RAPID TURNOVER OF LIFE-CYCLE-RELATED GENES IN THE BROWN ALGAE

A.P. Lipinska[1], M.L. Serrano-Serrano[2], Akira F. Peters[3], K. Kogame[4], J Mark Cock[1], Susana M. Coelho[1*]

[1]Sorbonne Université, UPMC Univ Paris 06, CNRS, Algal Genetics Group, Integrative Biology of Marine Models, Station Biologique de Roscoff, CS 90074, F-29688, Roscoff, France. [2]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland [3]Bezhin Rosko, 29250 Santec, France. [4]Department of Biological Sciences, Faculty of Sciences, Hokkaido University, Sapporo, 060-0810 Japan

[*]correspondence: coelho@sb-roscoff.fr

ABSTRACT

Background: Sexual life cycles in eukaryotes involve a cyclic alternation between haploid and diploid phases. While most animals possess a diploid life cycle, plants and algae alternate between multicellular haploid (gametophyte) and diploid (sporophyte) generations. In many algae, gametophytes and sporophytes are independent and free living, and may present dramatic phenotypic differences. The same shared genome can therefore be subject to different, even conflicting, selection pressures in each of the life cycle generations. Here, we have analysed the nature and extent of genome-wide generation-biased gene expression in four species of brown algae with contrasting levels of dimorphism between life cycle generations, in order to assess the potential role of generation-specific selection in shaping patterns of gene expression and divergence.

Results: We show that the proportion of the transcriptome that is generation-biased is associated with the level of phenotypic dimorphism between the life cycle stages. Importantly, our data reveals a remarkably high turnover rate for life-cycle-related gene sets across the brown algae and highlights the importance not only of co-option of regulatory programs from one generation to the other but also a key role for newly emerged, lineage-specific genes in the evolution of the

gametophyte and sporophyte developmental programs in this major eukaryotic group. Moreover, we show that generation-biased genes display distinct evolutionary modes, with gametophyte-biased genes evolving rapidly at the coding sequence level whereas sporophyte-biased genes exhibit changes in their patterns of expression.

Conclusion: Our analysis uncovers the characteristics, expression patterns and evolution of generation-biased genes and underline the selective forces that shape this previously underappreciated source of phenotypic diversity.

BACKGROUND

As a consequence of sexual reproduction, the vast majority of eukaryotes have life cycles involving an alternation between haploid and diploid phases [1,2]. The proportion of the life cycle spent in each phase varies dramatically depending on the species. In organisms with haplontic cycles, mitosis only occurs in the haploid stage. Haploid mitosis may lead to asexual (clonal) reproduction such as the case of *Chlamydomonas*, or involve somatic growth and cellular differentiation such as in *Volvox*. In these organisms, the zygote undergoes meiosis immediately after syngamy without undergoing any mitotic divisions. Conversely, in diplontic life cycles, mitosis only occurs during the diploid phase, and meiosis takes place immediately before gamete formation. Diploid mitosis leads to asexual reproduction in unicellular lineages (e.g. diatoms) and to somatic growth and differentiation in multicellular organisms such as Metazoans. Finally, in organisms with haploid-diploid life cycles, mitotic cell divisions occur during both the haploid and diploid phases. In land plants and some algae, these mitotic divisions can lead to the development of two distinct multicellular organisms, one haploid and the other diploid. The haploid organism is

47 generally referred to as the gametophyte, because it produces gametes, and the diploid organism
48 as the sporophyte, because it produces spores. Note, however, that the gametophyte and
49 sporophyte developmental programs are not absolutely linked to ploidy because ploidy and life
50 cycle generation have been shown to be uncoupled during variant life cycles [3,4]. The
51 gametophyte and sporophyte should therefore be thought of as genetically-controlled
52 developmental programs that are coordinated with, but not absolutely linked to, life cycle
53 progression.

54 The evolutionary advantages of life cycles with dominant haploid, dominant diploid, or
55 alternation between two phases have been subject to extensive theoretical work [5–10]. Models
56 exploring the evolution of haploidy and diploidy assume an alternation of generations with free-
57 living haploid and diploid phases, where expanding one phase reduces the other phase. These
58 models predict that purging of deleterious mutations favours expansion of the haploid phase when
59 recombination is rare, but that diploids are favoured when recombination is common because they
60 mask mutations from selection [6,11]. In contrast, niche differentiation between haploids and
61 diploids may favour the maintenance of biphasic life cycles, in which development occurs in both
62 phases [12]. For instance, gametophytes have been shown to exploit low-resource environments
63 more efficiently whereas sporophytes are more vigorous when resources are abundant [13]. The
64 interplay between genetic and ecological factors has been recently explored [7], in a model that
65 assumes different effects of mutations in haploids and diploids and competition between
66 individuals within a generation. The model predicts that temporal variations in ecological niches
67 stabilize alternation of generations. Empirical support for these models has come from the brown

68 alga *Ectocarpus* sp., where dimorphism between generations has been linked to the occupation of
69 different spatio-temporal niches [14].

70 In organisms with complex life cycles, an allele may be relatively beneficial when expressed
71 in one generation but deleterious when expressed in the other generation (generation-
72 antagonism), and in this case selection acts in opposite directions in haploids and diploids [7,15].
73 With this type of generation-dependent antagonistic selection, evolution favours the expansion of
74 whichever generation gains the greatest fitness advantage, on average, from the conflicting
75 selection pressures [15]. Generation antagonism is expected to be particularly relevant in
76 multicellular species where there is alternation of generations with morphologically dissimilar
77 gametophytes and sporophytes, as in the case of many plants and algae. When fitness optima differ
78 between the gametophyte and sporophyte generations for a shared trait, dimorphism can allow
79 each generation to express its optimum trait phenotype. Accordingly, the evolution of generation-
80 biased gene expression may be one mechanism that could help to resolve this intra-locus
81 'generation' conflict, in a similar manner to mechanisms that resolve sexual antagonism [16,17].
82 Another potential solution to resolve generation-conflict is gene duplication, followed by
83 divergence of the two loci towards distinct optima corresponding to each of the two generations.
84 An equivalent process has been shown to be important in the generation of sex-biased gene
85 expression [17,18]. While the role of sexual selection in shaping phenotypic diversity and in driving
86 patterns of evolution of gene expression has been studied extensively (e.g. [19,20]), we have
87 remained so far largely ignorant about the relationships between generation-biased selection,
88 generation-biased gene expression and phenotypic differentiation.

89 The brown algae (Phaeophyceae) are a group of complex multicellular eukaryotes that
90 diverged from plants and animals more than a billion years. Brown algal life cycles are
91 extraordinarily diverse, exhibiting a broad range of variation in terms of the relative complexities of
92 the gametophyte and sporophyte generations [21]. Here, we selected two pairs of brown algal
93 species from the orders Ectocarpales and Laminariales, which diverged about 95 Mya [22], to trace
94 the evolutionary history of generation-biased gene expression in the brown algal lineage. The
95 selected species exhibit markedly different levels of dimorphism between life cycle generations:
96 the Laminariales species *Macrocystis pyrifera* and *Saccharina japonica* have complex sporophyte
97 but highly reduced gametophyte generations, whereas the Ectocarpales species include
98 *Scytosiphon lomentaria,* which has a reduced sporophyte but a morphologically complex
99 gametophyte, and *Ectocarpus* sp. which has gametophyte and sporophyte generations of similar
100 complexity. We show that a large proportion of the transcriptome of brown algae exhibit
101 generation-biased expression, and that the set of life-cycle-biased genes turns over extremely
102 rapidly during evolution due to a combination of two processes: *de novo* birth of genes with
103 generation-biased expression and gain/loss of generation-biased expression by orthologous loci.
104 Our results uncover the characteristics, expression patterns and evolution of generation-biased
105 genes and underline the selective forces that shape this previously underappreciated source of
106 phenotypic diversity.

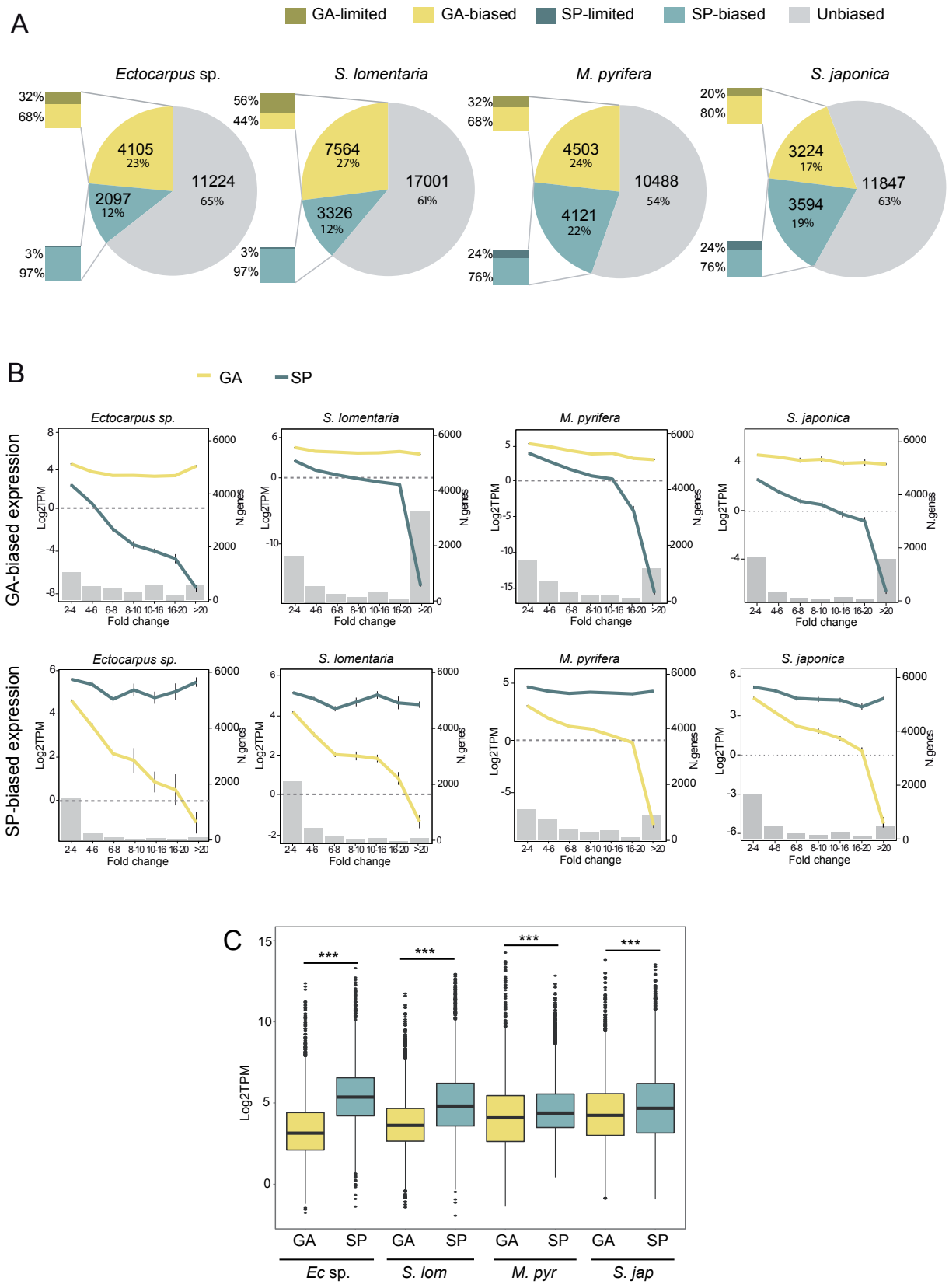
107 ## RESULTS

108 ### Measures of phenotypic differentiation between gametophyte and sporophyte generations

The number of different cell types in each generation and the ratios of the sizes of the gametophyte and the sporophyte at maturity were used as proxies to assess the degree of morphological complexity and the level of phenotypic dimorphism between life cycle generations in the four brown algal species studied (Table S1). Using these parameters, the Laminariales species *M. pyrifera* and *S. japonica* exhibited the highest level of phenotypic differentiation between generations, with the sporophyte being more complex than the gametophyte, both in terms of number of cell types and in terms of the size. As far as the Ectocarpales species were concerned, dimorphism between generations were also marked in *S. lomentaria*, but with the gametophyte being more complex than the sporophyte. *Ectocarpus* sp. exhibited the lowest level of differentiation between the gametophyte and sporophyte generation (Table S1; Figure S1).

