Convergent recruitment of TALE homeodomain life cycle regulators to direct sporophyte development in land plants and brown algae

Alok Arun et al
Figure 1. The oro life cycle mutation corresponds to a TALE homeodomain transcription factor gene. (A) Life cycles of wild type and oro mutant Ectocarpus. The wild type sexual cycle (upper panel) involves production of meio-spores by the diploid sporophyte via meiosis in unilocular (single-chambered) sporangia (US). The meio-spores develop as haploid, dioicous (male and female) gametophytes. The gametophytes produce gametes in plurilocular (multichambered) gametangia (PG), which fuse to produce a diploid sporophyte. Gametes that fail to fuse can develop parthenogenetically to produce a partheno-sporophyte, which can produce spores by apomeiosis or following endoreduplication to engender a new generation of gametophytes. PS, plurilocular sporangium (asexual reproduction). Gametes of the oro mutant (lower panel) are unable to initiate the sporophyte program and develop parthenogenetically to produce partheno-gametophytes. The mutation is recessive so a cross with a wild type gametophyte produces a heterozygous diploid sporophyte with a wild type phenotype. (B) Young gamete-derived parthenotes of wild type and oro strains. Arrowheads indicate round, thick-walled cells typical of the sporophyte for the wild type and long, wavy cells typical of the gametophyte for the oro mutant. Scale bars: 20 μm. (C) Representation of the interval on chromosome 14 between the closest recombining markers to the ORO locus (M_133_107 and M_133) showing the position of the single mutation within the mapped interval. DOI: https://doi.org/10.7554/eLife.43101.003
Figure 2. Phenotypic and genetic characterisation of sam life cycle mutants. (A-C) The sam-1 mutant exhibits gametophyte-like morphological characteristics. Different stages of (A) wild type gametophyte (strain Ec32), (B) wild type partheno-sporophyte (strain Ec32) and (C) sam-1 mutant (strain Ec374). PG, plurilocular gametangia; PS, plurilocular sporangium; US, unilocular sporangium. (D) sam mutants exhibit a gametophyte-like photopolarisation response to unidirectional light. Letters above the boxplot indicate significant differences (Wilcoxon test, p-value<0.01). n, number of replicates; i, number of individuals scored. (E) Representative images of congo red staining showing that the sam-1 mutant protoplasts are resistant to treatment with sporophyte conditioned medium (SCM). GCM, control gametophyte conditioned medium. (F) Abortion of unilocular sporangia in sam-1, sam-2 or sam-3 mutant sporophytes. Images are representative of n = 19 (Ec17), n = 23 (Ec768), n = 20 (Ec833) and n = 14 (Ec361) unilocular sporangia. Figure 2 continued on next page.
IUS, immature unilocular sporangium; MUS, mature unilocular sporangium. (G) Locations of the three sam mutations within the SAM gene. Scale bars: 20 μm (or 50 μm if indicated by 50).
DOI: https://doi.org/10.7554/eLife.43101.004
Figure 2—figure supplement 1. Morphological characteristics of sam mutants. (A–J) The sam-2 and sam-3 mutants exhibit gametophyte-like morphological characteristics. (A–E) sam-2 mutant (strain Ec364); F–J, sam-3
Figure 2—figure supplement 1 continued

mutant (strain Ec793); A-D and F-I, different stages of early development from germination to young, branched germling. E and J, plurilocular gametangia. Size bars indicate 20 μm for all panels except C, D and I where the size bar indicate 50 μm. See Figure 2 for the equivalent developmental stages of wild type sporophytes and gametophytes.

DOI: https://doi.org/10.7554/eLife.43101.005
Figure 3. Gene expression analysis. (A) Abundance of ORO and SAM transcripts during different stages of the life cycle. Error bars, standard error of the mean (SEM); TPM, transcripts per million. (B) Quantitative reverse transcription PCR analysis of generation marker genes. The graphs indicate mean transcript abundance relative to wild-type (wt) samples.

**Figure 3 continued on next page**

Arun et al. eLife 2019;8:e43101. DOI: https://doi.org/10.7554/eLife.43101
values ± standard error of transcript abundances for two gametophyte marker genes, Ec-23_004240 and Ec-21_006530, and two sporophyte marker genes, Ec-20_001150 and Ec-26_000310. Data from five independent experiments. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001. (C) Word cloud representations of the relative abundances (log2 gene number) of manually assigned functional categories in the set of genes that were differential regulated between the sporophyte and gametophyte generations (upper panel) and in the subset of those genes that encode secreted proteins (lower panel). Asterisks indicate functional categories that were significantly over- or under-represented in the two datasets. (D-E) Significantly overrepresented GO terms (D) and KEGG pathways (E) associated with generation-biased genes. (F) Expression patterns of the 200 most strongly generation-biased genes. *oro, oro mutant; sam, sam mutant; imm, immediate upright mutant; GA, gametophyte; SP, sporophyte.

