within the seed (**h,i**) and between the ovule and seed (**j**). Correlation plots show the Szymkiewicz–Simpson overlap coefficient between each tissue pair (**d,h**). Barplots

show enrichment testing of notch-enriched orthologs within genes specific to or common between tissues of the early ovule (**e**), tissues of the female gametophyte ('FG') (**f**) the total enriched genes of both (**g**), tissues within the post-fertilization seed (**i**) and all significantly-enriched ovule genes vs seed tissues (**j**). . Asterisks denote a significant increase ($P_{adj} < 0.05$) in the proportion of overlapping genes between candidate orthologs and the target dataset compared to the overlap with vegetative-specific genes. Individual p-values are provided (**Supplementary Data 8**). The assignment of individual genes to specific tissues is shown in **Supplementary Fig. 11**.

EP, embryo proper, MCE, micropylar endosperm, PEN, peripheral endosperm, CZE, chalazal endosperm, CZSC, chalazal seed coat, SC, seed coat



Extended Data Fig. 10. *Ceratopteris* orthologs of conserved *Arabidopsis* ovule genes are significantly enriched in the embryo-bearing archegonium.

a,b, Reciprocal identification of all *Ceratopteris* orthologs of *Arabidopsis* genes conserved between sporophyll (a) or notch candidates (b) and significantly-enriched in *Arabidopsis* reproductive organs/tissues (see Fig. 6). The number of conserved *Arabidopsis* genes per dataset is given in brackets. Plots show the proportion of *Ceratopteris* orthologs that are reproductive candidates in the sporophyll (a, blue) and notch (b, yellow) vs non-candidate genes (green). The number of *Ceratopteris* orthologs within each category is shown.

c,d, Enrichment testing within all *Ceratopteris* sporophyll candidate genes by expression cluster (see **Fig. 3**), testing *Ceratopteris* sporophyll candidate orthologs of conserved *Arabidopsis* stamen genes from all stamen tissues (**c**) or those specifically downregulated in *spl* mutant stamens (**d**).

e-i, Enrichment testing within all *Ceratopteris* notch candidate genes by expression cluster (see **Fig. 5**), testing *Ceratopteris* notch candidate orthologs of conserved *Arabidopsis* genes in the stamen (**e**), carpel (**f**), whole ovules (**g**) and individual tissues (**h-i**). Asterisks denote a significant increase ($P_{\text{adj}} < 0.05$) in the proportion of genes overlapping between candidate orthologs and each cluster compared to the overlap with all sporophyll or notch candidates, respectively. Individual p-values are provided (**Supplementary Data 9**).

SUPPLEMENTARY INFORMATION

Supplementary Fig. 1. Additional mRNA-seq samples and developmental ranges.

Supplementary Fig. 2. Correlation analysis between all sporophyte mRNA-seq samples.

Supplementary Fig. 3. Functional categorisation of enriched GO terms in sporophyll and notch reproductive candidate genes.

Supplementary Fig. 4. Distribution of enriched reproductive GO terms across sporophyll reproductive candidate expression categories.

Supplementary Fig. 5. Distribution of enriched gene regulatory GO terms across sporophyll reproductive candidate expression categories/clusters.

Supplementary Fig. 6. *Ceratopteris* embryo development is synchronised through restricted watering.

Supplementary Fig. 7. Correlation analysis between all gametophyte mRNA-seq samples.

Supplementary Fig. 8. Validation of notch reproductive candidate genes as previously-published embryo-related genes.

Supplementary Fig. 9. Distribution of enriched reproductive GO terms across notch reproductive candidate expression categories.

Supplementary Fig. 10. *Arabidopsis* gene orthology to *Ceratopteris* reproductive candidates by expression category.

Supplementary Fig. 11. Comparison of significantly enriched notch-enriched *Arabidopsis* orthologs between ovule and seed tissues.

Supplementary Fig. 12. Reciprocal enrichment testing of *Ceratopteris* notch-enriched orthologs of conserved *Arabidopsis* genes within notch-enriched candidates.

Supplementary Fig. 13. Growth of the *Ceratopteris* sporophyte under tissue culture conditions.

Supplementary Data 1. Sporophyte phenotypic characterisation data and statistical analysis.

Supplementary Data 2. Sporophyll reproductive candidate gene DE analysis.

Supplementary Data 3. GO term enrichment analysis of *Ceratopteris* reproductive candidate genes (sporophyll and notch)

Supplementary Data 4. Gametophyte fertilization characterisation data and statistical analysis.

Supplementary Data 5. Notch reproductive candidate gene DE analysis.

Supplementary Data 6. Pheatmap gene identities and additional statistical analysis

Supplementary Data 7. *Ceratopteris* gene orthology source data.

Supplementary Data 8. *Arabidopsis* gene ortholog enrichment testing source data and statistical analysis.

Supplementary Data 9. Conserved *Ceratopteris* gene ortholog enrichment testing source data and statistical analysis.