Patterns of generation-biased gene expression in gametophytes and sporophytes

We used DEseq2 to compare patterns of gene expression in gametophytes and sporophytes for each of the four brown algal species. The highest level of generation-biased gene expression was detected in *M. pyrifera*, where 46% of the transcriptome was generation-biased. Generation-biased genes represented 37%, 36% and 35% of the transcriptomes of *S. lomentaria, S. japonica* and *Ectocarpus* sp., respectively (Figure 1A). In general, more transcripts were gametophyte-biased than sporophyte-biased. This difference was most marked in *S. lomentaria*, where almost twice as many genes were gametophyte-biased (27% of the transcriptome) than were sporophyte-biased. In both *Ectocarpus* sp. and *S. lomentaria*, the fraction of sporophyte-biased transcripts was relatively low (12%) but the proportion was higher in species that have a more conspicuous sporophyte generation (i.e., both Laminariales), with 19-22% of the transcriptome being sporophyte-biased (Figure 1A).



131

*Figure 1. Generation-biased gene expression across the four-brown algal species. (**A**) Proportion of unbiased, gametophyte- and sporophyte-biased genes across the four-studied species. Bar inserts represent the proportion of generation-specific genes among the generation-biased genes in each species. (**B**) Mean gene expression levels (log2TPM) at several degrees of generation-bias (fold change, FC, represented by grey histograms) for gametophyte-biased (yellow) and sporophyte-biased (blue) genes in the four-studied species. The number of genes in each category of FC are represented in the right side of the graph. Error bars represent standard errors. GA: gametophyte; SP: sporophyte. (**C**) Boxplot showing the mean expression levels (log2TPM) of gametophyte- and sporophyte-biased genes.*

Generation-biased genes were defined as being generation-limited when the TPM for one of the two generations was below the 5th percentile (see Methods) (Figure 1A). Between 35% and 59% of the generation-biased genes were generation-limited, depending on the species. The proportion of generation-limited genes was larger for the gametophyte than for the sporophyte generation in *Ectocarpus* sp. and *S. lomentaria*. This trend was particularly marked for *S. lomentaria* (which has a dominant gametophyte generation) where more than half of the gametophyte-biased genes were gametophyte-limited. In contrast, in the two Laminariales species, similar proportions of generation-limited genes were observed in both generations (Figure 1A).

To examine the relationship between degree of generation-biased expression and transcript abundance (expression level), the generation-biased genes were grouped according to the fold change (FC) difference between gametophyte and sporophyte samples, and the mean expression levels in gametophytes and sporophytes (log_2TPM) were plotted for each group (Figure 1B). This analysis indicated that, overall, the most marked levels of generation-biased expression (high fold changes) were the result of down-regulation of genes in the generation where they were expressed more weakly, rather than strong up-regulation in the generation where they were expressed more strongly. However, for gametophyte-biased genes, the expression in sporophytes reached the lower threshold (about log_2TPM<0) much faster than the expression of sporophyte-biased genes in

gametophytes. In other words, when genes exhibited a moderate to high degree of gametophyte-biased expression, this was predominantly due to strong downregulation (silencing) of these genes in the sporophyte generation (Figure 1B).

Interestingly, in the Ectocarpales species more than 80% of the sporophyte-biased genes exhibited fold changes of between 2 and 6, whereas in the Laminariales species (which have a dominant sporophyte generation) a greater proportion of the sporophyte-biased genes exhibited very high fold changes between generations, with between 13% and 23% in *S. japonica* and *M. pyrifera,* respectively, exhibiting fold changes of more than 20 (Figure 1B; Table S2). Nevertheless, in all four species the majority of the generation-biased genes with very strong bias (FC>20) were gametophyte-biased (Figure 1B; Table S2), with as many as 50% of the gametophyte-biased genes in *S. lomentaria* belonging to this group (15% in *Ectocarpus* sp., 28% in *M. pyrifera* and 44% in *S. japonica*).

We also noted that, on average, sporophyte-biased genes were expressed at significantly higher levels than gametophyte-biased genes in all four species (Wilcoxon test, $p<2e-12$) (Figure 1C). In general, the complexity of each life cycle generation, both in terms of number of cell types and size of the organism, tended to be correlated with the number of generation-biased genes (Table S3). For instance, *S. lomentaria*, which has a dominant gametophyte generation, possessed the highest proportion of gametophyte-biased genes. Conversely, *M. pyrifera* and *S. japonica*, whose sporophytes are much larger than the gametophytes (Figure S1, Table S3), exhibit the highest proportion of sporophyte-biased genes.

High turnover of generation-biased gene sets in the brown algae

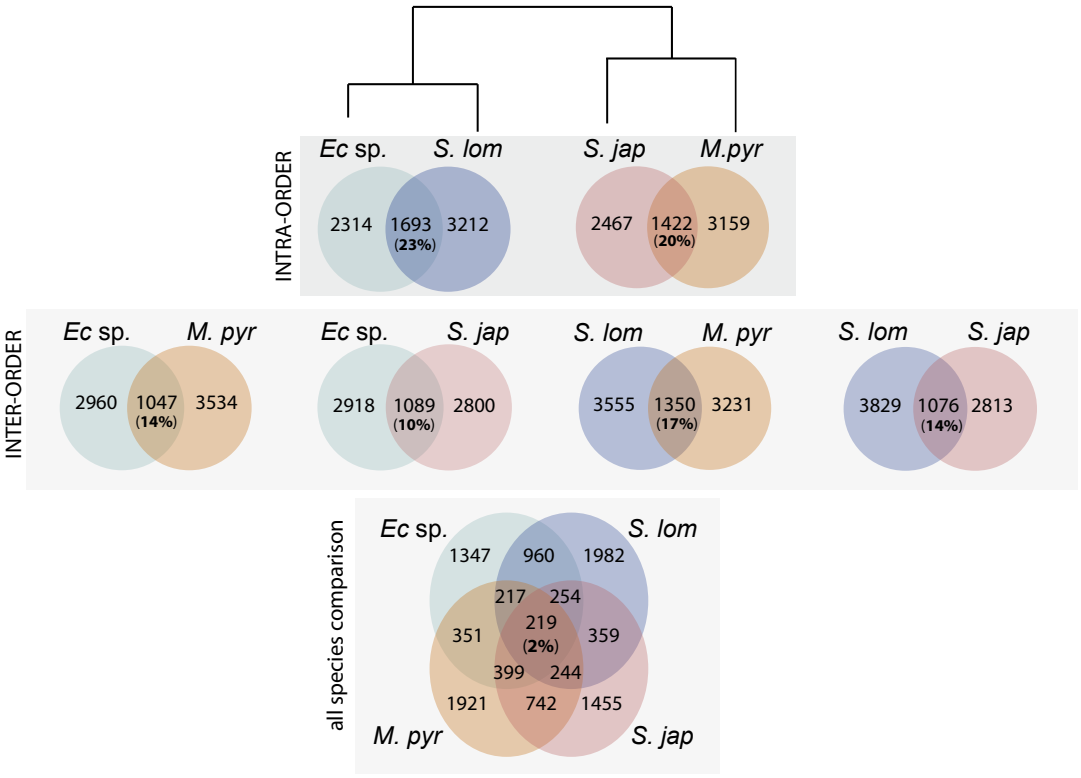
171 Comparisons of orthogroups (OGs) containing generation-biased genes showed that they
172 were poorly conserved between pairs of species (Figure 2A). Only 23% and 20% of the orthogroups
173 (OGs) containing generation-biased genes were shared by the Ectocarpales and Laminariales
174 species pairs respectively, and conservation between pairs of species from different orders was
175 even lower (14% to 17%). Only 2% of the generation-biased OGs were conserved across all four of
176 the study species (Figure 2A). Importantly, a large proportion of the generation-biased genes
177 (between 36% and 63%) did not have orthologs in the genome of any of the other three study
178 species nor in the genomes of four other distant Stramenopile species (Figure 2B; Table S4). We
179 refer to these taxonomically-restricted genes hereinafter as "orphan" genes. The orphan genes
180 were not included in the OGs analysis described above (Figure 2A) because most orphans are not
181 members of an OG. The analysis of OGs therefore actually overestimated the degree to which
182 generation-biased gene sets were conserved across species.

183 The sets of gametophyte-biased genes in *Ectocarpus* sp., *S. lomentaria* and *S. japonica* and
184 the set of sporophyte-biased genes in *S. japonica* were significantly enriched in orphan genes

185 compared to the whole genome (Fisher test, $p<0.03$, $p=0.0008$, respectively), suggesting that *de*

186 *novo* evolution has played an important role in the emergence of the generation-biased gene sets.

