DR. JOHN W STILLER (Orcid ID : 0000-0002-0668-8243)

Article type : Regular Article

Evolution and expression of core SWI/SNF genes in red algae[1]

John W. Stiller[2]

Department of Biology, East Carolina University, Greenville, North Carolina 27858, USA

Chunlin Yang

Department of Pharmacology and Toxicology, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA

Jonas Collén, Nathalie Kowalczyk

Sorbonne Université, CNRS, Integrative Biology of Marine Models (LBI2M), Station Biologique de Roscoff (SBR), 29680 Roscoff, France

Beth E. Thompson

Department of Biology, East Carolina University, Greenville, North Carolina 27858, USA

[1] Received _____. Accepted _____

[2] Author for correspondence: e-mail stillerj@ecu.edu, phone: 252-328-2738, fax:

Running title: SWI/SNF genes in red algae

Editorial Responsibility: C. Lane (Associate Editor)

Abstract

Red algae are the oldest identifiable multicellular eukaryotes, with a fossil record dating back more than a billion years. During that time two major rhodophyte lineages, bangiophytes and florideophytes, have evolved varied levels of morphological complexity. These two groups are distinguished, in part, by different patterns of multicellular development, with florideophytes exhibiting a far greater diversity of morphologies. Interestingly, during their long evolutionary history, there is no record of a rhodophyte achieving the kinds of cellular and tissue-specific differentiation present in other multicellular algal lineages. To date, the genetic underpinnings of unique aspects of red algal development are largely unexplored; however, they must reflect the complements and patterns of expression of key regulatory genes. Here we report comparative evolutionary and gene expression analyses of core subunits of the SWI/SNF chromatin-remodeling complex, which is implicated in cell differentiation and developmental regulation in more well studied multicellular groups. Our results suggest that a single, canonical SWI/SNF complex was present in the rhodophyte ancestor, with gene duplications and evolutionary diversification of SWI/SNF subunits accompanying the evolution of multicellularity in the common ancestor of bangiophytes and florideophytes. Differences in how SWI/SNF chromatin remodeling evolved subsequently, in particular gene losses and more rapid divergence of SWI3 and SNF5 in bangiophytes, could help to explain why they exhibit a more limited range of morphological complexity than their florideophyte cousins.

Key index words: chromatin-remodeling; development; EST; evolution; qRT-PCR; red algae; RNA-Seq; SWI/SNF

Abbreviations: BLAST, Basic Local Alignment Search Tool; EST, Expressed Sequence Tag; qRT-PCR, qRT-PCR, Quantitative Reverse Transcription Polymerase Chain Reaction; RNA-Seq, Massively Parallel RNA Sequencing; SWI/SNF, SWItch/Sucrose NonFermentable

Red algae attained multicellularity over a billion years ago (Butterfield 2000, Gibson et al. 2017) and since have acquired substantial genetic, morphological and ecological diversity. In all that time, however, no member of the Rhodophyta has managed to attain the kinds of cellular and tissue-level differentiation present in most other eukaryotic lineages that have evolved large, multicellular forms (Murray and Dixon 1992, Brawley et al. 2017). The most developmentally complex florideophyte red algae have multiaxial filamentous construction, reinforced by secondary connections (pit plugs) between filamentous axes, forming pseudoparenchyma that mimics tissues and stabilizes leaf-, stem- and bulb-like structures (Coomans and Hommersand 1990). The reasons for this are unclear but, at least in part, must reflect developmental programing that is different in the Rhodophyta compared to other multicellular plants and algae (Stiller et al. 2012).

A previous examination of expressed sequence tags (ESTs) from *Porphyra umbilicalis* and *P. purpurea* recovered most major gene families important for achieving developmental complexity in model organisms (Stiller et al. 2012); however, it also highlighted potentially important variation in SWI/SNF (SWItch/Sucrose Non-Fermentable) proteins comprising a major chromatin remodeling complex (Winston and Carlson 1992) that is generally conserved across diverse eukaryotes. Chromatin remodeling controls accessibility of DNA to transcriptional machinery and is important for regulating cellular and developmental processes in animals, fungi and green plants (Jerzmanowski 2007, Ho and Crabtree 2010). In these groups several multi-subunit complexes are responsible for remodeling chromatin, each organized around a specific catalytic ATPase subunit (Clapier and Cairns 2009). The SWI/SNF complex, originally identified in yeast, was the first chromatin-remodeling

complex to be discovered, and is the one characterized most thoroughly. It has been shown to play key roles in multiple aspects of development in multicellular animals and land plants (Archacki et al. 2009, Clapier and Cairns 2009, Ho and Crabtree 2010); however, to our knowledge, evolutionary conservation and patterns of expression of SWI/SNF genes have not been characterized in red algae prior to this study.