Supplementary Data 10. *Ceratopteris* mRNA-seq sample information and QC data

REFERENCES

1. **Hilton J, Bateman RM 2006.** Pteridosperms are the backbone of the seed-plant phylogeny. *Journal of the Torrey Botanical Society* 133: 119-168. DOI: 10.3159/1095-5674(2006)133[119:PATBOS]2.0.CO;2.
2. **Wickett NJ, Mirareb S, Nguyen N, Warnow T, Carpenter E, Matasci N, Ayyampalayam S, Barker MS, Burleigh JG, Gitzendanner MA, Ruhfel BR, Wafula E, Der JP, Graham SW, Mathews S, Melkonian M, Soltis DE, Soltis PS, Miles NW, Rothfels CJ, Pokorny L, Shaw AJ, DeGirononimo L, Stevenson DW, Surek B, Villarreal JC, Roure B, Philippe H, dePamphilis CW, Chen T, Deyholos MK, Baucom RS, Kutchan TM, Augustin MM, Wang J, Zhang Y, Tian Z, Yan Z, Wu X, Sun X, Wong GKS, Leebens-Mack J 2014.** Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Science* 111: E4859-E4868. DOI: 10.1073/pnas.1323926111.
3. **Pryer KM, Schneider H, Smith AR, Cranfill R, Wolf PG, Hunt JS, Sipes SD 2001.** Horestails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature* 409: 618-622. DOI: 10.1038/35054555.
4. **One Thousand Plant Transcriptomes Initiative 2019.** One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574: 679–685. DOI: 10.1038/s41586-019-1693-2.
5. **Rudall PJ 2021.** Evolution and patterning of the ovule in seed plants. *Biological Reviews* 96: 943-960. DOI: 10.1111/brv.12684.

6. **Singh H, Johri BM 1972.** Development of gymnosperm seeds. In: Seed Biology: Importance, development and germination Volume 1(pp 21-75). Editor Kozlowski TT, Elsevier. DOI: 10.1016/B978-0-124-24301-9.X5001-3.
7. **Linkies A, Graeber K, Knight C, Leubner-Metzger G 2010.** The evolution of seeds. New Phytologist 186: 817-831. DOI: 10.1111/j.1469-8137.2010.03249.x.
8. **Matilla AJ 2019.** Seed coat formation: its evolution and regulation. Seed Science Research 29: 215-226. DOI: 10.1017/S0960258519000254.
9. **Qiu YL, Taylor AB, McManus HA 2012.** Evolution of the life cycle in land plants. Journal of Systematics and Evolution 50: 179-194. DOI: 10.1111/j.1759-6831.2012.00188.x.
10. **Conway SJ, Di Stilio VS 2020.** An ontogenetic framework for functional studies in the model fern *Ceratopteris richardii*. Developmental Biology 457: 20-29. DOI: 10.1016/j.ydbio.2019.08.017.
11. **Johnson, GP, Renzaglia, KS 2008.** Embryology of *Ceratopteris richardii* (Pteridaceae, tribe Ceratopterideae), with emphasis on placental development. Journal of Plant Research 121:581-592. DOI: 10.1007/s10265-008-0187-3.
12. **Herr JM 1995.** The origin of the ovule. American Journal of Botany 82: 547-564. DOI: 10.1002/j.1537-2197.1995.tb15676.x
13. **Taylor EL, Taylor TN, Krings M 2009.** Palaeobotany: The biology and evolution of fossil plants. Cambridge MA: Academic Press.
14. **DiMichele WA, Davis JI, Olmstead RG 1989.** Origins of heterospory and the seed habit: the role of heterochrony. Taxon 38: 1-11. DOI: 10.2307/1220881.
15. **Haig D, Westoby H 1989.** Selective forces in the emergence of the seed habit. Biological Journal of the Linnean Society 38: 215-238. DOI: 10.1111/j.1095-8312.1989.tb01576.x.
16. **Bateman RM, DiMichele WA 1994.** Heterospory: the most iterative key innovation in the evolutionary history of the plant kingdom. Biological Reviews 69: 345-417. DOI: 10.1111/j.1469-185X.1994.tb01276.x.
17. **Petersen KB, Burd M 2017.** Why did heterospory evolve? Biological Reviews 92: 1739-1754/ DOI: 10.1111/brv.12304.
18. **Hickok LG, Warne TR, Fribourg RS 1995.** The biology of the fern *Ceratopteris* and its use as a model system. International Journal of Plant Sciences 156: 332-345. DOI: 10.1086/297255.
19. **Ambrose BA, Stevenson DW 2024.** The evolution and development of sporangia- the fundamental reproductive organ of land plant sporophytes. Current Opinion in Plant Biology 81: 102563. DOI: 10.1016/j.pbi.2024.102563.
20. **Friedman WE, Williams JH 2003.** Modularity of the angiosperm female gametophyte and its bearing on the early evolution of endosperm in flowering plants. Evolution 57: 216-230. DOI: 10.1111/j.0014-3820.2003.tb00257.x.