A



B

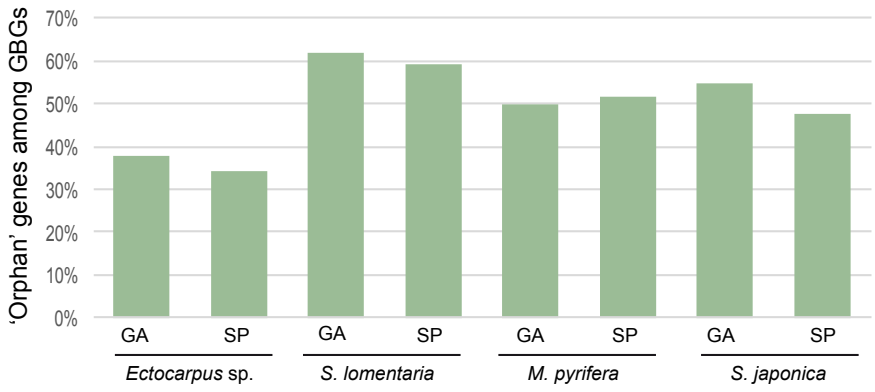


*Figure 2. OGs with generation-biased genes are poorly conserved across brown algal species and the generation-biased gene sets include many orphan genes. (**A**) Shared OGs with generation-biased genes across the four-studied species. Venn diagrams representing the number of shared versus species-specific generation-biased OGs. Comparisons were made at several evolutionary distances. (**B**) Proportion of orphan (taxonomically-restricted) genes within the generation-biased gene sets of each of the four-studied species. GBGs: generation-biased genes.*

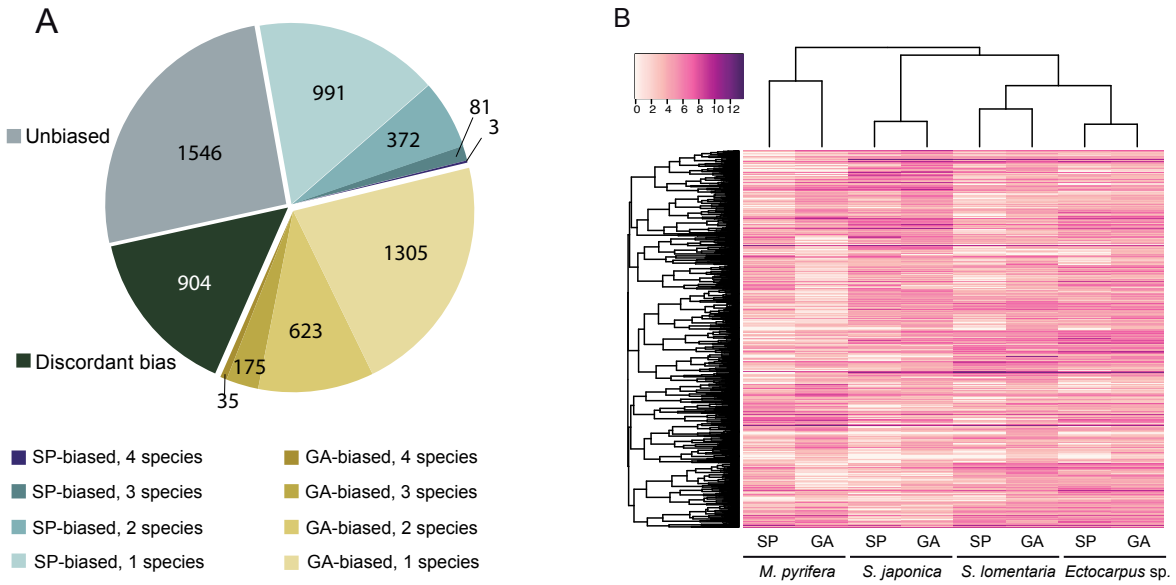
Evolution of generation-biased expression of orthologous genes across brown algal species

To further analyse the evolutionary history of the generation-biased genes, we focused on genes for which there was a clear orthologous relationship across the four species. A large set of 15,888 orthologous groups (OGs), identified using OrthoMCL, was screened to identify 6,035 single copy orthologous genes with either 1:1:1:1 or 1:1:1:0 occurrence across the four brown algal species (see methods for details). We will refer to this set of OGs as "all single orthologues" (ASOs).

The ASO dataset was used to assess the conservation of generation-biased gene expression across the four species. Of the 6,035 ASOs, 4,489 (74%) included genes that were generation-biased in at least one of the species. However, only 35 gametophyte-biased genes and three sporophyte-biased genes consistently exhibited patterns of generation-biased expression across all four species (Figure 3A). The number of genes with conserved generation-biased expression increased to 175 gametophyte-biased and 81 sporophyte-biased when we took into account orthologous genes with generation-bias in three species (with the ortholog missing or unbiased in the fourth species) (Figure 3A). Fifteen percent of the ASOs (904 of the 6,035) exhibited discordant generation-biased expression patterns, so that, for example, the orthologue of a gene that was sporophyte-biased in one species was gametophyte-biased in at least one of the other three species (Figure 3A).

205 We used hierarchical clustering of expression levels for all the members of the 1:1:1:1 ASO
206 dataset with at least one generation-biased member in one of the studied species to visualize global
207 transcription patterns within and among the four species. In this analysis, samples clustered
208 primarily by phylogenetic relatedness and not according to life cycle stage (Figure 3B), reflecting
209 the low level of conservation of generation-biased expression patterns of gene expression across
210 the lineages.

211 Taken together, these analyses indicated that the gametophyte-biased transcriptome tended
212 to be more conserved than the sporophyte-biased transcriptome, but overall, generation-biased
213 expression of the ASO dataset was extremely poorly conserved across the four brown algal species.



214

Figure 3. Conservation of generation-biased gene expression across species. ***(A)*** *Numbers of ASOs showing unbiased, discordant bias or different degrees of shared bias between the four studied species. GA: gametophyte; SP: sporophyte (****B****) Hierarchical clustering and heatmap of gene expression for all the members of the 1:1:1:1 ortholog dataset with at least one generation-biased member in one of the studied species (Heatmap3 package, R). The dendogram was constructed using hierarchical clustering with 1000 bootstraps (pvclust package, R).*

Generation-biased gene expression within the Ectocarpales and Laminariales

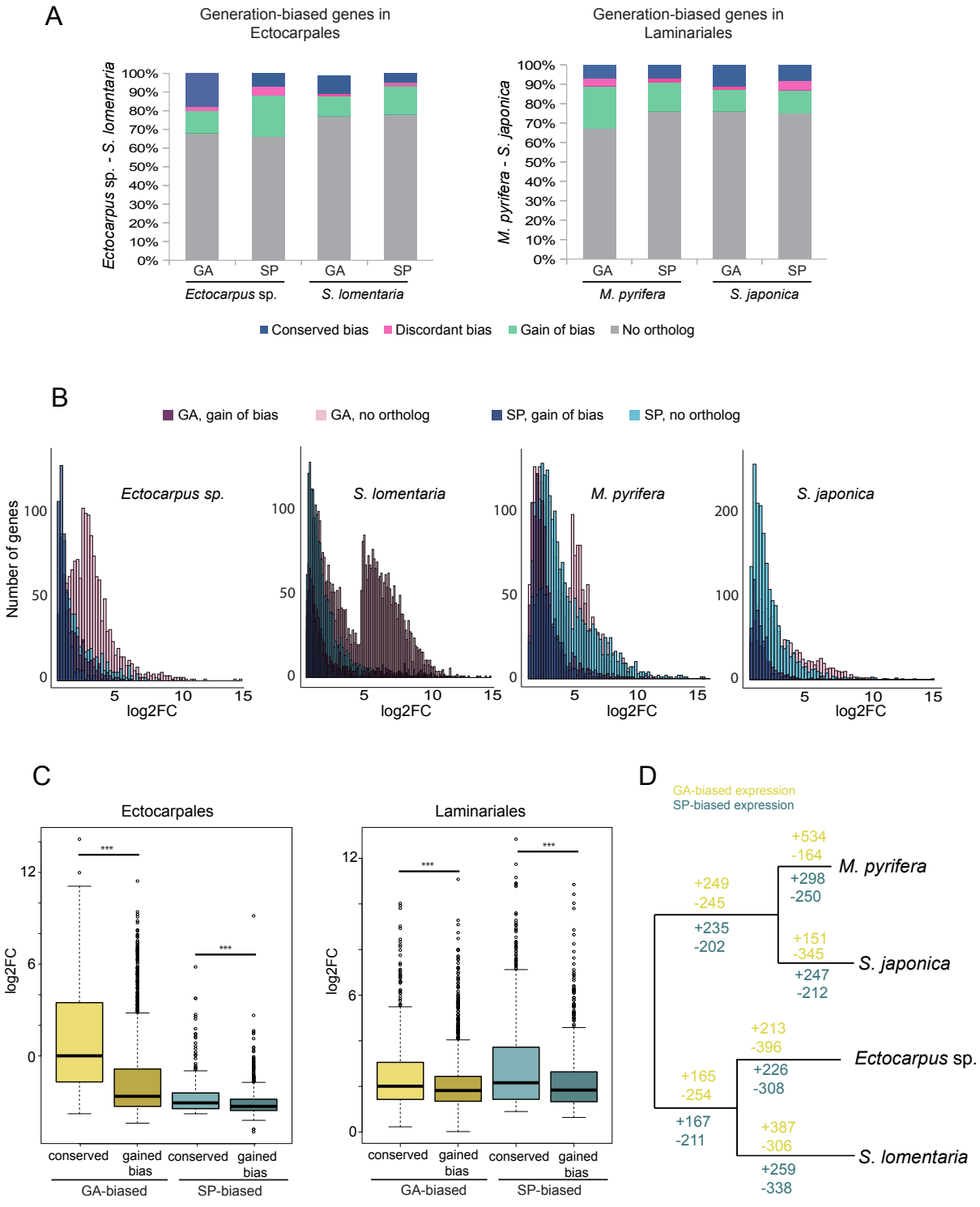
To analyse divergences of generation-biased expression patterns within orders, we used the OrthoMCL analysis to identify single copy (1:1) orthologues shared either by the two Ectocarpales (6,438 OGs) or by the two Laminariales (5,061 OGs) species. These sets of 1:1 OGs were termed "pairwise single orthologues" (PSOs) (Table S5).

Between 22-34% (Ectocarpales) and 24-33% (Laminariales) of the generation-biased genes were PSOs, whereas the remaining generation-biased genes had no ortholog in the other species from the same order (Figure 4A). Between 11-22% (for both Ectocarpales and Laminariales) of the generation-biased genes gained either sporophyte- or gametophyte-biased expression in one of the species (Figure 4A), whereas discordant generation-biased expression was observed for 1.5-5% (Ectocarpales) and 2-5% (Laminariales) of the generation-biased genes (Figure 4A).

Overall, species-specific gametophyte-biased genes presented greater bias, measured as fold change (Figure 4B). This was particularly pronounced in *S. lomentaria,* where about half of the species-specific genes with gametophyte-bias had expression levels at least 50 times higher ($\log_2$FC>5.7) than in the sporophyte. In Laminariales, conversely, it was the species-specific genes with sporophyte-biased expression that presented overall higher fold changes (Figure 4B). In other words, newly emerged, species-specific genes showed a strong magnitude of bias in gametophytes of Ectocarpales and sporophytes of Laminariales.

A correlation was observed between the level of bias and the conservation of generation-biased genes within each lineage (Ectocarpales and Laminariales), i.e. the mean fold change of genes that were conservatively generation-biased (both gametophyte-biased or both sporophyte-

236 biased) across the species pairs was significantly higher than that of genes were generation-biased
237 in one of the species but unbiased in the other (Figure 4C; Wilcoxon test, p<1e-05). Taken together,
238 these analyses suggested that the rapid evolution of generation-biased gene sets involves not only
239 the emergence of new generation-biased genes but also the emergence, in species-specific
240 fashion, of novel generation-biased expression patterns associated with existing orthologous
241 genes.