Canonical SWI/SNF complexes contain four subunits broadly conserved across eukaryotic evolution: SNF2, SWI3, SNF5 and SWP73 (Jerzmanowski 2007, Clapier and Cairns 2009), and all four core SWI/SNF subunits were identified in the completed *Cyanidioschyzon merolae* genome (Stiller et al. 2012). *Porphyra* ESTs generated from blades grown under diverse conditions contained SNF2 and SWP73 sequences; however, BLAST similarity searches yielded no evidence for expression of either SWI3 or SNF5 (Stiller et al. 2012). This was intriguing because SWI3 and SNF5, along with SNF2, can form a minimal core complex in vitro with chromatin-remodeling activity comparable to the full SWI/SNF complex (Phelan et al. 1999, Sudarsanam and Winston 2000). Moreover, while SNF2 in *Porphyra purpurea* showed greater expression in the sporophyte than gametophyte, the reverse was true for SWP73 (Stiller et al. 2012).

To further elucidate the evolution and potential importance of SWI/SNF chromatin-remodeling proteins in rhodophytes, we undertook a broader investigation of red algal genomic and transcriptomic resources now available, along with experimental (qRT-PCR) analyses of SWI/SNF subunit expression in *Cyanidioschyzon merolae*.

Materials and Methods

Bioinformatic analyses of SWI/SNF subunits in red algae

Experimentally validated SWI/SNF protein sequences from human, budding yeast and *Arabidopsis* were used as queries in BLASTP and TBLASTN searches to recover homologous sequences from seven published red algal genomes (Table S1 in the Supporting Information). In an effort to expand comparative SWI/SNF analyses, particularly to additional florideophytes and members of the SRCP clade (Brawley et al. 2017), eight expansive EST data sets also were queried (Table S1). When no sequence with significant similarity to the three validated sequences was found, unvalidated but strongly conserved sequences from more closely related red algae also were used as BLAST queries to look for potentially more divergent homologs. All rhodophyte sequences recovered were used as reciprocal BLAST queries to verify they were most similar to their respective SWI/SNF homologs in all other organisms included in this study.

To verify a monophyletic association of rhodophyte homologs, and to examine relationships among rhodophyte paralogs when present, sequences of the four core SWI/SNF subunits were aligned using MUSCLE (Edgar 2004) in MEGA 7 (Kumar et al. 2016), then trimmed to include only highly conserved domains. Maximum-likelihood (ML) bootstrap (1000 iterations) phylogenetic analyses were carried out in PHYML 3.1 (Guindon et al. 2010) under the WAG model of amino acid substitution (based on Model Selection in MEGA 7.0) and an invariant+gamma estimate for rate variation among sites. Conserved functional protein domains of rhodophyte SWI3 sequences were predicted using Interpro (Mitchell et al. 2015) to further investigate the likely timing and subsequent evolution of the gene duplication that resulted in two paralogs in the florideophytes *Chondrus crispus*, *Gracilariopsis chorda,* and *Hildenbrandia rubra*. Inferred domains and their locations were visualized using Procite Image Creator in the ExPASy Bioinformatics Resource Portal (Sigrist et al. 2013).

SWI/SNF transcriptomics in multicellular rhodophytes.

RNA-seq data from the *Chondrus* and *Porphyra* genome projects (Collen et al. 2013, Brawley et al. 2017) were analyzed for expression of SWI/SNF genes from materials harvested under different developmental and environmental conditions. *Porphyra umbilicalis* RNA-seq reads generated from vegetative blade centers (NCBI, SRX115569), developing neutral spores (SRX115570) and mature spores (SRX115571) were examined. Individual reads were mapped to full-length SNF2, SWI3, SNF5 and SWP73 transcripts retrieved from the *P. umbilicalis* genome browser (*Phytozome*) using BLASTN searches. Stringent parameters (minimum word size of 64/expected threshold of 1) were used to prevent cross-mapping to paralogs with similar, conserved protein domains (this issue was limited to core domains of SNF2 paralogs), and total mapped reads were normalized to both overall transcript lengths and total number of reads in the targeted data set.