DOI: https://doi.org/10.7554/eLife.43101.006
Evidence for the production of full-length ORO and SAM transcripts during the gametophyte generation. Immature and mature male and female gametophyte Illumina RNA-seq data was mapped onto the ORO and SAM gene sequences using Tophat2. Blue boxes, ORO and SAM coding exons; orange, RNA-seq reads; purple, gaps introduce during mapping corresponding to introns.

DOI: https://doi.org/10.7554/eLife.43101.007
**Figure 4.** Detection of ORO-SAM heterodimerisation in vitro using a pull-down assay. (A) ORO and SAM constructs used for the pull-down experiments. (B) Pull-down assay between SAM and different versions of the ORO protein. (C) Pull-down assay between different versions of the SAM protein and full-length ORO protein. Note that all ORO proteins were fused with the HA epitope. FL, full-length; HD, homeodomain.

DOI: https://doi.org/10.7554/eLife.43101.008
Figure 5. ORO and SAM conservation and domain structure. (A) Unrooted maximum likelihood tree of ORO, SAM and Ec-04_000450 orthologues from diverse brown algal species and the raphidophyte Heterosigma akashiwo. (B) Domain structure of the ORO and SAM TALE homeodomain transcription.

Figure 5 continued on next page.
factors. Conservation: strong (blue), less strong (orange), secondary structure: α-helix (green), β-strand (red). Q1-4, A1 and G1: regions rich in glutamine, alanine and glycine, respectively. (C) Conserved domains in ORO and SAM proteins. Cok, Cladosiphon okamuranus; Csi, Colpomenia sinuosa; Dvi, Desmarestia viridis; Dun, Dictyopteris undulata; Esp, Ectocarpus sp.; Hea, Heterosigma akashiwo; Hfu, Hizikia fusiformis; Iok, Ishige okamurai; Kcr, Kjellmaniella crassifolia; Pfa, Petalonia fascia; Pla, Punctaria latifolia; Sja, Saccharina japonica; Smu, Sargassum muticum; Sva, Sargassum vachellianum; Sdo, Scytosiphon dotyi; Slo, Scytosiphon lomentaria; Upi, Undaria pinnatifida.
DOI: https://doi.org/10.7554/eLife.43101.009
Figure 5—figure supplement 1. Intron conservation in homeobox genes. (A) Conservation of introns in Ectocarpus (Ec), C. okamuranus (Co) and S. japonica (Sj) ORO and SAM genes. Schematic representation of the coding regions of ORO and SAM genes showing the positions and phase of introns. Conserved intron positions, based on sequence similarity, are indicated by grey lines. Introns in similar positions but not linked by a grey line are also likely to be ancestral but it is not possible to verify homology because these regions of the proteins are too diverged. Protein identifiers are Ec-ORO, Ec-14_005920; Co-ORO, Cok_S_s017_4976.t2; Sj-ORO, SJ07622; Ec-SAM, Ec-27_006660; Co-SAM, Cok_S_s018_5094mod; Sj-SAM, SJ10977mod where the suffix ‘mod’ indicates that the original gene model has been modified (see Supplementary file 6). (B) Positions of homeobox introns in stramenopile homeobox genes, life cycle regulators from the green lineage, fungal mating type regulators and selected metazoan homeobox genes. Introns positions are colour coded according to phase: 0, red; 1, blue; 2, orange. The numbering at the bottom indicate the spacing at which the introns are located.
Figure 5—figure supplement 1 continued

conserved 60 residues of the homeodomain and xxx indicates the three additional amino acids in TALE HD TFs. Numbers in brackets indicate the number of introns in the entire coding region of each gene. The asterisk indicates a stop codon. Esp, Ectocarpus sp.; Cok, Cladosiphon okamuranus; Sja, Saccharina japonica; Noc, Nanochloropsis oceanica; Ptr, Phaeodactylum tricornutum; Pmu, Pseudo-nitzschia multiseries; Cre, Chlamydomonas reinhardtii; Ppa, Physcomitrella patens; Sce, Saccharomyces cerevisiae; Uma, Ustilago maydis; Cne, Cryptococcus neoformans; Dme, Drosophila melanogaster. Note that Phytophthora infestans gene 05545 has two homeoboxes. Supplementary files.
DOI: https://doi.org/10.7554/eLife.43101.010