242

*Figure 4. Conservation of generation-bias across the Ectocarpales (Ectocarpus sp. and S. lomentaria) and Laminariales (S. japonica and M. pyrifera). (**A**) Proportion of the PSOs among generation-biased genes. In grey, genes which had no ortholog in the other species of the same order, colour indicates PSOs. For the PSOs, genes with 'Conserved bias' (dark blue) exhibited the same bias in the two species. Genes with 'Discordant bias' (pink) were gametophyte-biased in one species and sporophyte-biased in the other species. 'Gained bias' genes (green) were generation-biased in one species but not in the other species. (**B**) Distribution of magnitude of generation-bias (log2 fold changes) for PSOs that gained bias (ortholog is unbiased in the other species) and species-specific genes with biased expression (no ortholog in the other species). Orthology was established per lineage in pairwise comparisons between Ectocarpus sp.-S. lomentaria and S. japonica-M. pyrifera. (**C**) Overall level of generation-biased expression (log2FC) for PSOs that are conserved versus PSOs that gained bias in Ectocarpales and Laminariales. (**D**) Representations of gene gain/loss events across the branches of the Ectocarpales and Laminariales phylogeny. Expected numbers of events are based on multiple stochastic mappings (see methods for detail).*

Evolutionary history of generation-biased gene sets

We used a phylogenetic stochastic mapping approach to investigate the evolution of generation-biased gene expression. Phylogenetic stochastic mapping allows reconstruction of the history of trait changes (in our case, generation-bias) based on the estimation of the probabilities and expectations of gain and loss events for each branch of an underlying phylogenetic tree [23]. Rates of gain and loss were equal for both gametophyte- and sporophyte-bias, as determined by a likelihood ratio test between the ER and ARD models (all p-values >0.05). The stochastic mapping results highlighted widespread and rapid turnover of generation-biased genes during evolution of the Laminariales and Ectocarpales (Figure 4D). Specifically, more events of gain of bias were observed for gametophyte-biased genes, and, conversely, sporophyte-biased genes presented more events of loss of bias. Overall, gametophyte-biased genes presented a more rapid turnover with a higher total number of events compared with sporophyte-biased (3409 versus 2953).

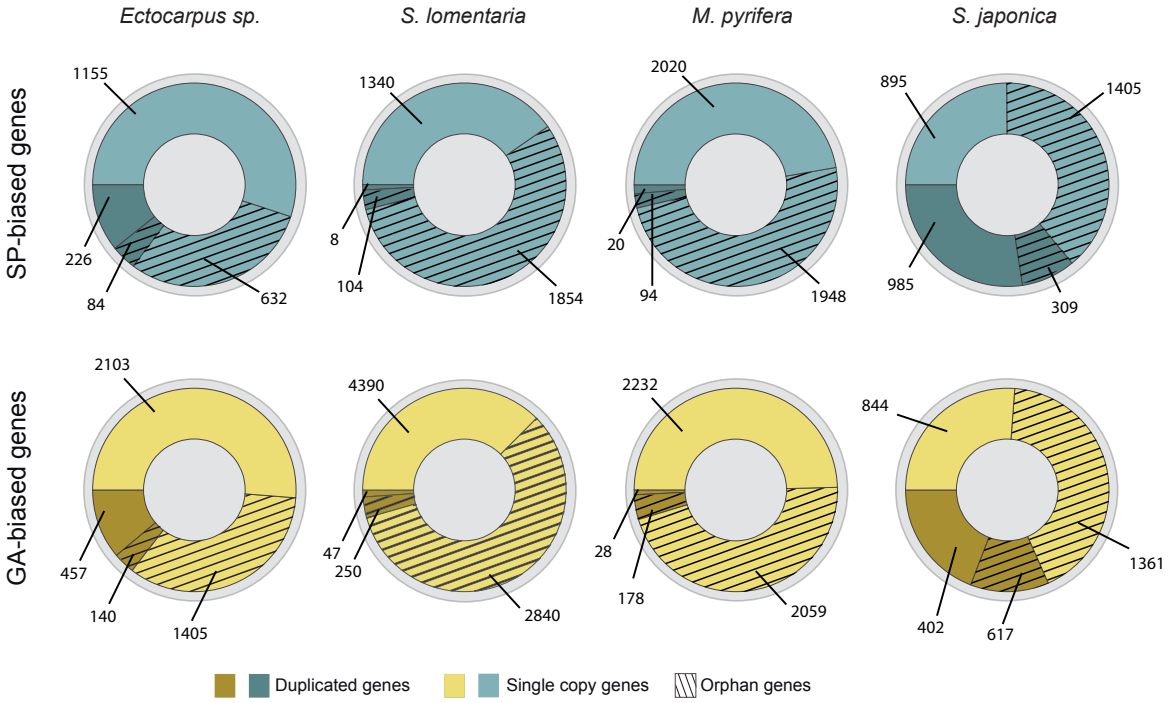
Duplicated generation-biased genes

Gene duplication and generation-specific co-option of paralogs may be a mechanism to resolve potential generation antagonism due to evolutionary divergence between the two

258 generations. Analysis of in-paralogs identified by OrthoMCL indicated that the generation-biased
259 gene sets of *Ectocarpus* sp., *S. lomentaria* and *M. pyrifera* were not enriched in members of
260 duplicated gene pairs compared with the rest of the genes in each genome (Fisher test, p=1).
261 However, duplicated genes constituted 32 and 36% of gametophyte and sporophyte-biased genes
262 in *S. japonica* (Table S4, Figure 5), which was significantly more than expected by chance (Fisher
263 test, p<2e-16 for both comparisons).

264 Discordant bias was observed for 10-44% of generation-biased in-paralog pairs. *S. japonica*
265 had significantly more generation-biased in-paralog pairs that showed discordant bias than the
266 other three species (Chisqr test, p<0.007). The set of in-paralogs with discordant bias was
267 completely different for each species, indicating that duplication of genes followed by acquisition

268 of two, opposite generation-biased expression patterns by the resulting in-paralogs occurred
269 independently in each of the four lineages.



270

Figure 5. Proportion of single copy versus duplicated genes and amount of orphan (taxonomically restricted) genes among gametophyte and sporophyte-biased genes in each of the brown algal species.

271 ## Predicted functions of generation-biased genes

272 An analysis of gene ontology (GO) terms associated with the generation-biased genes was
273 carried out using Blast2GO [24] to search for enrichment in particular functional groups. First,
274 Blast2GO analysis was carried out for each species, in order to relate gene function to the
275 phenotypic generation dimorphisms specific to each species.

276 The GO terms associated with *M. pyrifera* and *S. japonica* sporophyte-biased genes were
277 enriched in biological processes related to reproduction, carbohydrate metabolism, protein

modification, growth and development, signalling, cell communication, response to external stimulus and homeostasis (Fisher exact test, $p<0.05$, Table S6). Interestingly, a similar set of GO terms was enriched for the gametophyte-biased genes of *S. lomentaria*, in addition to categories related to sexual reproduction and cilium motility (Figure S2, Table S6). This result suggest that similar genetic processes are at work in the morphologically complex, long-lived, dominant generations of these three species, despite the large morphological differences between gametophytes of *S. lomentaria* and sporophytes of Laminariales and the limited number of shared generation-biased genes.

Analysis of the gametophyte-biased genes from all the studied species identified six GO terms related to biological processes that were consistently significantly enriched in all species (Fisher exact test $p<0.05$). These terms included microtubule and flagellar movement-related categories and corresponded to between 10% and 50% of gametophyte-biased genes with assigned ontology (Table S7). Three GO terms were consistently enriched for the sporophyte-biased genes of all the studied species. These terms, which were related to carbohydrate metabolism and small GTPase signalling processes, corresponded to between 10% and 20% of the sporophyte-biased genes in each species.

Structural characteristics of the generation-biased genes

Several structural characteristics (GC and GC3 content, coding region size and intron number) were compared between sporophyte-biased, gametophyte-biased and unbiased genes (Figure S3; Table S8). Gametophyte-biased genes tended to have longer coding regions, to possess more introns and to have a lower GC3 content than unbiased genes in all four species (Wilcoxon test, p-

299 value<6e-04). In contrast, sporophyte-biased genes did not present a consistent trend in relation
300 to the unbiased genes in any of the species studied. Excluding orphan genes from the analysis did
301 not significantly change the results indicating that the structural characteristics of gametophyte-
302 biased genes were not due to the fact that these genes are evolutionary younger (Figure S4).

303 ## Evolutionary features of the generation-biased genes

304 The evolutionary dynamics of generation-biased genes was investigated by calculating ratios
305 of pairwise synonymous to nonsynonymous substitution rates and comparing these data with gene
306 expression divergence (see Methods). This analysis, which was applied to all the PSOs that could
307 be examined for each order (Ectocarpales and Laminariales), indicated that gametophyte-biased
308 genes evolve faster (i. e., had higher dN/dS ratios) than unbiased or sporophyte-biased genes in
309 both Ectocarpales species (Figure 6A; Wilcoxon test, p<1e-11) and in *M. pyrifera* (Wilcoxon test
310 p<0.005). The kelp *S. japonica* was the only exception, no significant difference was observed
311 between the rates of evolution of gametophyte-biased and unbiased genes for this species (Figure
312 6A).

313 Sex-biased genes have been shown to exhibit accelerated rates of evolution in *Ectocarpus* sp.
314 [25] and many of the gametophyte-biased genes also exhibit a sex-biased pattern of expression
315 (809 genes or about 20%) because this is the sexual generation of the life cycle. However, when
316 these sex-biased genes were removed from the dN/dS analysis, the gametophyte-biased genes still
317 showed faster evolutionary rates than unbiased or sporophyte-biased genes (Figure S5A; Wilcoxon
318 test p-value=6.3e-13 for the gametophyte versus unbiased comparison, p-value=2e-09 for the
319 gametophyte versus sporophyte-biased comparison) indicating that the faster evolutionary rates

320 of gametophyte-biased genes were not solely due to the presence of sex-biased genes. Similar
321 results were obtained when we subdivided the generation-biased genes into conserved bias and
322 species-specific bias (i.e. genes that showed generation bias in one species but were unbiased in
323 the other) (Figure S5B, S5C**).** This latter analysis suggested that the faster evolutionary rates of
324 gametophyte-biased genes were not correlated with the degree of conservation of expression
325 across the lineages.