SWI/SNF gene expression in *Chondrus crispus* was analyzed from samples harvested during summer spring tides at 10 h (just after emersion), 13 h (at low tide), and 15 h (just before immersion) near Roscoff, France (48°43'32" N 3°58'11" W). At this site *Chondrus* grows in shallow rock pools or similar habitats, which means that the seaweeds are only lightly desiccated. In addition, for samples collected at 10 h, ESTs were recovered from thalli incubated in situ at doubled salinity by addition of NaCl (normal is 33), or in reduced light (40% of natural light filtered through a black mesh), or a combination of the two manipulations. Meteorological conditions were consistent during three days of sampling, with sunshine and occasional clouds, a light breeze, air temperatures around 19°C, and a water temperature of 17°C. Maximum irradiance was 1300 µmoles photons $\cdot$ m^{-2} $\cdot$ s^{-1}. Manipulations were performed with biological triplicates. RNA was extracted using the RNeasy plant kit (Qiagen, Germantown, MD, USA) according to manufacturer's protocol with the following modification: 100 mg of frozen tissue were ground in liquid nitrogen and resuspended in extraction buffer. EST data were generated, and analyzed as described in Lipinska et al. (2015).

qRT-PCR analyses of SWI/SNF in Cyanidioschyzon merolae.

Cyanidioschyzon cells (N-1804) were grown in bubbling culture in Allen medium 570 (Minoda et al. 2004) at 42ºC with 24 h of continuous light. Cultures were examined daily for optical density and harvested during exponential growth and after they had reached stationary growth. Harvested cells were collected by centrifugation at 3000*g* for 1-2 min at 4°C, and flash-frozen in liquid nitrogen.

RNA was extracted using Trizol (500 μL · 50 mg cells^{-1}) according to manufacturer's instructions (Invitrogen, Carlsbad, CA, USA), DNase treated and purified using the Qiagen RNase-Free DNase set and RNeasy MinElute Cleanup Kit. For each sample, 2 μg total RNA was used as template for reverse transcription using Invitrogen SuperScript III Kit with Oligo(dT) primers. cDNA corresponding to 50 ng of total RNA was used as template for quantitative PCR using SsoFast EvaGreen supermix (BioRad, Hercules, CA, USA). All qPCR reactions were performed in triplicate as a technical control. Efficiencies for each primer set (see Table S2 in the Supporting Information for primer sequences) were determined using a standard curve from a dilution series of genomic DNA; primer efficiencies were all between 90-110%. To measure relative RNA levels of subunits, we used standard curves to determine the amount of genomic DNA that corresponded to Ct values obtained from qPCR reactions using cDNA templates. Because gene copy number for each subunit should be the same in a given volume of genomic DNA, this allowed us to determine relative amounts of cDNA corresponding to each subunit in each sample.

Results

SWI/SNF subunits in red algae.

All seven red algal genomes examined in this study contain at least one homolog of each of the four major SWI/SNF gene families analyzed. Rhodophyte EST data sets (Table S1) interrogated generally were missing one or more SWI/SNF subunit, and/or had only short, fragmented sequences that could not be analyzed rigorously. In some cases, EST data contained potential contaminants that were closer matches to homologs from other eukaryotic lineages. The one exception was *Hildenbrandia rubra*,

where homologs were found for all four major SWI/SNF components, all with closest reciprocal matches to other rhodophytes. Because the significance of missing sequences could not be assessed, and divergent sequences could not be assigned reliably to a red algal genomic scaffold, other than *Hildenbrandia*, only data from published, annotated genomes were used for further computational and experimental analyses.

Bioinformatic analyses of red algal genomes confirmed previous findings based on *P. umbilicalis* ESTs for both SWP73 and SNF2 (Stiller et al. 2012). All seven rhodophyte genomes examined (plus *Hildenbrandia* ESTs) contain at least two SNF2 paralogs that trace to a duplication event in the ancestor of extant red algae (Fig. 1A). One of these copies was duplicated a second time in bangiophytes resulting in three SNF2 paralogs in *P. umbilicalis* and *Pyropia yezoensis* (Fig. 1A). SWP73 occurs as a single, conserved gene in all red algal genomes examined (Fig. 1B). Although SWI3 was not identified previously among *Porphrya* ESTs (Stiller et al. 2012), BLAST searches of both gene models and extensive EST data from the *P. umbilicalis* genome (Brawley et al. 2017) returned one sequence with significant similarity to yeast, human, *Arabidopsis* and *Cyandioschyzon* SWI3 homologs (Fig. 1C). Once this sequence was identified in *Porphyra*, a reexamination of the original EST dataset (Chan et al. 2012) found evidence for at least weak SWI3 expression there as well. A fragment with high similarity to the *Porphyra* sequence also was found on a short contig from the *Pyropia* genome (Fig. 1C). SWI3 occurs as a single copy in the red algae we examined, except in *Chondrus, Gracilariopsis and Hildenbrandia* where two relatively diverged paralogs are present (Fig. 1C). One of these paralogs, which we designate SWI3A (Fig. 1C), is more conserved with respect to *Arabidopsis*, yeast, human, and most other rhodophyte homologs (Table S3); however, sequences from the two bangiophytes, *Porphyra* and *Pyropia*, are more similar to SWI3B from *Chondrus* than to SWI3A (4e-61 versus 7e-27, respectively, for the full *Porphyra* sequence; Table S3 in the Supporting Information), and clustered with the florideophyte SWI3B paralogs in phylogenetic analyses (Fig. 1C).