21. **Ranganath RM 2003.** Female gametophyte development in higher plants- meiosis and mitosis break the cellular barrier. *Plant Biology* 5: 42-49. DOI: 10.1055/s-2003-37976.
22. **Favaro R, Pinyopich A, Battaglia R, Kooiker M, Borghi L, Ditta G, Yanofsky MF, Kater MM, Colombo L 2003.** MADS-box protein complexes control carpel and ovule development in *Arabidopsis*. *The Plant Cell* 15: 2603-2611. DOI: 10.1105/tpc.015123.
23. **Pinyopich A, Ditta GS, Savidge B, Liljegren SJ, Baumann E, Wisman E, Yanofsky MF 2003.** Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature* 424: 85-88. DOI: 10.1038/nature01741.
24. **Hasebe M, Wen CK, Kato M, Banks JA 1998.** Characterization of MADS homeotic genes in the fern *Ceratopteris richardii*. *Proceedings of the National Academy of Sciences USA* 95: 6222-6227. DOI: 10.1073/pnas.95.11.6222.
25. **Theissen G, Melzer R, Rümpler F 2016.** MADS-domain transcription factors and the floral quartet model of flower development: linking plant development and evolution. *Development* 143: 3259-3271. DOI: 10.1242/dev.134080.
26. **Xie Z, Li X, Glover BJ, Bai S, Rao GY, Luo J, Yang J 2008.** Duplication and functional diversification of *HAP3* genes leading to the origin of the seed-developmental regulatory gene, *LEAFY COTYLEDON1 (LEC1)*, in nonseed plant genomes. *Molecular Biology and Evolution* 25: 1581-1592. DOI: 10.1093/molbev/msn105.
27. **Rodríguez-Pelayo C, Ambrose B, Vasco A, Alzate JF, Pabón-Mora N 2022.** Evolution and expression of *LEAFY* genes in ferns and lycophytes. *EvoDevo* 13: 2. DOI: 10.1186/s13227-021-00188-9.
28. **Withers KA, Falls K, Youngstrom CE, Nguyen T, DeWald A, Yarvis RM, Simons GP, Flanagan R, Bui LT, Irish EE, Cheng CL 2023.** A *Ceratopteris* *EXCESS MICROSPOROCTES1* suppresses reproductive transition in the fern vegetative leaves. *Plant Science* 353: 111812. DOI: 10.1016/j.plantsci.2023.111812.
29. **Li FW, Brouwer P, Carretero-Paulet L, Cheng S, De Vries J, Delaux PM, Eily A, Koppers N, Kuo LY, Li Z, Simenc M, Small I, Wafula E, Angarita S, Barker MS, Bräutigam A, dePamphilis C, Gould S, Hosmani PS, Huang YM, Huettel B, Kato Y, Liu X, Maere S, McDowell R, Mueller LA, Nierop KGJ, Rensing SA, Robison T, Rothfels CJ, Sigel EM, Song Y, Timilsena PR, Van de Peer Y, Wang H, Wilhelmsson PKI, Wolf PG, Xu X, Der JP, Schluepmann H, Wong GKS, Pryer KM 2018.** Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nature Plants* 4: 460-472. DOI: 10.1038/s41477-018-0188-8.
30. **Fang Y, Qin X, Liao Q, Du R, Luo X, Zhou Q, Li Z, Chen H, Jin W, Yuan Y, Sun P, Zhang R, Zhang J, Wang L, Cheng S, Yang X, Yan Y, Zhang X, Zhang Z, Bai S, Van de Peer Y, Lucas WJ, Huang S, Yan J 2022.** The genome of homosporous maidenhair fern sheds

- light on the euphyllophyte evolution and defences. *Nature Plants* 8: 1024-1037. DOI: 10.1038/s41477-022-01222-x.
31. **Huang X, Wang W, Gong T, Wickell D, Kuo LY, Zhang X, Wen J, Kim H, Lu F, Zhao H, Chen S, Li H, Wu W, Yu C, Chen S, Fan W, Chen S, Bao X, Li L, Zhang D, Jiang L, Khadka D, Yan X, Liao Z, Zhou G, Guo Y, Ralph J, Sederoff RR, Wei H, Zhu P, Li FW, Ming R, Li Q 2022.** The flying spider-monkey tree fern genome provides insights into fern evolution and arborescence. *Nature Plants* 8: 500-512. DOI: 10.1038/s41477-022-01146-6.
32. **Marchant DB, Chen G, Cai S, Chen F, Schafran P, Jenkins J, Shu S, Plott C, Webber J, Lovell TJ, He G, Sandor L, Williams M, Rajasekar S, Healey A, Barry K, Zhang Y, Sessa E, Dhakal RR, Wolf PG, Harkess A, Li FW, Rössner C, Becker A, Gramzow L, Xue D, Wu Y, Tong T, Wang Y, Dai F, Hua S, Wang H, Xu S, Xu F, Duan H, Theissen G, McKain MR, Li Z, McKibben MTW, Barker MS, Schmitz RJ, Stevenson DW, Zumajo-Cardona C, Ambrose BA, Leebens-Mack JH, Grimwood J, Schmutz J, Soltis PS, Soltis DE, Chen ZH 2022.** Dynamic genome evolution in a model fern. *Nature Plants* 8: 1038-1051. DOI: 10.1038/s41477-022-01226-7.
33. **Fernández P, Amice R, Bruy D, Christenhusz MJM, Leitch IJ, Leitch AL, Pokorny L, Hidalgo O, Pellicer J 2024.** A 160 Gbp fork fern genome shatters size record for eukaryotes. *iScience* 27: 109889. DOI: 10.1016/j.isci.2024.109889.
34. **Aya K, Kobayashi M, Tanaka J, Ohyanagi H, Suzuki T, Yano K, Takano T, Yano K, Matsuoka M 2015.** De novo transcriptome assembly of a fern, *Lygodium japonicum*, and a web resource database, LjtransDB. *Plant and Cell Physiology* 56: e5. DOI: 10.1093/pcp/pcu184.
35. **Geng Y, Cai C, McAdam SAM, Banks JA, Wisecaver JH, Zhou Y 2021.** A de novo transcriptome assembly of *Ceratopteris richardii* provides insights into the evolutionary dynamics of complex gene families in land plants. *Genome Biology and Evolution* 13: evab042. DOI: 10.1093/gbe/evab042.
36. **Ali ZM, Tan QW, Lim PK, Chen H, Pfeifer L, Julca I, Lee JM, Classen B, de Vries S, de Vries J, Koh TS, Chin LL, Vinter F, Alvarado C, Layens A, Mizrachi E, Motawie MS, Joergensen B, Ulvskov P, Van de Peer Y, Ho BC, Sibout R, Mutwil.** Comparative transcriptomics in ferns reveals key innovations and divergent evolution of the secondary cell walls. Preprint at <https://www.biorxiv.org/content/10.1101/2024.08.27.609851v1> (2024).
37. **Fu Q, Chen LQ 2019.** Comparative transcriptome analysis of two reproductive modes in *Adiantum reniforme* var. *sinense* targeted to explore possible mechanism of apogamy. *BMC Genetics* 20: 55. DOI: 10.1186/s12863-019-0762-8.