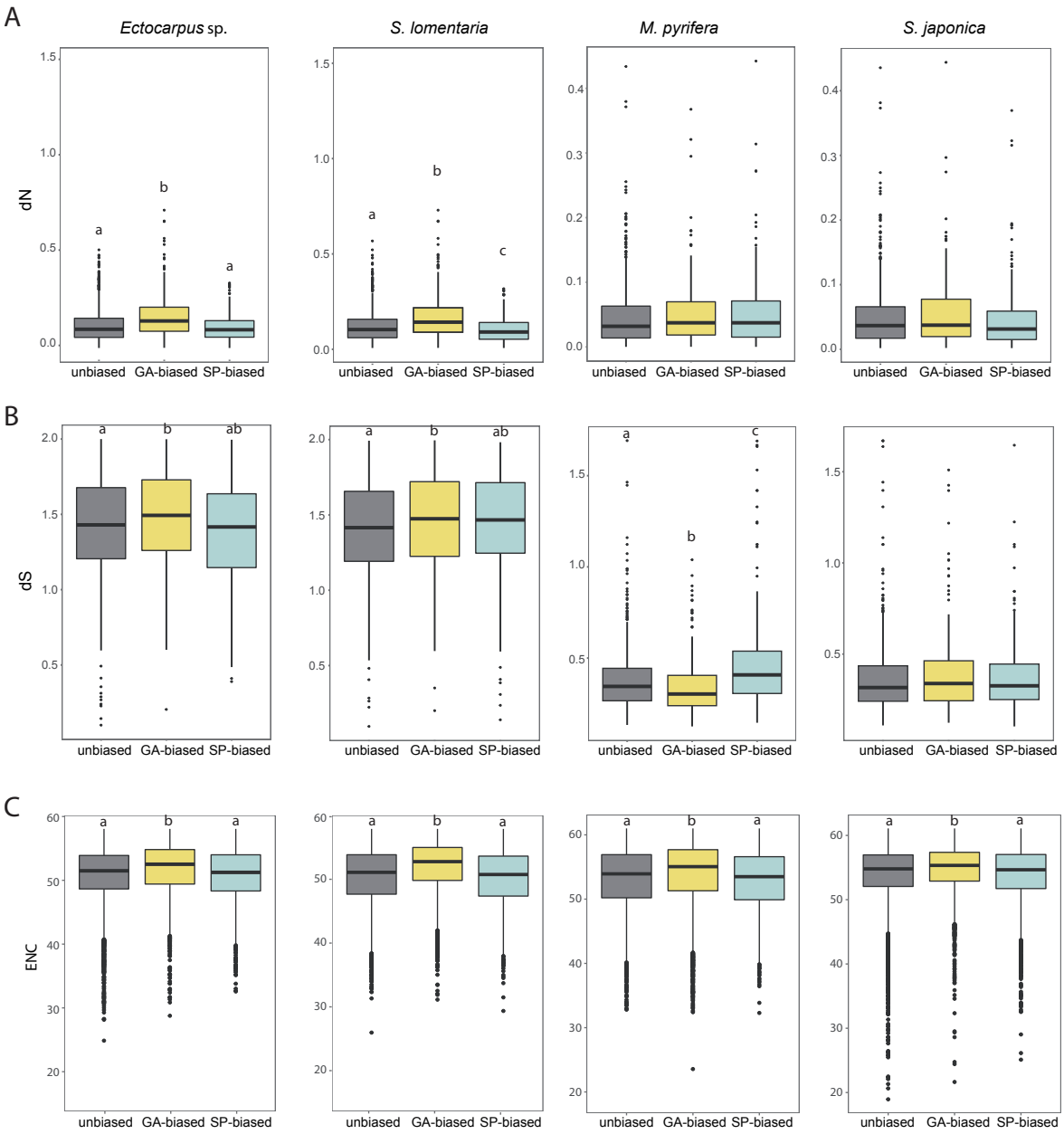


Figure 6. Evolution of generation-biased genes. (A) Evolutionary rates measured as dN/dS between species pairs (Ectocarpus sp./S. lomentaria, M. pyrifera/S. latissima) for unbiased, gametophyte-biased and sporophyte-biased genes in the four brown algal species. (B) Gene expression divergence measured as Euclidean distances for unbiased, gametophyte- and sporophyte-biased genes in each of the four brown algal species. Different letters above the plots indicate significant differences (Wilcoxon test; P<0.05).

The higher evolutionary rates of gametophyte-biased genes were due to the accumulation of both synonymous and non-synonymous changes (Figure S6A, S6B). There was no correlation

between dN/dS and fold change of generation-biased gene expression between generations (Spearman's rho=-0.037). Codon usage bias (CUB), measured as the effective number of codons (ENC), indicated that gametophyte-biased genes had significantly lower codon usage bias (i.e. higher ENC) than sporophyte-biased and unbiased genes in both Ectocarpales and Laminariales (pairwise Wilcoxon test p<3.0e-11) (Figure S6C). We found no significant difference in CUB between sporophyte-biased and unbiased genes.

Overall, order-specific genes (i.e., 'young' genes) exhibited higher evolutionary rates (dN/dS) than genes conserved across the four species ('old' genes) for both the Ectocarpales (Wilcoxon test, p-value=2.8e-13) and Laminariales (Wilcoxon test, p-value=6.5-06), suggesting that younger genes are evolving more rapidly (Figure S7).

To assess whether increased protein divergence rates were due to increased positive selection or relaxed purifying selection, we performed a maximum likelihood analysis using codeml in PAML4. In addition to the four study species, we searched for orthologs of the ASOs in published transcriptome data for three additional *Ectocarpus* species (*E. fasciculatus,* un unnamed *Ectocarpus* species from New Zealand and *Ectocarpus siliculosus)* and the recently published genome of *C. okamuranus* [26]. The 522 conserved orthologs identified by this analysis included 393 orthologs that exhibited generation bias in at least one of the studied species. For 42 of these comparisons both pairs of models (M1a-M2a, M7-M8) suggested positive selection and a total of 111 genes were predicted to be evolving under positive selection as indicated by the model M7-M8 alone (Table S9). Among the 111 genes identified by the M7-M8 model, 83 exhibited generation-bias in at least one species. Taken together, our analysis is consistent with the idea that a subset of the generation-biased genes exhibit signatures of positive selection, although the set of generation-

biased genes was not significantly enriched in genes that were predicted to be under positive selection (Fisher's exact test, P=0.8045).

When patterns of gene expression were considered, measured as Euclidian distances for PSOs within each order, sporophyte-biased genes showed overall significantly larger divergence that unbiased or gametophyte-biased genes (Figure 6B; Wilcoxon test p<3e-7). Furthermore, gametophyte-biased genes exhibited the most conserved expression patterns within each order, even in comparison to unbiased genes.

Taken together, our data suggests that gametophyte- and sporophyte-biased genes have distinct patterns of evolution: gametophyte-biased genes tend to exhibit rapid evolution of their coding sequence whereas sporophyte-biased genes tend to exhibit changes to their patterns of expression.

DISCUSSION

Here, we have used four phylogenetically-diverse brown algal species with different levels of generation dimorphism and complexity to investigate genome-wide generation-biased gene expression patterns and to assess the potential role of generation-specific selection in shaping these patterns of gene expression.

Differential gene expression underlies phenotypic dimorphism between life cycle generations

Between 35% and 46% of the genome of the brown algae species studied here was differentially regulated during the gametophyte and sporophyte generations of the life cycle. This is substantially more than in the moss *Funaria hygrometrica,* where only 4% of the genome is

differentially expressed [27], and even exceeds the situation in *Arabidopsis,* where about 30% of the genome is generation-biased [28]. Comparative analysis between these two land plants showed that the relative proportion of generation-biased genes assigned to the two life cycle generations was lower in the moss than in *Arabidopsis* [27], consistent with the lower level of phenotypic dimorphism between generations in the former. Likewise, we found that for the brown algae studied here the relative numbers of generation-biased genes during each generation was broadly correlated with size differences between the two generations. Despite this tendency, however, the absolute number of gametophyte-biased genes was consistently greater than the number of sporophyte-biased genes across all four studied species, even in kelps where the gametophyte is much less complex, morphologically, than the sporophyte. This tendency is difficult to explain but we note that a previous analysis using *Ectocarpus* sp. also indicated that the gametophyte developmental program deployed more genes than the sporophyte program [4]. Based on this difference it was proposed that the switch to the sporophyte program involves predominantly gene repression. Again, this hypothesis is consistent with our analysis indicating that gametophyte-biased gene expression tends to be the result of downregulation during the sporophyte generation.

Generation-conflict and generation-biased gene expression

In many brown algal species, free-living gametophytes and sporophytes display extensive morphological and physiological dimorphisms (reviewed in [1,29,30]) and this phenotypic diversity reflects in many cases different ecological niche preferences for gametophytes and sporophytes (e.g.[14,31]). Gametophyte and sporophyte development and function are controlled by a common genome, with a large number of genes carrying out functions during both generations. When there are marked morphological and physiological differences between the two generations, as is the

393 case for most of the species studied here, this can lead to conflict due to genes being submitted to
394 different selection pressures during the different generations of the life cycle. Generation-biased
395 gene expression is one mechanism to reduce inter-generational conflict, allowing gene products to
396 be targeted specifically to one generation (although this does not necessarily mean that every
397 generation-biased gene arose due to generation antagonism). Gene duplication, followed by
398 acquisition of generation-bias through neo-functionalization, can play an important role in the
399 resolution of generation conflict and we found some evidence for this, at least for one of the kelp
400 species. Note that gene duplication followed by neo-functionalisation has also been proposed as
401 one of the mechanisms that allow resolution of sexual antagonism (reviewed in [17]).

402 ### Generation-biased genes turnover rapidly during the evolution of the brown algae

403 Perhaps the most striking result of our analysis is the remarkably limited number of
404 generation-biased genes that were shared by all the four of the studied species, indicating a rapid
405 turnover of life cycle-biased genes in brown algae. This turnover appears to be due to a
406 combination of two processes: emergence of new genes with generation-bias and gain/loss of bias
407 in existing, orthologous genes. The former mechanism appears to have been of paramount
408 importance as between 36% and 63% of generation-biased genes were orphans, depending on the
409 species. Phylogenetic stochastic mapping results were consistent with rapid loss and gain of bias in
410 orthologous genes, and highlighted a particularly high rate of turnover for gametophyte-biased
411 genes.

412 The ancestor of brown algae is thought to have alternated between multicellular, isomorphic
413 gametophyte and sporophyte generations without a clearly dominant generation [22]. From this

414 morphologically simple ancestor, there would have been a tendency, in most brown algal lineages,
415 to evolve towards increased complexity of either the gametophyte or the sporophyte generation
416 [22]. Our data indicate that this increase in size and developmental complexity was accompanied
417 by an overall increase in the proportion of the transcriptome that become gametophyte- or
418 sporophyte-biased, depending on the lineage.

419 Recent analysis of *Ectocarpus* sp. developmental mutants have indicated that the evolution
420 of the sporophyte and gametophyte genetic programs involved both co-option of genetic programs
421 from one generation to the other and generation-specific innovations [34,35]. We observed that a
422 subset of the expressed genes in the four brown algal species exhibited switching of bias between
423 life cycle generations in the different lineages, in line with the idea that the evolution of the
424 generation-specific developmental programs in the brown algae has, to some extent, involved
425 sharing of genes between generations. But importantly, our evidence indicates that the evolution
426 of brown algal gametophyte and sporophyte developmental programs has been predominantly
427 driven by the emergence of lineage-specific, orphan genes, suggesting a more important role for
428 generation-specific innovations during the evolution of the sporophyte and gametophyte genetic
429 programs.

430 The evolutionary origins of the sporophyte and gametophyte developmental programs in
431 land plants have been intensively studied, particularly with regard to the question of whether each
432 generation has independently evolved its own developmental pathways or, alternatively, whether
433 there has been recruitment of developmental programs from one generation to the other during
434 evolution [36–38]. It is currently thought that the developmental networks that implement land
435 plant sporophyte programs were mainly recruited from the gametophyte generation, which was

436 initially the dominant generation [27,36,39] although there is also evidence that there have been
437 sporophyte-specific innovations [27,40]. We suggest however, based on the observations
438 presented here for the brown algae, that it may be an oversimplification to think in terms of one
439 generation gradually recruiting programs from the other generation, and that a more extensive
440 sampling of generation-biased gene sets in land plants may reveal a more dynamic situation
441 involving important amounts of both lineage-specific gene evolution and lineage-specific switching
442 of generation-biased expression patterns.