When only rhodophyte sequences were included, phylogenetic analysis of SWI3 grouped florideophyte SWI3B paralogs with SWI3 from *Porphyra* (only a short fragment of the gene was recovered from *Pyropia*) with somewhat stronger bootstrap support (Fig. S1 in the Supporting Information). Moreover, this group of SWI3 homologs shares several structural features relative to other rhodophyte SWI3 sequences. The generally conserved SMARCC C-terminal domain is not recognized in any of them, and BRCT motifs are found more distally within the SMARCC N-terminal domain than in other SWI3 sequences (Fig. S1). Florideophyte SWI3B and *Porphyra* SWI3 tend to be more diverged with respect to other eukaryotic and rhodophyte homologs (Fig. S2 in the Supporting Information), making it plausible that their apparent phylogenetic affinity and common domain loss reflect parallel rather than shared evolutionary histories. Nevertheless, current evidence suggests that bangiophyte SWI3 is orthologous with SWI3B from florideophytes, meaning the gene duplication occurred in the common ancestor of the Bangiophyceae and Florideophyceae, and that SWI3A was lost from bangiophytes.

As reported previously for EST data (Stiller et al. 2012), no SNF5 homolog was found in the complete *Porphyra* genome searching with human, yeast, *Arabidopsis* or *Cyanidioschyzon* sequences. With the more closely related *Chondrus* genome available (Collen et al. 2013), analyses of SNF5 in *Porphyra* and other multicellular red algae proved more complicated. Two SNF5 sequences were identified in *Chondrus*; they are reciprocal best matches in BLAST searches (3e-12) and no other similar sequences were found in the *Chondrus* genome. One (SNF5A; XP_005711230) is much more strongly conserved with respect to validated eukaryotic homologs (8e-33 to human SNF5, 4e-30 to *C. merolae*) than the other (SNF5B; XP_005711230, 1e-05 to human SNF5, no significant match at < 1e-05 in *C. merolae*). Nevertheless, a SNF5 domain is detected in the more divergent *Chondrus* sequence in NCBI BLASTP searches, and the sequences corresponding to SNF5 homologues are the only matches with even marginal significance (from 6e-06 to 1e-04). Both *Chondrus* SNF5 paralogs also were found in the recently published genome from a second florideophyte, *Gracilariopsis forma* (Lee et al. 2018). A single copy of SNF5 was found in *Hildenbrandia* (Fig. 1D), although a second copy could be present in the genome but not represented in available EST data.

A BLASTP search querying with *Chondrus* SNF5A found no significant match in *Porphyra*, although weak similarity (E = 0.01) was found to one gene model (OSX68760). Interestingly, this *Porphyra* sequence is more similar to the second, less conserved florideophyte SNF5B paralog (1e-08), a relationship supported by phylogenetic analysis of SNF5 sequences (Fig. 1D). As with SWI3, once this putative SNF5 sequence was identified in *Porphyra,* expression was confirmed in both EST (Chan et al. 2012) and RNA-seq (Brawley et al. 2017) data. Although no gene model has been annotated from the *Pyropia* genome (Nakamura et al. 2013), two small genomic segments (contigs 25786 and 39096) contain open reading frames with significant similarity to the *P. umbilicalis* SNF5 sequence. Phylogenetic analysis confirmed the relationship between *Porphyra* and *Pyropia* sequences (Fig. 1D), and BLAST searches against *Pyropia* ESTs (SRX437137) verified mRNA expression.

Relative expression levels of SWI/SNF subunits.

RNA-seq data from the *P. umbilicalis* genome project indicate that all SWI/SNF genes are expressed, but do not suggest coordinated functioning in a SWI/SNF complex. All three SNF2 paralogs are comparably expressed under all developmental conditions; interestingly they all show some increased expression in spore-differentiating regions of the blade (Fig. S3 in the Supporting Information). SWI3 has much lower relative expression in vegetative cells, but a proportionally larger increase in differentiating cells. Although SNF5 and SWP73 show a similar pattern to SWI3, they are expressed at much lower levels under all conditions. In vegetative cells, SNF5 expression is 75 to 100 times lower than the three SNF2 genes (Fig. S3). Unless there are dramatic differences in mRNA translation rates, or in stability of these proteins, relative expression levels cast doubt on the stoichiometry required for cooperative action in a canonical SWI/SNF complex in *Porphyra*.