38. **Qian Q, Yang Y, Zhang W, Hu Y, Li Y, Yu H, Hou X 2021.** A novel *Arabidopsis* gene RGAT1 is required for GA-mediated tapetum and pollen development. *New Phytologist* 231: 137-151. DOI: 10.1111/nph.17314.
39. **Hedden P, Thomas SG 2012.** Gibberellin biosynthesis and its regulation. *Biochemical Journal* 444: 11-25. DOI: 10.1042/BJ20120245.
40. **Couteau F, Belzile F, Horlow C, Grandjean O, Vezon D, Doutriaux MP 1999.** Random chromosome segregation without meiotic arrest in both male and female meiocytes of a *dmc1* mutant of *Arabidopsis*. *The Plant Cell* 11: 1623-1634. DOI: 10.1105/tpc.11.9.1623.
41. **Higgins JD, Armstrong SJ, Franklin FCH, Jones GH 2004.** The *Arabidopsis* MutS homolog AtMSH4 functions at an early step in recombination: evidence for two classes of recombination in *Arabidopsis*. *Genes and Development* 18: 2557-2570. DOI: 10.1101/gad.317504.
42. **Chelysheva L, Gendrot G, Vezon D, Doutriaux MP, Mercier R, Grelon M 2007.** Zip4/Spo22 is required for Class I CO formation but not for synapsis completion in *Arabidopsis thaliana*. *PLoS Genetics* 3: e83. DOI: 10.1371/journal.pgen.0030083.t002.
43. **Lu X, Liu X, An L, Zhang W, Sun J, Pei H, Meng H, Fan Y, Zhang C 2008.** The *Arabidopsis* MutS homolog AtMSH5 is required for normal meiosis. *Cell Research* 18: 589-599. DOI: 10.1038/cr.2008.44.
44. **Zamariola L, De Storme N, Tiang CL, Armstrong SJ, Franklin FCH, Geelan D 2013.** SGO1 but not SGO2 is required for maintenance of centromere cohesion in *Arabidopsis thaliana* meiosis. *Plant Reproduction* 26: 197-208. DOI: 10.1007/s00497-013-0231-x.
45. **Fujita H, Takemura M, Tani E, Nemoto K, Yokota A, Kohchi T 2003.** An *Arabidopsis* MADS-box protein, AGL24, is specifically bound to and phosphorylated by Meristematic Receptor-Like Kinase (MRLK). *Plant and Cell Physiology* 44: 735-742. DOI: 10.1093/pcp/pcg092.
46. **Gregis V, Sessa A, Dorca-Fornell C, Kater MM 2009.** The *Arabidopsis* floral meristem identity genes AP1, AGL24 and SVP directly repress class B and C floral homeotic genes. *The Plant Journal* 60: 626-637. DOI: 10.1111/j.1365-313X.2009.03985.x.
47. **Liu C, Xi W, Shen L, Tan C, Yu H 2009.** Regulation of floral patterning by flowering time genes. *Developmental Cell* 16: 711-722. DOI: 10.1016/j.devcel.2009.03.011.
48. **Canales C, Bhatt AM, Scott R, Dickinson H 2002.** EXS, a putative LRR receptor kinase, regulates male germline cell number and tapetal identity and promotes seed development in *Arabidopsis*. *Current Biology* 12: 1718-1727. DOI: 10.1016/S0960-9822(02)01151-X.
49. **Zhao DZ, Wang GF, Speal B, Ma H 2002.** The *EXCESS MICROSPOROCTES1* gene encodes a putative leucine-rich repeat receptor protein kinase that controls somatic and reproductive cell fates in the *Arabidopsis* anther. *Genes and Development* 16: 2021-2031. DOI: 10.1101/gad.997902.