443 Interestingly, despite the marked differences between the generation-biased gene sets of the
444 four studied brown algae, the enriched GO terms associated with genes expressed during the more
445 morphologically complex, long-lived, dominant generation tended to be similar. The predicted
446 functions of both the sporophyte-biased genes of kelps and the gametophyte-biased genes of *S.*
447 *lomentaria* were enriched in GO terms associated with polysaccharide and cell wall biosynthesis,
448 developmental processes, cell signalling and cell communication. These conserved, enriched GO
449 terms could reflect developmental and morphological processes common to dominant life cycle
450 generations, such as extended multicellular growth.

451 ## Rapid evolution of the coding regions of gametophyte-biased genes

452 On average, gametophyte-biased genes were found to be evolving significantly more rapidly
453 (higher dN/dS) than sporophyte-biased and unbiased genes in all the species studied except *S.*
454 *japonica*. This was surprising because purifying selection is expected to be more efficient for genes
455 expressed during the haploid phase of the life cycle due to the absence of masking of recessive and
456 partially recessive mutations [42,43]. However, accelerated evolution of gametophyte-biased genes

457 has also been previously reported in land plant systems [28,41] and a number of hypothesis have
458 been put forward to explain this phenomenon. It has been suggested, for example, that
459 gametophyte-biased genes are under relaxed constrain because of lower expression breadth and
460 low level of tissue complexity [27,41,44–46]. This hypothesis is unlikely to explain our observations
461 because in *S. lomentaria* the gametophyte generation is the dominant phase of the life cycle and is
462 larger and more complex than the sporophyte. It has also been proposed that strong selection on
463 reproductive traits during gametogenesis (i.e., during the gametophyte-generation) may explain
464 the faster rates of evolution of gametophyte-biased genes in plants [41]. However, when sex-biased
465 genes were excluded from the *Ectocarpus* sp. gametophyte-biased gene dataset, the remaining
466 genes still exhibited a significantly higher rate of evolution than that of unbiased or sporophyte-
467 biased rates. Finally, land plant gametophyte-biased gene sets are enriched in young genes [34]
468 and this may affect evolution rate as young genes are known to evolve more rapidly. However, rapid
469 evolution of young genes is unlikely to explain the faster evolutionary rates of brown algal
470 gametophyte-biased genes because the gametophyte- and sporophyte-biased gene sets contained
471 similar numbers of young (faster evolving) genes.

472 We did note, however, that gametophyte-biased genes have overall lower levels of expression
473 than sporophyte-biased genes, and expression levels have been negatively correlated with
474 evolutionary rates [47–49]. Moreover, since that gametophyte-biased genes turn over more rapidly
475 than sporophyte-biased genes, one interesting possibility is that gametophyte-biased genes may
476 be less associated with complex gene interaction networks, and therefore be more dispensable and
477 thus under less constraint [50–52]. More information about gene interaction networks will be
478 needed for the brown algae in order to test this hypothesis.

Sporophyte-biased and gametophyte-biased genes exhibit different patterns of evolution.

In contrast to the gametophyte-biased genes, sporophyte-biased genes did not exhibit overall accelerated rates of evolution of their coding sequences but they did exhibit significantly higher levels of diversification of expression patterns (measured as Euclidean distance), compared to both unbiased and gametophyte-biased genes. Therefore, whilst the gametophyte-biased genes exhibited accelerated evolution of their coding regions, the sporophyte-biased genes appeared to have experienced accelerated evolution of their regulatory sequences. Decoupling of protein sequence evolution and expression pattern evolution has been observed in other eukaryotes (e.g. [50,51], but see [52,53]) but it is not clear why the mechanisms of evolution should differ between gametophyte-biased and sporophyte-biased genes in these brown algal species. It has been suggested that mutations that change protein sequences and mutations affecting gene regulation play different roles during evolution, with genes involved in physiological traits tending to exhibit the former and genes involved in morphological traits evolve primarily in terms of gene expression [58,61]. An in-depth functional analysis of brown algae genes using experimental approaches will be crucial to understand if the different evolutionary modes of gametophyte- and sporophyte-biased genes are associated with different functions of the gene networks underlying each generation.

CONCLUSIONS

This study afforded the first comparative analysis of generation-biased gene expression across several species with complex life cycles to understand the role of generation-specific selection in shaping patterns of gene expression and divergence. Our analyses revealed that an

extensive proportion of the genome exhibits generation-biased expression in the brown algae and the relative proportion of genes that are generation-biased was correlated with the degree of phenotypic dimorphism between generations. Life cycle-biased genes turn over very rapidly during evolution due to a combination of two processes: the gain/loss of generation-biased expression by orthologous loci and the emergence, *de novo,* of genes with generation-biased expression. Our results are consistent with the idea that the evolution of the genetic program associated with each generation appears to have involved the recruitment of genes across generations but the exceptionally high number of generation-biased orphan genes emphasizes an important role for generation-specific developmental innovations in each lineage. Finally, our analysis has revealed that gametophyte and sporophyte exhibit overall distinct evolutionary modes, with gametophyte-biased genes evolving rapidly predominantly at the level of their sequence and sporophyte-biased genes diverging mostly at the level of their patterns of expression.

METHODS

Biological material and generation of genomic and transcriptomic sequence data

The algal strains used, sequencing statistics and accession numbers are listed in Table S10. We used published RNA-seq datasets for gametophytes and sporophytes of the model brown alga *Ectocarpus* sp. [25,62], *S. japonica* (Teng et al., 2017), *M. pyrifera* (Lipinska et al., 2017) and for gametophytes of *S. lomentaria* (Lipinska et al., 2015). *S. lomentaria* sporophytes (strain Zy2) were derived from a controlled laboratory cross between the Asari6 female gametophyte and the Asari9 male gametophyte. Both gametophytes were field collected. Sporophyte clones were grown in 20°C 14h:10h light:dark conditions, in half-strength Provasoli enriched seawater [63] which allowed

521 them to be maintained as immature thalli (absence of meiotic structures). For the gametophyte
522 and sporophyte of *Ectocarpus* sp., gametophyte and sporophyte of *S. lomentaria* and gametophyte
523 of both kelps we used whole thallus for RNA extractions. Total RNA was extracted using the Qiagen
524 Mini kit (http://www.qiagen.com) as previously described (Coelho, et al. 2012). RNA was
525 sequenced with Illumina HiSeq 2000 paired-end technology with a read length of 125 bp (Fasteris,
526 Switzerland) and is available under the accession number detailed in Table S10.

527 For the sporophytes of kelps, we used RNA-seq data produced from replicate samples of
528 blades tissue of adult individuals from natural populations (SRA references provided in Table S10).

529 Sets of reference genes for each species were derived from the published genomes of
530 *Ectocarpus* sp. and *S. japonica* [64,65] or from draft genome assemblies for *S. lomentaria* and *M.*
531 *pyrifera* [62] with gene prediction based on mapping of RNA-seq data using Stringtie [66]. Quality
532 filtering of the raw reads was performed with FastQC
533 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc), and adapter sequences were
534 trimmed using Trimmomatic (leading and trailing bases with quality below 3 and the first 12 bases
535 were removed, minimum read length 50 bp [67]. *Ectocarpus sp.* and *S. japonica* reads were aligned
536 to the reference genomes [65,68] using Tophat2 [59–61]. Gene models for *S. lomentaria* and *M.*
537 *pyrifera* were predicted using Transdecoder (http://transdecoder.sf.net) based on mapping of the
538 RNA-seq data (Tophat2) to the draft genome sequences generated previously for these two species
539 [62]. Gene expression levels were represented as TPMs. Genes with expression values below the
540 5th percentile of all TPM values calculated per species were considered not to be expressed and
541 were removed from the analysis.

542 Identification of generation-biased and generation-limited genes

543 The filtering steps described above yielded a set of expressed genes in the transcriptome that
544 were then classified based on their generation-expression patterns. Genes were considered to be
545 gametophyte-biased or sporophyte-biased if they exhibited at least a 2-fold difference in expression
546 between generations with a false discovery rate (FDR) of < 0.05. Generation-biased genes were
547 defined as generation-limited when the TPM was below the 5^{th} percentile for one of the
548 generations.

549 Gene orthology

550 OrthoMCL (Fischer, 2011) was used to assess orthologous relationships between the genes
551 of the four-studied species (blastp, e-value < 1e-5). OrthoMCL identified a total of 15,888
552 orthogroups (OGs), of which 3,290 contained only one gene per species and therefore represented
553 the set of 1:1:1:1 OGs. An additional 2,745 OGs had only one member in three of the studied
554 species but no ortholog (i.e. the gene was missing) in the fourth species (1:1:1:0 OGs). We
555 considered that these 1:1:1:0 OGs, which most likely represent single copy ancestral genes that
556 were lost in one of the species, also provided useful information about conservation of generation-
557 biased gene expression because they consisted of members from two different orders
558 (Ectocarpales and Laminariales). We therefore combined the two sets of OGs (1:1:1:1 and 1:1:1:0)
559 to create the "all single orthologues" (ASO) dataset, which was composed of a total of 6,035 OGs.
560 The ASO dataset was employed to assess conservation of generation-biased gene expression across
561 the four-studied species.

562 For pairwise comparisons within orders, we selected OGs that contained only one member
563 in each of the two species (6,438 OGs for the Ectocarpales and 5,061 OGs for the Laminariales).
564 We refer to the OGs in these datasets as "pairwise single orthologues" (PSOs).

565 Orphan (or *de novo*) genes (i. e., taxonomically restricted genes) were defined as genes
566 present in the genome of only one species and having no BLASTp match (10-4E value cutoff) with
567 a range of other stramenopile genome-wide proteomes from public databases (indicating that they
568 are likely to have evolved since the split from the most recent common ancestor): the brown alga
569 *Cladosiphon okamuranus* [26] the eustigmatophyte *Nannochloropsis gaditana* [72], the
570 pelagophyte *Aureococcus anophagefferens* [64] and the diatom *Thalassiosira pseudonana* [74].
571 Duplicated genes were identified in each species using the in-paralog list generated by OrthoMCL.

572 ## Prediction of gene function

573 InterProScan [75] and BLAST2GO [24] were used to assign protein function annotations to
574 genes in all four-studied species. Fisher's exact test with a p-value cutoff of 0.05 was used to detect
575 enrichment of specific GO-terms in various groups of generation-biased genes.

576 The visualization of gene ontology data used for Figure S2 was generated using Revigo [76].

577 ## Evolutionary analysis

578 To estimate the evolutionary rates (non-synonymous to synonymous substitutions, dN/dS)
579 for generation-biased and unbiased genes, pairwise analyses were carried on the PSOs for each
580 order (Ectocarpales and Laminariales). Orthologous protein sequences were aligned with Tcoffee
581 (M-Coffee mode [77], the alignments curated with Gblocks [78] and then translated back to
582 nucleotide sequence using Pal2Nal [79] or TranslatorX [80]. Sequences that produced a gapless

583 alignment exceeding 100 bp in length were retained for pairwise dN/dS (ω) analysis using
584 Phylogenetic Analysis by Maximum Likelihood (PAML4, CodeML, F3x4 model of codon frequencies,
585 runmode = −2) [81]. Genes with saturated synonymous substitution values (dS > 2) were excluded
586 from the analysis.