In *Chondrus*, extensive RNA-seq data generated from blades experiencing different abiotic stresses indicate less overall variation in relative expression among the various SWI/SNF subunits, consistent with their functioning in a coordinated manner. Moreover, there is a suggestion that the

different SWI3 paralogs could function independently in different SWI/SNF complexes. Among 18 RNA-seq replicates across differing light and salinity conditions, only SNF2A and SWI3A mRNA levels varied significantly (sometimes at a two-fold difference relative to the other SWI/SNF subunits, and/or a four-fold difference relative to the average across all data points; Fig. S4 in the Supporting Information). Moreover, expression of these two genes generally are correlated across *Chondrus* RNA-seq data, but not necessarily with genes encoding the other core SWI/SNF components (see Fig. S4, 15h samples). This observation suggests that SNF2A and SWI3A might function together in an alternative chromatin remodeling complex. Curiously, no evidence for expression of SWP73 was found in RNA-seq data under any conditions examined, although expression of the *Chondrus* SWP73 gene model was supported by pooled EST data generated for the full genome project (Collen et al. 2013).

Given the ambiguous overall results from multicellular rhodophytes, we undertook a direct, experimental analysis of expression of SWI/SNF subunits in the unicellular species *C. merolae,* which has proven a viable model for empirical investigation of transcriptional mechanisms in rhodophytes (Yang et al. 2014). qRT-PCR confirmed that all four conserved SWI/SNF subunits in *C. merolae* are expressed as mRNAs (Fig. 2). There was some variation in expression levels among different SWI/SNF components; for example, SNF2-2 was expressed five to eight-fold higher than SNF5 across all samples (Fig. 2). Nevertheless, relative expression levels of the four core subunits appear to vary in the same directions and patterns across samples and replicates. Although *Cyanidioschyzon* provides little opportunity to examine how expression may vary at different developmental stages, relative transcript levels of all four SWI/SNF genes are comparable during exponential growth and at stationarity (Fig. 2).

Discussion

The *Cyanidioschyzon, Galdieria,* and *Porphyridium* genomes all contain evolutionarily conserved genes for the four core SWI/SNF subunits, and these unicellular forms generally are recovered as basal and/or sister to the a bangiophyte/florideophyte clade in phylogenetic analyses (Yoon et al. 2006, Qui et al. 2015, Yang et al. 2016, Brawley et al. 2017). qRT-PCR results from *Cyanidioschyzon* provide no evidence to reject a null model that SWI/SNF interact in a typical chromatin-remodeling complex in both growing and stationary conditions. Given the relative conservation of SWI/SNF subunits in unicellular taxa, and quantitative expression results in *C. merolae*, it is reasonable to conclude that a typical SWI/SNF complex was present in the ancestor of red algae, and likely continues to function in extant unicellular forms.

The situation appears less straightforward in multicellular rhodophytes. Here we propose a model for SWI/SNF evolution in red algae (Fig. 3), based upon evolutionary relationships and expression patterns of SWI/SNF genes from seven published rhodophyte genomes. As noted above, most red algal EST data are missing otherwise conserved SWI/SNF genes and, therefore, could not be used reliably in a thorough analysis of SWI/SNF gene evolution. It should be noted, however, that no partial results from BLAST searches of rhodophyte ESTs contradict our overall the model derived from genomic comparisons.

The presence of two very different copies of both SWI3 and SNF5 in *Chondrus and Gracilariopsis,* and particularly the highly divergent nature of one of the SNF5 paralogs, suggest that chromatin remodeling is more complex in the Florideophyceae than in unicellular rhodophytes. This is to be expected given the great morphological and developmental diversity that has evolved in florideophytes (Yang et al. 2016). What is less clear is whether and how the various SWI/SNF components interact. Their apparently correlated expression across diverse conditions in *Chondrus* suggests that SNF2A and SWI3A could work together in a complex, presumably along with the more conserved of the SNF5 paralogs. These three components from humans can achieve chromatin remodeling at rates comparable to the full SWI/SNF complex in vitro (Phelan et al. 1999, Sudarsanam

and Winston 2000). On the other hand, transcriptomic results suggest that SWP73 might participate in chromatin remodeling only under limited conditions in *Chondrus*. In well-studied systems, there is evidence that SWI/SNF complexes of different composition can form in vivo. For example, mutations in different subunits of the SWI/SNF complex contribute to cancer in tissue-specific ways in mammals (Kadoch and Crabtree 2015). In yeast, SNF5 can contribute to a SWI/SNF regulatory submodule, and the presence or absence of this module alters patterns of gene expression (Sen et al. 2017).