- 1043 50. **Banks J 1999.** Gametophyte development in ferns. Annual Review of Plant Biology 50:
1044 163-186. DOI: 10.1146/annurev.arplant.50.1.163.
- 1045 51. **Geng Y, Yan A, Zhou Y 2022.** Positional cues and cell division dynamics drive meristem
1046 development and archegonium formation in *Ceratopteris* gametophytes. Communications
1047 Biology 5: 650. DOI: 10.1038/s42003-022-03627-y.
- 1048 52. **Li W, Chen C, Markmann-Mulisch U, Timofejeva L, Schmelzer E, Ma H, Reiss B 2004.**
1049 The *Arabidopsis AtRAD51* gene is dispensable for vegetative development but required
1050 for meiosis. Proceedings of the National Academy of Science USA 101: 10596-10601.
1051 DOI: 10.1073/pnas.0404110101.
- 1052 53. **Siaud N, Dray E, Gy I, Gérard E, Takvorian N, Doutriaux MP 2004.** Brca2 is involved in
1053 meiosis in *Arabidopsis thaliana* as suggested by its interaction with Dmc1. The EMBO
1054 Journal 23:1392-1401. DOI: 10.1038/sj.emboj.7600146.
- 1055 54. **Bui LT, Pandzic D, Youngstrom CE, Wallace S, Irish EE, Szövényi P, Cheng CL 2017.**
1056 A fern AINTEGUMENTA gene mirrors BABY BOOM in promoting apogamy in *Ceratopteris*
1057 *richardii*. The Plant Journal 90: 122-132. DOI: 10.1111/tpj.13479.
- 1058 55. **Plackett ARG, Conway SJ, Hewett Hazelton KD, Rabbinoiwitsch EH, Langdale JA, Di**
1059 **Stilio VS 2018.** LEAFY maintains apical stem cell activity during shoot development in the
1060 fern *Ceratopteris richardii*. eLife 7: e39625. DOI: 10.7554/eLife.39625.
- 1061 56. **Clark SE, Running MP, Meyerowitz EM 1993.** *CLAVATA1*, a regulator of meristem and
1062 flower development in *Arabidopsis*. Development 119:397-418. DOI:
1063 10.1242/dev.119.2.397.
- 1064 57. **Greb T, Clarenz O, Schafer E, Muller D, Herrero R, Schmitz G, Theres K 2003.**
1065 Molecular analysis of the *LATERAL SUPPRESSOR* gene in *Arabidopsis* reveals a
1066 conserved control mechanism for axillary meristem formation. Genes and Development
1067 17: 1175-1187. DOI: 10.1101/gad.260703.
- 1068 58. **Cui Y, Bi YM, Brugière N, Arnoldo M, Rothstein SJ 2000.** The S locus glycoprotein and
1069 the S-receptor kinase are sufficient for self-pollen rejection in Brassica. Proceedings of the
1070 National Academy of Science USA 97: 3713-3717. DOI: 10.1073ypnas.050480297.
- 1071 59. **Vivian-Smith A, Kultonow AM 1999.** Genetic analysis of growth-regulator-induced
1072 parthenocarpy in *Arabidopsis*. Plant Physiology 121: 437-451. DOI: 10.1104/pp.121.2.437.
- 1073 60. **Goetz M, Hooper LC, Johnson SD, Rodrigues JCM, Vivian-Smith A, Kultonow AM**
1074 **2007.** Expression of aberrant forms of *AUXIN RESPONSE FACTOR8* stimulates
1075 parthenocarpy in *Arabidopsis* and tomato. Plant Physiology 145: 351-366. DOI:
1076 10.1104/pp.107.104174.
- 1077 61. **Figueiredo DD, Batista RA, Roszak PJ, Köhler C 2015.** Auxin production couples
1078 endosperm development to fertilization. Nature Plants 1: 15184. DOI:
1079 10.1038/nplants.2015.184.

62. **Liu YH, Offler CE, Ruan YL 2013.** Regulation of fruit and seed response to heat and drought by sugars as nutrients and signals. *Frontiers in Plant Science* 4: 282. DOI: 10.3389/fpls.2013.00282.
63. **Griffiths CA, Sagar R, Geng Y, Primavesi LF, Patel MK, Passarelli MK, Gilmore IS, Steven RT, Bunch J, Paul MJ, Davis BG 2016.** Chemical intervention in plant sugar signalling increases yield and resilience. *Nature* 540: 574-578. DOI: 10.1038/nature20591.
64. **Cordle AR, Irish EE, Cheng CL 2007.** Apogamy induction in *Ceratopteris richardii*. *International Journal of Plant Sciences* 168: 361-369. DOI: 10.1086/511049.
65. **Honys D, Twell D 2003.** Comparative analysis of the Arabidopsis pollen transcriptome. *Plant Physiology* 132: 640-652. DOI: 10.1104/pp.103.020925.
66. **Honys D, Twell D 2004.** Transcriptome analysis of haploid male gametophyte development in *Arabidopsis*. *Genome Biology* 5: R85. DOI: 10.1186/gb-2004-5-11-r85.
67. **Wellmer F, Reichmann JL, Alves-Ferreira M, Meyerowitz EM 2004.** Genome-wide analysis of spatial gene expression in Arabidopsis flowers. *The Plant Cell* 16: 1314-1326. DOI: 10.1105/tpc.021741.
68. **Yu HJ, Hogan P, Sundaresan V 2005.** Analysis of the female gametophyte transcriptome of Arabidopsis by comparative expression profiling. *Plant Physiology* 139: 1853-1869. DOI: 10.1104/pp.105.067314.
69. **Alves-Ferreira M, Wellmer F, Banhara A, Kumar V, Reichmann JL, Meyerowitz EM 2007.** Global expression profiling applied to the analysis of Arabidopsis stamen development. *Plant Physiology* 145: 747-762. DOI: 10.1104/pp.107.104422.
70. **Johnston AJ, Meier P, Gheyselinck J, Wuest SEJ, Federer M, Schlagenhauf E, Becker JD, U Grossniklaus 2007.** Genetic subtraction profiling identifies genes essential for *Arabidopsis* reproduction and reveals interaction between the female gametophyte and the maternal sporophyte. *Genome Biology* 8: R204. DOI: 10.1186/gb-2007-8-10-r204.
71. **Peiffer JA, Kaushik S, Sakai H, Arteaga-Vasquez M, Sanchez-Leon N, Ghazal H, Vielle-Calzada JP, Meyers BC 2008.** A spatial dissection of the Arabidopsis floral transcriptome by MPSS. *BMC Plant Biology* 8: 43. DOI:10.1186/1471-2229-8-43.
72. **Skinner DJ, Gasser CS 2009.** Expression-based discovery of candidate ovule development regulators through transcriptional profiling of ovule mutants. *BMC Plant Biology* 9: 29. DOI: 10.1186/1471-2229-9-29.
73. **Le BH, Cheng C, Bui AQ, Wagmaister JA, Henry KF, Pelletier J, Kwong L, Belmonte M, Kirkbride R, Horvath S, Drews GN, Fischer RL, Okamuro JK, Harada JJ, Goldberg RB 2010.** Global analysis of gene activity during *Arabidopsis* seed development and identification of seed-specific transcription factors. *Proceedings of the National Academy of Sciences* 107: 8063-8070. DOI: 10.1073/pnas.1003530107.