587 The positive selection analysis was carried out using CodeML (PAML4, F3x4 model of codon
588 frequencies) using additional orthologs of the 1:1 best ortholog set from OrthoMCL found in the
589 transcriptomes of three *Ectocarpus* species (*E. fasciculatus,* an unnamed *Ectocarpus* species from
590 New Zealand and *E. siliculosus;*) and in the genome of *Cladosiphon okamuranus* [26]. The analysis
591 was therefore based on data from seven species in total. Protein alignment and curation was
592 performed as described above. Gapless alignments longer than 100 bp containing sequences from
593 at least three species were retained for subsequent analysis. CodeML paired nested site models
594 (M1a, M2a; M7, M8) [81] of sequence evolution were used and the outputs compared using the
595 likelihood ratio test. The second model in each pair (M2a and M8) is derived from the first by
596 allowing variable dN/dS ratios between sites to be greater than 1, making it possible to detect
597 positive selection at critical amino acid residues.

598 The effective number of codons (ENC) was calculated using ENCprime [82] with ribosomal
599 genes as background nucleotide composition.

600 ## Euclidean distances

601 Euclidean distances were estimated for all the PSOs for each of the two orders (Ectocarpales
602 and Laminariales) following the approach of [83]. The following formula was used:

$$\text{EucD} = \sqrt{\sum_{j=1}^{k}\left(x_{1j} - x_{2j}\right)^2}$$

Where x_{ij} is the expression level of the gene under consideration (TPM) in species i (i.e. species 1 or species 2) during stage j (i.e. gametophyte or sporophyte) and k is the total number of stages (i.e. two, gametophyte and sporophyte). All statistical analysis was performed using RStudio (R version 3.4.2).

Stochastic mapping approach to assess the evolutionary dynamics of generation-biased expression

We conducted an evolutionary analyses of the presence and absence of generation-biased gene expression as a dynamic between gain and loss of phyletic patterns [84]. To estimate the evolutionary dynamics of each event we tested whether the rates of gain (0→1) and loss (1→0) of bias were equal (ER model) or different (ARD model), and implemented stochastic mapping for each gene using the *phytools* R package [85]. The number of events on each branch only included those transitions that effectively produced a change of state at the start and end of the specified branch.

DECLARATIONS

Ethics approval and consent to participate: not applicable

Availability of data and materials: All data generated or analysed during this study are included in this published article and its supplementary information files. Details of SRA references for sequence data are included in Table S10. All the datasets used and analysed during the current study are also available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests.

Funding: This work was supported by the CNRS, Sorbonne Université and the ERC (grant agreement 638240). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author's contributions. AL, AFP and KK prepared the biological material and performed experiments. AL and MS performed the computational analysis. AL, MS and SMC analysed data. SMC wrote the manuscript with valuable input from AL and JMC. SMC coordinated the study. All authors read and approved the final manuscript.

Acknowledgements: The authors thank Nicolas Salemin (University of Lausanne) for fruitful discussions and advice on phylogenetic methods, Dieter Muller, Erasmo Macaya and Delphine Scornet for providing the photographs of *M. pyrifera* and *Ectocarpus* sp. shown in Figure S1.

ADDITIONAL FILES

Additional File 1 (.xlx): Supplemental tables.

Table S1. Phenotypic differences between gametophytes (GA) and sporophytes (SP) of the four-studied species.

Table S2. Gametophyte- and sporophyte-biased genes identified for each of the four studied species (DEseq2 FC<2, padj<0.05, TPM>5th percentile).

Table S3. Generation-biased gene expression and morphological complexity of sporophytes and gametophytes across the four-studied species. GA: gametophyte; SP: sporophyte.

642 **Table S4**. Number of duplicated, single copy and orphan genes among the generation-biased
643 genes in each of the studied brown algal species. SP: sporophyte; GA: gametophyte.

644 **Table S5**. Orthology statistics based on the OrthoMCL analysis. GBGs (generation-biased
645 genes); GA (gametophyte); SP (sporophyte); PSOs (pairwise single orthologs); ASOs (all single
646 orthologs).

647 **Table S6**. Enriched gene ontology (GO) terms associated with the generation-biased genes
648 (GBGs) identified for each of the studied species (Fisher test $p<0.05$).

649 **Table S7**. Enrichment in GO terms common to all gametophytes and sporophytes.

650 **Table S8.** Structural characteristics of the generation-biased genes identified for each of the
651 studied species

652 **Table S9**. PAML codeml analysis with the F3X4 substitution model. UB: unbiased; GA:
653 gametophyte-biased, SP: sporophyte-biased.

654 **Table S10**. Strains used in this study and DNA and RNA sequencing data statistics.

655 **Additional file 2 (.doc): Supplemental Figures**

656 **Figure S1.** Brown algal species used in this study. (**A**) *Ectocarpus* sp. sporophyte (**B**) *Ectocarpus*
657 sp. gametophyte (**C**) *S. lomentaria* gametophyte; (**D**) *S. lomentaria* sporophyte; (**E**) *S. japonica*
658 sporophyte; (**F**) *M. pyrifera* sporophyte; (**G**) *S. japonica* gametophyte; (**H**) *M. pyrifera* gametophyte.

659 **Figure S2.** Visualisation of GO terms associated with generation-biased genes for *S.*
660 *lomentaria* gametophytes and *S. japonica* and *M. pyrifera* sporophytes. The scatterplot shows the
661 cluster representatives (i.e. terms remaining after the redundancy reduction) in a two-dimensional

662 space derived by applying multidimensional scaling to a matrix of the GO terms' semantic
663 similarities [76]. Bubble colour indicates p values and bubble size indicates the number of genes
664 assigned with each GO term.

665 **Figure S3**. Structural characteristics (% GC, %GC3, CDS size and intron number) of unbiased,
666 gametophyte- and sporophyte-biased genes across brown algal species.

667 **Figure S4.** Structural characteristics (% GC, %GC3, CDS size and intron number) of unbiased,
668 gametophyte- and sporophyte-biased genes across brown algal species excluding orphan genes
669 from the datasets.

670 **Figure S5**. (**A**) Evolutionary rates (dN/dS) for generation-biased genes when sex-biased genes
671 were excluded from the analysis. GA: gametophyte, SP: sporophyte; SBG: sex-biased genes. (**B** and
672 **C**) Comparison of evolutionary rates of unbiased genes with those of genes that exhibited either
673 conserved bias (generation-biased in both species) or species-specific bias (i.e., were generation-
674 biased in one species but unbiased in the other) in Ectocarpales (**B**) and Laminariales (**C**). SBGs: sex-
675 biased genes.

676 **Figure S6.** Non-synonymous substitutions (**A**), synonymous substitutions (**B**) and codon usage
677 bias (**C**) for unbiased, gametophyte- and sporophyte-biased genes in the four-studied species.
678 Different letters above the plots indicate significant differences ($P<0.05$). Statistical significance was
679 calculated by pairwise comparisons using the Wilcoxon rank sum test with Holm adjustment.

680 **Figure S7**. Evolutionary rates (dN/dS) in Laminariales and Ectocarpales orthologous genes
681 with different evolutionary ages. Yellow dots indicate orthologous genes that are present in the
682 four-studied species ('old' genes, i.e., ASOs), blue dots indicate order-specific orthologous genes

683 ('young' genes, i.e. PSOs excluding ASOs). Overall, young genes evolve faster than older genes.

684 Asterisks indicate a significant difference (Wilcoxon test, p-value<0.007).

685 REFERENCES

686 1. Coelho SM, Peters AF, Charrier B, Roze D, Destombe C, Valero M, et al. Complex life cycles of multicellular
687 eukaryotes: New approaches based on the use of model organisms. Gene. 2007;406:152–70.

688 2. Valero M, Richerd S, Perrot V, Destombe C. Evolution of alternation of haploid and diploid phases in life
689 cycles. Trends Ecol Evol. 1992;7:25–9.

690 3. Arun A, Peters NT, Scornet D, Peters AF, Mark Cock J, Coelho SM. Non-cell autonomous regulation of life
691 cycle transitions in the model brown alga Ectocarpus. New Phytol. 2013;197:503–10.

692 4. Coelho SM, Godfroy O, Arun A, Le Corguillé G, Peters AF, Cock JM. OUROBOROS is a master regulator of
693 the gametophyte to sporophyte life cycle transition in the brown alga Ectocarpus. Proc Natl Acad Sci U S A.
694 2011;108:11518–23.

695 5. Perrot V, Richerd S, Valero M. Transition from haploidy to diploidy. Nature. 1991;351:315–7.

696 6. Otto SP, Goldstein DB. Recombination and the evolution of diploidy. Genetics. 1992;131:745–51.

697 7. Rescan M, Lenormand T, Roze D. Interactions between Genetic and Ecological Effects on the Evolution of
698 Life Cycles. Am Nat. 2016;187:19–34.

699 8. Jenkins CD. Selection and the evolution of genetic life cycles. Genetics. 1993;133:401–10.

700 9. Scott MF, Rescan M. Evolution of haploid-diploid life cycles when haploid and diploid fitnesses are not
701 equal. Evol Int J Org Evol. 2017;71:215–26.

702 10. Mable BK, Otto SP. The evolution of life cycles with haploid and diploid phases. BioEssays. 1998;20:453–
703 62.

704 11. Paquin CE, Adams J. Relative fitness can decrease in evolving asexual populations of S. cerevisiae. Nature.
705 1983;306:368–70.

706 12. Hughes, Otto. Ecology and the Evolution of Biphasic Life Cycles. Am Nat. 1999;154:306–20.

707 13. Destombe C, Godin J, Nocher M, Richerd S, Valero M. Differences in response between haploid and
708 diploid isomorphic phases of Gracilaria verrucosa (Rhodophyta: Gigartinales) exposed to artificial
709 environmental conditions. 1993.

710 14. Couceiro L, Le Gac M, Hunsperger HM, Mauger S, Destombe C, Cock JM, et al. Evolution and
711 maintenance of haploid-diploid life cycles in natural populations: The case of the marine brown alga
712 Ectocarpus. Evol Int J Org Evol. 2015;69:1808–22.

15. Immler S, Otto SP. The evolution of sex chromosomes in organisms with separate haploid sexes. Evolution. 2015;69:694–708.

16. Connallon T, Knowles LL. Intergenomic conflict revealed by patterns of sex-biased gene expression. Trends Genet TIG. 2005;21:495–9.

17. Parsch J, Ellegren H. The evolutionary causes and consequences of sex-biased gene expression. Nat Rev Genet. 2013;14:83–7.

18. Connallon T, Clark AG. The resolution of sexual antagonism by gene duplication. Genetics. 2011;187:919–37.

19. Safran RJ, Vortman Y, Jenkins BR, Hubbard JK, Wilkins MR, Bradley RJ, et al. The maintenance of phenotypic divergence through sexual selection: An experimental study in barn swallows Hirundo rustica. Evol Int J Org Evol. 2016;70:2074–84.

20. Harrison PW, Wright AE, Zimmer F, Dean R, Montgomery SH, Pointer MA, et al. Sexual selection drives evolution and rapid turnover of male gene expression. Proc Natl Acad Sci U S A. 2015;112:4393–8.

21. Cock JM, Godfroy O, Macaisne N, Peters AF, Coelho SM. Evolution and regulation of complex life cycles: a brown algal perspective. Curr Opin Plant Biol. 2014;17:1–6.