Whether a core SWI/SNF complex can form in members of the Bangiophyceae is questionable. The only SNF5 sequence present in *Porphyra* and *Pyropia* is highly diverged from experimentally validated sequences in diverse eukaryotes, and from putative SNF5 homologs in other red algae. Moreover, it is expressed at extremely low levels compared to SNF2 core ATPases, with which it would interact in a SWI/SNF complex. This makes it unclear whether these sequences function as SNF5 proteins at all and, if so, whether they can interact physically with the other SWI/SNF core proteins that are relatively conserved in these bangiophytes. Although there is no evidence for major compensatory substitutions to account for the dramatic changes to SNF5 primary structure, the SWI3 homolog in *Porphyra* is more divergent than those of most other reds (Fig. S2). Interestingly, apparent orthologs of these divergent bangiophyte SWI3 and SNF5 sequences also are present in *Chondrus* and *Gracilariopsis*, along with a second, more strongly conserved copy in each case. This suggests that florideophytes could retain a typical SWI/SNF complex, and that the more divergent SWI3 and SNF5 sequences have evolved independent function(s), perhaps in an alternative SWI/SNF complex. More specialized complexes, active only under certain conditions, could help to explain the lack of evidence for consistent expression of SWP73 in *Chondrus*, as well as the very low expression of SNF5 and SWP73 (and SWI3, to a lesser degree) in normal, vegetative blades in *Porphyra*. If the ancestor of bangiophytes and florideophytes had evolved two different SWI/SNF complexes, it appears that only the more diverged version was retained in the bangiophytes (Fig. 3), and it may be important for chromatin-remodeling only at certain developmental stages.

In addition to missing strongly conserved SWI3 and SNF5 homologs, bangiophytes are

unique among red algae in having a duplication of a SNF2 paralog from this subfamily (Fig. 1A; Stiller et al. 2012). Moreover, all three SNF2 paralogs in *Porphyra* are expressed at much higher levels than other SWI/SNF proteins (Fig. S3). It is possible that these observations are related. SNF2 ATPase/helicase subunits form the active core of SWI/SNF complexes but also can alter chromatin structure and DNA protein interactions on their own (Phelan et al. 1999, Flaus et al. 2006). Interestingly, all three *Porphyra* SNF2 proteins are larger than other red algal homologs, and the extended sequences contain recognizable functional domains in all cases. Such major structural changes could block interactions with other SWI/SNF proteins, promote interactions with alternative proteins, and/or provide additional functions that enhance the ability of SNF2 to remodel chromatin on its own. Thus, both the additional copy and unusual sequence elements of *Porphyra* SNF2 proteins could compensate functionally for a diminished role of the canonical SWI/SNF complex.

Based on comparative genomics, it has been proposed that the ancestor of red algae went through an early genetic bottleneck that resulted in a compacted genome and multiple gene losses (Bhattacharya et al. 2013, Collen et al. 2013, Qiu et al. 2015). Whereas most major developmental regulators are present in rhodophytes, such as MADS-box and homeodomain gene families (Stiller et al. 2012), notably reduced cytoskeleton-related functions has been suggested as a general constraint on red algal cell size and morphological complexity (Brawley et al. 2017). After diverging from florideophytes, bangiophytes also appear to have lost SWI3 and SNF5 genes, which could account for a more rapid divergence of the otherwise conserved SWI/SNF complex. This also could have spurred the subsequent evolution of novel chromatin-remodeling mechanisms, as reflected by the third copy of the SNF2 helicase present only in bangiophytes. Florideophytes diversified broadly whereas bangiophytes remained relatively static in morphology and life history over more than a billion years of evolution (Sutherland et al. 2011). Although a complex set of factors are likely responsible for restricting bangiophyte evolutionary diversification, the differential evolution of chromatin remodeling machinery could help to explain this difference in the two major multicellular lineages of the Rhodophyta.

Acknowledgements

Porphyra genome-level analyses were supported by the U.S. Department of Energy Joint Genome Institute under Contract No. DE-AC02-05CH11231 and NSF Research Collaboration Network grant 0741907.

Literature cited

Archacki, R., Sarnowski, T. J., Halibart-Puzio, J., Brzeska, K., Buszewicz, D., Prymakowska-Bosak, M., Koncz, C., & Jerzmanowski, A. 2009. Genetic analysis of functional redundancy of BRM ATPase and ATSWI3C subunits of *Arabidopsis* SWI/SNF chromatin remodeling complexes. *Planta* 229: 1281–92.

Bhattacharya, D., Price, D. C., Chan, C. X., Qiu, H., Rose, N., Ball, S., Weber, A. P. M., Arias, M. C., Henrissat, B., Coutinho, P. M., Krishnan, A., Zäuner, S., Morath, S., Hilliou, F., Egizi, A., Perrineau, M. M. & Yoon, H. S. 2013. Genome of the red alga *Porphyridium purpureum*. *Nat. Commun.* 4:1941.