74. **Wuest SE, Vijverberg K, Schmidt A, Weiss M, Gheyselinck J, Lohr M, Wellmer F, Rahnenführer J, von Mering C, Grossniklaus U 2010.** *Arabidopsis* female gametophyte gene expression map reveals similarities between plant and animal gametes. *Current Biology* 20: 506-512. DOI: 10.1016/j.cub.2010.01.051.
75. **Schmidt A, Wuest SE, Vijverberg K, Baroux C, Kleen D, Grossniklaus U 2011.** Transcriptome analysis of the *Arabidopsis* Megaspore Mother Cell uncovers the importance of RNA helicases for plant germline development. *PLoS Biology* 9: e1001155. DOI: 10.1371/journal.pbio.1001155.
76. **Xiang D, Venglat P, Tibiche C, Yang H, Risseuw E, Cao Y, Babic V, Cloutier M, Keller W, Wang E, Selvaraj G, Datla R 2011.** Genome-wide analysis reveals gene expression and metabolic network dynamics during embryo development in *Arabidopsis*. *Plant Physiology* 156: 346-356. DOI: 10.1104/pp.110.171702.
77. **Belmonte MF, Kirkbride RC, Stone SL, Pelletier JM, Bui AQ, Yeung EC, Hashimoto M, Fei J, Harada CM, Munoz MD, Le BH, Drews GN, Brady SM, Goldberg RB, Harada JJ 2013.** Comprehensive developmental profiles of gene activity in regions and subregions of the *Arabidopsis* seed. *Proceedings of the National Academy of Sciences* 110: E435-E444. DOI: 10.1073/pnas.1222061110.
78. **Schiefthaler U, Balasubramanian S, Sieber P, Chevalier D, Wisman E, Schneitz K 1999.** Molecular analysis of *NOZZLE*, a gene involved in pattern formation and early sporogenesis during sex organ development in *Arabidopsis thaliana*. *Proceedings of the National Academy of Science USA* 96: 11664-11669. DOI: 10.1073/pnas.96.20.11664.
79. **Yang YC, Ye D, Xu J, Sundaresan V 1999.** The *SPOROCTELESS* gene of *Arabidopsis* is required for initiation of sporogenesis and encodes a novel nuclear protein. *Genes and Development* 13: 2108-2117. DOI: 10.1101/gad.13.16.2108.
80. **Bao S, Hua C, Shen L, Yu H. 2019.** New insights into gibberellin signaling in regulating flowering in *Arabidopsis*. *Journal of Integrative Plant Biology* 62: 118-131. DOI: 10.1111/jipb.12892.
81. **Ferrándiz C, Gu Q, Martienssen R, Yanofsky MF 2000.** Redundant regulation of meristem identity and plant architecture by *FRUITFULL*, *APETALA1* and *CAULIFLOWER*. *Development* 127: 725-734. DOI: 10.1242/dev.127.4.725.
82. **Becker A, Theissen G 2003.** The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Molecular Phylogenetics and Evolution* 29: 464-489. DOI: 10.1016/S1055-7903(03)00207-0.
83. **Seeliger K, Dukowic-Schulze S, Wurzel-Wildersinn R, Pacher M, Puchta H 2012.** BRCA2 is a mediator of RAD51- and DMC1-facilitated homologous recombination in *Arabidopsis thaliana*. *New Phytologist* 193: 364-375. DOI: 10.1111/j.1469-8137.2011.03947.x.

84. **Kitajima TS, Kawashima SA, Watanabe Y 2004.** The conserved kinetochore protein shugoshin protects centromeric cohesion during meiosis. *Nature* 427: 510-517. DOI: 10.1038/nature02312.
85. **Tsubouchi T, Zhao H, Roeder GS 2006.** The meiosis-specific Zip4 protein regulates crossover distribution by promoting synaptonemal complex formation together with Zip2. *Developmental Cell* 10: 809-819. DOI: 10.1016/j.devcel.2006.04.003.
86. **Trapp O, Seeliger K, Puchta H 2011.** Homologs of breast cancer genes in plants. *Frontiers in Plant Science* 2: 19. DOI: 10.3389/fpls.2011.00019.
87. **Grishaeva TM, Bogdanov YF 2018.** Conservation of meiosis-specific nuclear proteins in eukaryotes: comparative approach. *Nucleus* 61: 175-182. DOI: 10.1007/s13237-018-0253-8.
88. **Abe K, Osakabe K, Nakayama S, Endo M, Tagiri A, Todoriki S, Ichikawa H, Toki S 2005.** Arabidopsis *RAD51C* gene is important for homologous recombination in meiosis and mitosis. *Plant Physiology* 139: 896-908. DOI: 10.1104/pp.105.065243.
89. **Zhao L, He J, Cai H, Lin H, Li Y, Liu R, Yang Z, Qin Y 2014.** Comparative expression profiling reveals gene functions in female meiosis and gametophyte development in Arabidopsis. *The Plant Journal* 80: 615-628. DOI: 10.1111/tpj.12657.
90. **Mendes MA, Petrella R, Cucinotta M, Vignati E, Gatti S, Pinto SC, Bird DC, Gregis V, Dickinson H, Tucker MR, Colombo L 2020.** The RNA-dependent DNA methylation pathway is required to restrict *SPOROCTELESS/NOZZLE* expression to specify a single female germ cell precursor in *Arabidopsis*. *Development* 147: dev194274. DOI: 10.1242/dev.194274.
91. **Okano Y, Aono N, Hiwatashi Y, Murata T, Nishiyama T, Ishikawa T, Kubo M, Hasebe M 2009.** A polycomb repressive complex 2 gene regulates apogamy and gives evolutionary insights into early land plant evolution. *Proceedings of the National Academy of Sciences USA* 106: 16321-16326. DOI: 10.1073/pnas.0906997106.
92. **Warne TR, Hickok LG 1987.** (2-Chloroethyl)phosphonic Acid Promotes Germination of Immature Spores of *Ceratopteris richardii* Brongn. *Plant Physiology* 83: 723-725. DOI: 10.1104/pp.83.4.723.
93. **Plackett ARG, Rabbinowitsch EH, Langdale JA 2015.** Protocol: Genetic transformation of the fern *Ceratopteris richardii* through microparticle bombardment. *Plant Methods* 11: 37. DOI: 10.1186/s13007-015-0080-8.
94. **Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K. Tomancak P. 2012.** Fiji: an open-source platform for biological-image analysis. *Nature Methods* 9: 676-682. DOI: 10.1038/nmeth.2019.