22. Silberfeld T, Leigh JW, Verbruggen H, Cruaud C, de Reviers B, Rousseau F. A multi-locus time-calibrated phylogeny of the brown algae (Heterokonta, Ochrophyta, Phaeophyceae): Investigating the evolutionary nature of the “brown algal crown radiation”. Mol Phylogenet Evol. 2010;56:659–74.

23. Cohen O, Pupko T. Inference of gain and loss events from phyletic patterns using stochastic mapping and maximum parsimony--a simulation study. Genome Biol Evol. 2011;3:1265–75.

24. Conesa A, Götz S. Blast2GO: A comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics. 2008;2008:619832.

25. Lipinska A, Cormier A, Luthringer R, Peters AF, Corre E, Gachon CMM, et al. Sexual dimorphism and the evolution of sex-biased gene expression in the brown alga ectocarpus. Mol Biol Evol. 2015;32:1581–97.

26. Nishitsuji K, Arimoto A, Iwai K, Sudo Y, Hisata K, Fujie M, et al. A draft genome of the brown alga, Cladosiphon okamuranus, S-strain: a platform for future studies of “mozuku” biology. DNA Res Int J Rapid Publ Rep Genes Genomes. 2016;23:561–70.

27. Szovenyi P, Rensing SA, Lang D, Wray GA, Shaw AJ. Generation-biased gene expression in a bryophyte model system. Mol Biol Evol. 2011;28:803–12.

28. Szovenyi P, Ricca M, Hock Z, Shaw JA, Shimizu KK, Wagner A. Selection is no more efficient in haploid than in diploid life stages of an angiosperm and a moss. Mol Biol Evol. 2013;30:1929–39.

29. Thornber CS. Functional properties of the isomorphic biphasic algal life cycle. Integr Comp Biol. 2006;46:605–14.

30. Clayton M. N. Evolution and Life Histories of Brown Algae. Bot Mar. 2009;31:379.

747 31. Bessho K, Iwasa Y. Optimal seasonal schedules and the relative dominance of heteromorphic and
748 isomorphic life cycles in macroalgae. J Theor Biol. 2010;267:201–12.

749 32. Luthringer R, Cormier A, Peters AF, Cock JM, Coelho SM. Sexual dimorphism in the brown algae.
750 Perspectives in Phycology. 2015;1:11–25.

751 33. Immler S, Arnqvist G, Otto SP. Ploidally antagonistic selection maintains stable genetic polymorphism.
752 Evol Int J Org Evol. 2012;66:55–65.

753 34. Godfroy O, Uji T, Nagasato C, Lipinska AP, Scornet D, Peters AF, et al. DISTAG/TBCCd1 Is Required for Basal
754 Cell Fate Determination in Ectocarpus. Plant Cell. 2017;

755 35. Macaisne N, Liu F, Scornet D, Peters AF, Lipinska A, Perrineau M-M, et al. The Ectocarpus IMMEDIATE
756 UPRIGHT gene encodes a member of a novel family of cysteine-rich proteins that have an unusual
757 distribution across the eukaryotes. Dev Camb Engl. 2017;

758 36. Dolan L. Body building on land — morphological evolution of land plants. Curr Opin Plant Biol.
759 2009;12:4–8.

760 37. Pires ND, Dolan L. Morphological evolution in land plants: new designs with old genes. Philos Trans R Soc
761 B Biol Sci. 2012;367:508–18.

762 38. Shaw AJ, Szovenyi P, Shaw B. Bryophyte diversity and evolution: windows into the early evolution of land
763 plants. Am J Bot. 2011;98:352–69.

764 39. Niklas KJ, Kutschera U. The evolution of the land plant life cycle. New Phytol. 2010;185:27–41.

765 40. Sano R, Juarez CM, Hass B, Sakakibara K, Ito M, Banks JA, et al. KNOX homeobox genes potentially have
766 similar function in both diploid unicellular and multicellular meristems, but not in haploid meristems. Evol
767 Dev. 2005;7:69–78.

768 41. Gossmann TI, Saleh D, Schmid MW, Spence MA, Schmid KJ. Transcriptomes of Plant Gametophytes Have
769 a Higher Proportion of Rapidly Evolving and Young Genes than Sporophytes. Mol Biol Evol. 2016;33:1669–
770 78.

771 42. Gerstein AC, Cleathero LA, Mandegar MA, Otto SP. Haploids adapt faster than diploids across a range of
772 environments. J Evol Biol. 2011;24:531–40.

773 43. Orr HA, Otto SP. Does Diploidy Increase the Rate of Adaptation? Genetics. 1994;136:1475–80.

774 44. Duret L, Mouchiroud D. Determinants of substitution rates in mammalian genes: expression pattern
775 affects selection intensity but not mutation rate. Mol Biol Evol. 2000;17:68–74.

776 45. Zhang L, Li W-H. Mammalian housekeeping genes evolve more slowly than tissue-specific genes. Mol
777 Biol Evol. 2004;21:236–9.

778 46. Slotte T, Bataillon T, Hansen TT, St Onge K, Wright SI, Schierup MH. Genomic determinants of protein
779 evolution and polymorphism in Arabidopsis. Genome Biol Evol. 2011;3:1210–9.

780 47. Krylov DM, Wolf YI, Rogozin IB, Koonin EV. Gene loss, protein sequence divergence, gene dispensability,
781 expression level, and interactivity are correlated in eukaryotic evolution. Genome Res. 2003;13:2229–35.

782 48. Cherry JL. Expression Level, Evolutionary Rate, and the Cost of Expression. Genome Biol Evol.
783 2010;2:757–69.

784 49. Pal C, Papp B, Lercher MJ. An integrated view of protein evolution. Nat Rev Genet. 2006;7:337–48.

785 50. Alvarez-Ponce D, Feyertag F, Chakraborty S. Position Matters: Network Centrality Considerably Impacts
786 Rates of Protein Evolution in the Human Protein-Protein Interaction Network. Genome Biol Evol.
787 2017;9:1742–56.

788 51. Hirsh AE, Fraser HB. Protein dispensability and rate of evolution. Nature. 2001;411:1046–9.

789 52. Alvarez-Ponce D, Fares MA. Evolutionary rate and duplicability in the Arabidopsis thaliana protein-protein
790 interaction network. Genome Biol Evol. 2012;4:1263–74.

791 53. Arunkumar R, Josephs EB, Williamson RJ, Wright SI. Pollen-Specific, but Not Sperm-Specific, Genes Show
792 Stronger Purifying Selection and Higher Rates of Positive Selection Than Sporophytic Genes in Capsella
793 grandiflora. Mol Biol Evol. 2013;30:2475–86.

794 54. Martínez EA, Santelices B. Selective mortality on haploid and diploid microscopic stages of Lessonia
795 nigrescens Bory (Phaeophyta, Laminariales). J Exp Mar Biol Ecol. 1998;229:219–39.

796 55. Parente MI, Neto AI, Fletcher RL. MORPHOLOGY AND LIFE HISTORY OF SCYTOSIPHON LOMENTARIA
797 (SCYTOSIPHONACEAE, PHAEOPHYCEAE) FROM THE AZORES1. J Phycol. 2003;39:353–9.

798 56. Kogame K. A taxonomic study of Japanese Scytosiphon (Scytosiphonales, Phaeophyceae), including two
799 new species. Phycol Res. 1998;46:39–56.

800 57. Tirosh I, Barkai N. Evolution of gene sequence and gene expression are not correlated in yeast. Trends
801 Genet. 2008;24:109–13.

802 58. Liao B-Y, Weng M-P, Zhang J. Contrasting genetic paths to morphological and physiological evolution.
803 Proc Natl Acad Sci U S A. 2010;107:7353–8.

804 59. Nuzhdin SV, Wayne ML, Harmon KL, McIntyre LM. Common pattern of evolution of gene expression level
805 and protein sequence in Drosophila. Mol Biol Evol. 2004;21:1308–17.

806 60. Khaitovich P, Hellmann I, Enard W, Nowick K, Leinweber M, Franz H, et al. Parallel patterns of evolution
807 in the genomes and transcriptomes of humans and chimpanzees. Science. 2005;309:1850–4.

808 61. Wray GA. The evolutionary significance of cis-regulatory mutations. Nat Rev Genet. 2007;8:206–16.

809 62. Lipinska AP, Toda NRT, Heesch S, Peters AF, Cock JM, Coelho SM. Multiple gene movements into and out
810 of haploid sex chromosomes. Genome Biol. 2017;18:104.

811 63. Coelho SM, Scornet D, Rousvoal S, Peters NT, Dartevelle L, Peters AF, et al. How to cultivate Ectocarpus.
812 Cold Spring Harb Protoc. 2012;2012:258–61.

64. Cormier A, Avia K, Sterck L, Derrien T, Wucher V, Andres G, et al. Re-annotation, improved large-scale assembly and establishment of a catalogue of noncoding loci for the genome of the model brown alga Ectocarpus. New Phytol. 2017;214:219–32.

65. Ye N, Zhang X, Miao M, Fan X, Zheng Y, Xu D, et al. Saccharina genomes provide novel insight into kelp biology. Nat Commun. 2015;6:6986.

66. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc. 2016;11:1650–67.

67. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinforma Oxf Engl. 2014;30:2114–20.

68. Cock JM, Sterck L, Rouzé P, Scornet D, Allen AE, Amoutzias G, et al. The Ectocarpus genome and the independent evolution of multicellularity in brown algae. Nature. 2010;465:617–21.

69. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinforma Oxf Engl. 2009;25:1105–11.

70. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.

71. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013;14:R36.

72. Radakovits R, Jinkerson RE, Fuerstenberg SI, Tae H, Settlage RE, Boore JL, et al. Draft genome sequence and genetic transformation of the oleaginous alga Nannochloropis gaditana. Nat Commun. 2012;3:686.

73. Gobler CJ, Berry DL, Dyhrman ST, Wilhelm SW, Salamov A, Lobanov AV, et al. Niche of harmful alga Aureococcus anophagefferens revealed through ecogenomics. Proc Natl Acad Sci U S A. 2011;108:4352–7.

74. Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, et al. The genome of the diatom Thalassiosira pseudonana: ecology, evolution, and metabolism. Science. 2004;306:79–86.

75. Zdobnov EM, Apweiler R. InterProScan--an integration platform for the signature-recognition methods in InterPro. Bioinforma Oxf Engl. 2001;17:847–8.

76. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. PLOS ONE. 2011;6:e21800.

77. Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol. 2000;302:205–17.

78. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 2000;17:540–52.

79. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 2006;34:W609-612.

80. Abascal F, Zardoya R, Telford MJ. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res. 2010;38:W7-13.

847 81. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007;24:1586–91.

848 82. Satapathy SS, Sahoo AK, Ray SK, Ghosh TC. Codon degeneracy and amino acid abundance influence the
849 measures of codon usage bias: improved Nc (Nc) and ENCprime (N(')c) measures. Genes Cells Devoted Mol
850 Cell Mech. 2017;22:277–83.

851 83. Pereira V, Waxman D, Eyre-Walker A. A problem with the correlation coefficient as a measure of gene
852 expression divergence. Genetics. 2009;183:1597–600.

853 84. Cohen O, Ashkenazy H, Belinky F, Huchon D, Pupko T. GLOOME: gain loss mapping engine. Bioinformatics.
854 2010;26:2914–5.

855 85. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol
856 Evol. 2012;3:217–23.

857

858