Brawley, S. H., Blouin, N. A., Ficko-Blean, E., Wheeler, G. L., Lohr, M., Goodson, H. V., Jenkins, J. W., Blaby-Haas, C. E., Helliwell, K. E., Chan, C. X., Marriage, T. N., Bhattacharya, D., Klein, A. S., Badis, Y., Brodie, J., Cao, Y., Collén, J., Dittami, S. M., Gachon, C. M. M., Green, B. R., Karpowicz, S. J., Kim, J. W., Kudahl, U. J., Lin, S., Michel, M., Mittag, M., Olson, B. J. S. C., Pangilinan, J. L., Peng, Y., Qiu, H., Shu, S., Singer, J. T., Smith, A. G., Sprecher, B. N., Wagner, V., Wang, W., Wang, Z. Y, Yan, J., Yarish, C., Zäuner-Riek, S., Zhuang, Y., Zou, Y., Lindquist, E. A., Grimwood, J., Barry, K. W., Rokhsar, D. S., Schmutz, J., Stiller, J. W., Grossman, A. R., & Prochnik, S. E. 2017. Insights into the red algae and eukaryotic evolution from the genome of *Porphyra umbilicalis* (Bangiophyceae, Rhodophyta).” *Proc. Natl. Acad. Sci. USA* 114:E6361-70.

Butterfield, N. J. 2000. "*Bangiomorpha pubescens n. gen., n. sp.*: Implications for the evolution of sex, multicellularity, and the mesoproterozoic/neoproterozoic radiation of eukaryotes." *Paleobiol.* 26: 386–404.

Chan, C. X., Blouin, N. A., Zhuang, Y., Zäuner, S., Prochnik, S. E., Lindquist, E., Lin, S., Benning, C., Lohr , M., Yarish, C., Gantt, E., Grossman, A. R., Lu, S., Müller, K., Stiller, J. W., Brawley, S. W., Bhattacharya, D. 2012. *Porphyra* (Bangiophyceae) transcriptomes provide insights into red algal development and metabolism. *J. Phycol.* 48:1328–42.

Clapier, C. R., & Cairns, B. R. 2009. The biology of chromatin remodeling complexes. *Ann. Rev. Biochem.* 78:273–304.

Collen, J., Porcel, B., W. Carre, S. G. Ball, C. Chaparro, T. Tonon, Barbeyron, T., Michel, G., Noel, B., Valentin, K., Elias, M., Artiguenave, F., Arun, A., Aury, J-M., Barbosa-Neto, J. F., Bothwell, J. H., Bouget, F. Y., Brillet, L., Cabello-Hurtado, F., Capella-Gutiérrez, S., Charrier, B., Cladière, L., Cock, J. M., Coelho, S. M., Colleoni, C., Czjzek, M., Da Silva, C., Delage, L., Denoeud, F., Deschamps, P., Dittami, S. M., Gabaldón, T., Gachon, C. M. M., Groisillier, A., Hervé, C., Jabbari, K., Katinka, M., Kloareg, B., Kowalczyk, N., Labadie, K., Leblanc, C., Lopez, P. J., McLachlan, D. H., Meslet-Cladiere, L., Moustafa, A., Nehr, Z., Collén, P. N., Panaud, O., Partensky, F., Poulain, J., Rensing, S. A., Rousvoal, S., Samson, G., Symeonidi, A., Weissenbach, J., Zambounis, A., Wincker, P. & Boyen, B. 2013. Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida. In *Proc. Natl. Acad. Sci. USA* 110:5247–52.

Coomans, R. J., & Hommersand, M. H. 1990. Vegetative growth and organization. *In* Cole, K.M., & Sheath, R. G. [Eds.] *Biology of Red Algae*, Cambridge University Press, NY, USA, pp. 275-304.

Edgar, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput." *Nucl. Acids Res.* 32:1792–97.

Flaus, A., Martin, D. M. A., Barton, G. J. & Owen-Hughes, T. 2006. Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucl. Acids Res.* 34:2887–2905.

Gibson, T. M., Shih, P. M., Cumming, V. M., Fischer, W. M., Crockford, P. W., Hodgskiss, M. S. W., Wörndle, S., Creaser, R. A., Rainbird, R. H., Skulski, T. M. & Halverson, G. P. 2017. Precise age of *Bangiomorpha pubescens* dates the origin of eukaryotic photosynthesis. *Geology* 46:135-38.

Guindon, S., Dufayard, J-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* 59:307–21.

Ho, L., & Crabtree, G. R. 2010. Chromatin remodelling during development. *Nature* 463:474–84.