95. **R Core Team (2021)**. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
96. **Fox J, Weisberg S. 2019**. *An R Companion to Applied Regression*, Third edition (Sage).
97. **Shapiro SS, Wilk MB 1965**. An analysis of variance test for normality (complete samples). *Biometrika* 52: 591–611. DOI: 10.1093/biomet/52.3-4.591.
98. **Levene H 1960**. Robust tests for equality of variances. In *Contributions to Probability and Statistics*, I. Olkin, ed. (Stanford University Press), pp. 278–292.
99. **Benjamini Y, Hochberg Y 1995**. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)* 57: 289–300. DOI: 10.1111/j.2517-6161.1995.tb02031.x.
100. **Kim D, Paggi JM, Park C, Bennett C, Salzberg SL 2019**. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology* 37, 907-915. DOI: 10.1038/s41587-019-0201-4.
101. **Emms DM, Kelly S 2019**. Orthofinder: phylogenetic orthology inference for comparative genomics. *Genome Biology* 20: 238. DOI: 10.1186/s13059-019-1832-y.
102. **Bernardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, Huala E 2015**. The Arabidopsis Information Resource: making and mining the “gold standard” annotated reference plant genome. *Genesis* 53: 474-485. DOI: 10.1002/dvg.22877.
103. **Chamala S, Chanderbali AS, Der JP, Lan T, Walts B, Albert VA, dePamphilis CW, Leebens-Mack J, Rounsley S, Schuster SC, Wing RA, Xiao N, Moore R, Soltis PS, Soltis DE, Barbazuk WB 2013**. Assembly and validation of the genome of the nonmodel basal angiosperm *Amborella*. *Science* 342: 1516-1517. DOI: 10.1126/science.1241130.
104. **Schmutz J., Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelan JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang XC, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA 2010**. Genome sequence of the palaeopolyploid soybean. *Nature* 463: 178-183. DOI: 10.1038/nature08670.
105. **Worden AZ, Lee JH, Mock T, Rouzé P, Simmons MP, Aerts AL, Allen AE, Cuvelier ML, Derelle E, Everett MV et al. 2009**. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science* 324, 268-272. DOI: 10.1126/science.1167222.
106. **Palenik B, Grimwood J, Aerts A, Rouzé P, Salamov A, Putnam N, Dupont C, Jorgensen R, Derelle E, Rombauts S et al. 2007**. The tiny eukaryote *Ostreococcus*

- provides genomic insights into the paradox of plankton speciation. Proceedings of the National Academy of Science USA 104: 7705-7710. DOI: 10.1073/pnas.0611046104.
107. **Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y, Zheng L et al. 2007.** The TIGR Rice Genome Annotation Resource: improvements and new features., Nucleic Acids Research 35: D883-D887. DOI: 10.1093/nar/gkl976.
108. **Lang D, Ullrich KK, Murat F, Fuchs J, Jenkins J, Haas FB, Piednoel M, Gundlach H, Van Bel M, Meyberg R et al. 2018.** The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. The Plant Journal 93: 515-533. DOI: 10.1111/tpj.13801.
109. **Plomion C, Aury JM, Amselem J, Leroy T, Murat F, Duplessis S, Faye S, Francillonne N, Labadie K, Le Provost G et al. 2018.** Oak genome reveals facets of long lifespan. Nature Plants 4: 440-452. DOI: 10.1038/s41477-018-0172-3.
110. **Hosmani P S, Flores-Gonzalez M, van de Geest H, Maumus F, Bakker LV, Schijlen E, van Haarst J, Cordewener J, Sanchez-Perez G, Peters S, Fei Z, Giovannoni JJ, Mueller LA, Saha S. 2019.** An improved de novo assembly and annotation of the tomato reference genome using single-molecule sequencing, Hi-C proximity ligation and optical maps. Preprint at <https://www.biorxiv.org/content/10.1101/767764v1> (2019).
111. **Banks JA, Nishiyama T, Hasebe M, Bowman JL, Gribskov M, dePamphilis C, Albert VA, Aono N, Aoyama T, Ambrose BA et al. 2011.** The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. Science 332: 960-963. DOI: 10.1126/science.1203810.
112. **Wang W, Haberer G, Gundlach H, Gläßer C, Nussbaumer T, Luo MC, Lomsadze A, Borodovsky M, Kerstetter RA, Shanklin J et al. 2014.** The *Spirodela polyrhiza* genome reveals insights into its neotenus reduction fast growth and aquatic lifestyle. Nature Communications 5: 3311. DOI: 10.1038/ncomms4311.
113. **Prochnik SE, Umen J, Nedelcu AM, Hallmann A, Miller SM, Nishii I, Ferris P, Kuo A, Mitros T, Fritz-Laylin LK et al. 2010.** Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*. Science 329: 223-226. DOI: 10.1126/science.1188800.
114. **Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R 2015.** InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics 16:169. DOI: 10.1186/s12859-015-0611-3.
115. **Kolde R 2018.** pheatmap: Pretty Heatmaps. R package version 1.0.12, <https://github.com/raivokolde/pheatmap>.
116. **Alexa A, Rahnenfuhrer J 2024.** topGO: Enrichment Analysis for Gene Ontology. R package version 2.58.0. <https://github.com/ycl6/topGO-feat>.

117. **Wickham H 2016.** ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, <https://ggplot2.tidyverse.org>.
118. **Yan L 2023.** ggvenn: Draw Venn Diagram by 'ggplot2'. R package version 0.1.10, <https://CRAN.R-project.org/package=ggvenn>.
119. **Kassambara A 2023.** ggcorrplot: Visualization of a Correlation Matrix using 'ggplot2'. R package version 0.1.4.1, <https://CRAN.R-project.org/package=ggcorrplot>.
120. **Ray A, Robinson-Beers K, Ray S, Baker SC, Lang JD, Preuss D, Milligan SB, Gasser CS 1994.** *Arabidopsis* floral homeotic gene BELL (*BEL1*) controls ovule development through negative regulation of AGAMOUS gene (AG). Proceedings of the National Academy of Sciences USA 91: 5761-5765. DOI: 10.1073/pnas.91.13.5761.
121. **Reiser L, Modrusan Z, Margossian L, Samach A, Ohad N, Haughn GW, Fisher RL 1995.** The *BELL1* gene encodes a homeodomain protein involved in pattern formation in the *Arabidopsis* ovule primordium. Cell 83: 735-742. DOI: 10.1016/0092-8674(95)90186-8.
122. **Wei SJ, Chai S, Zhu RM, Duan CY, Zhang Y, Li S 2020.** HUA ENHANCER1 mediates ovule development. Frontiers in Plant Science 11: 397. DOI: 10.3389/fpls.2020.00397.
123. **Golden TA, Schauer SE, Lang JD, Pien S, Mushegian AR, Grossniklaus U, Meinke DW, Ray A 2002.** *Short Integuments1/suspensor1/carpel Factory*, a Dicer Homolog, Is a Maternal Effect Gene Required for Embryo Development in *Arabidopsis*. Plant Physiology 130: 808-822. DOI: 10.1104/pp.003491.
124. **Robinson-Beers K, Pruitt RE, Gasser CS 1992.** Ovule development in wild-type *Arabidopsis* and two female-sterile mutants. The Plant Cell 4: 1237-1249. DOI: 10.1105/tpc.4.10.1237.
125. **Pagnussat GC, Yu HJ, Ngo QA, Rajani S, Mayalagu S, Johnson CS, Capron A, Xie LF, Ye D, Sundaresan V 2005.** Genetic and molecular identification of genes required for female gametophyte development and function in *Arabidopsis*. Development 132: 603-614. DOI: 10.1242/dev.01595.
126. **Oh SA, Johnson A, Smertenko A, Rahman D, Park SK, Hussey PJ, Twell D 2005.** A divergent cellular role for the FUSED kinase family in the plant-specific cytokinetic phragmoplast. Current Biology 15: 2107-2111. DOI: 10.1016/j.cub.2005.10.044.
127. **Schröder JA, Bonnet DMV, Jullien PE 2023.** Non-autonomous small RNA silencing in *Arabidopsis* female gametes. Current Biology 33: 183-188. DOI: 10.1016/j.cub.2022.11.044.
128. **Gross-Hardt R, Kägi C, Baumann N, Moore JM, Baskar R, Gagliano WB, Jürgens G, Grossniklaus U 2007.** *LACHESIS* restricts gametic cell fate in the female gametophytes of *Arabidopsis*. PLoS Biology 5: e47. DOI: 10.1371/journal.pbio.0050047.

129. **Griffith ME, Mayer U, Capron A, Ngo QA, Surendrarao A, McClinton R, Jürgens G, Sundaresan V 2007.** The *TORMOZ* gene encodes a nuclear protein required for regulated division planes and embryo development in *Arabidopsis*. *The Plant Cell* 19: 2246-2263. DOI: 10.1105/tpc.106.042697.
130. **Autran D, Baroux C, Raissig MT, Lenormand T, Wittig M, Grob S, Steimer A, Barann M, Klostermeier UC, Leblanc O, Vielle-Calzada JP, Rosenstiel P, Grimanelli D, Grossniklaus U 2011.** Maternal epigenetic pathways control parental contributions to *Arabidopsis* early embryogenesis. *Cell* 145: 707-719. DOI: 10.1016/j.cell.2011.04.014
131. **Chen B, Maas L, Figueiredo D, Zhong Y, Reis R, Li M, Horstman A, Riksen T, Weeman M, Liu H, Siemons C, Chen S, Angenent GC, Boutilier K 2022.** BABY BOOM regulates early embryo and endosperm development. *Proceedings of the National Academy of Science USA* 119, e2201761119. DOI: doi.org/10.1073/pnas.2201761119.
132. **Tzafrir I, McElver JA, Liu C, Wu JQ, Martinez A, Patton DA, Meinke DW 2002.** Diversity of TITAN functions in *Arabidopsis* seed development. *Plant Physiology* 128: 38-51. DOI: 10.1104/pp.010911.