Jerzmanowski, A. 2007. SWI/SNF chromatin remodeling and linker histones in plants. *Biochim. Biophys. Acta* 1769: 330–45.

Kadoch, C., and Crabtree, G. R. 2015. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci. Adv.* 1: e1500447.

Kumar, S., Stecher, G. & Tamura, K. 2016. MEGA7: Molecular evolutionary genetics analysis, version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33: 1870-74.

Lee, J., Yang, E. C., Graf, L., Yang, J. H., Qiu, H., Zelzion, U., Chan, C. X., Stephens, T. G., Weber, A. P. M., Boo, G. A., Boo, S. M., Kim, K. M., Shin, Y., Jung, M., Lee, S. J., Yim, H. S., Lee, J. H., Bhattacharya, D. & Yoon, H. S. 2018. Analysis of the draft genome of the red seaweed *Gracilariopsis chorda* provides insights into genome size evolution in rhodophyta. *Mol. Biol. Evol.* 35: 1869-86.

Lipinska, A., Cormier, A., Luthringer, R., Peters, A. F., Corre, E., Gachon, C. M., Cock J. M. & Coelho, S. M. 2015. Sexual dimorphism and the evolution of sex-biased gene expression in the brown alga Ectocarpus. *Mol. Biol. Evol.* 32:1581-97.

Minoda, A., Sakagami, R., Yagisawa, F. Kuroiwa, T. & Tanaka, K. 2004. Improvement of culture conditions and evidence for nuclear transformation by homologous recombination in a red alga, *Cyanidioschyzon merolae* 10D. *Plant Cell Physiol.* 45:667–71.

Mitchell, A., Chang, H. Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., McAnulla, C., McMenamin, C., Nuka, G., Sangrador-Vegas, S. P. A., Scheremetjew, M., Rato, C., Yong, S. Y., Bateman, A., Punta, M., Attwood, T. K., Sigrist, C. J. A., Redaschi, N., Rivoire, C., Xenarios, I., Kahn, D., Guyot, D., Bork, P., Letunic, I., Gough, J., Oates, M., Haft, D., Huang, H., Natale, D. A., Wu, C. H., Orengo, C., Sillitoe, I., Huaiyu Mi, H., Thomas, P. D. & Finn, R. D. 2015. The InterPro protein families database: The classification resource after 15 Years. *Nucl. Acids Res.* 43:D213-221.

Murray, S. N., & Dixon, P. S. 1992. The Rhodophyta: some aspects of their biology III. *Oceanogr Mar. Biol. Ann. Rev.* 30:1-148.

Nakamura, Y., Sasaki, N., Kobayashi, M., Ojima, N., Yasuike, M., Shigenobu, Y., Satomi, M., Fukuma, Y., Shiwaku, K., Tsujimoto, A., Kobayashi, T., Nakayama, I., Ito, F., Nakajima, K., Sano, M., Wada, T., Kuhara, S., Inouye, K., Gojobori, T. & Ikeo, K. 2013. The first symbiont-free genome sequence of marine red alga, Susabi-Nori (*Pyropia yezoensis*). *PLoS ONE* 8:e57122.

Phelan, M. L., Sif, S., Narlikar, G. J. & Kingston, R. E. 1999. Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Mol. Cell.* 3:247–53.

Qiu, H., Price, D. C., Yang, E. C., Yoon, H. S. & Bhattacharya, D. 2015. Evidence of ancient genome reduction in red algae (Rhodophyta). *J. Phycol.* 51:624–36.

Sen, P., Luo, J., Hada, A., Hailu, S. G., Dechassa, M. K., Persinger, J., Brahma, S., Paul, S., Ranish, J. & Bartholomew, B. 2017. Loss of Snf5 induces formation of an aberrant SWI/SNF complex. *Cell Rep.* 18:2135–47.

Sigrist, C. J., de Castro, E. Cerutti, L. Cuche, B. A., Hulo, N., Bridge, A., Bougueleret, L.& Xenarios, I. 2013. New and continuing developments at PROSITE. *Nucl. Acids Res.* 41:D344-347.

Stiller, J. W., Perry, J., Rymarquis, L. A., Accerbi, M., Green, P. J., Prochnik, S., Lindquist, E., Chan, C. X., Yarish, C., Lin, S., Zhuang, Y., Blouin, N. A. & Brawley, S. H. 2012. Major developmental regulators and their expression in two closely related species of *Porphyra* (Rhodophyta). *J. Phycol.* 48:883–96.

Sudarsanam, P., & Winston, F. 2000. The Swi/Snf family - nucleosome-remodeling complexes and transcriptional control. *Trends Genet.* 16:345–51.

Sutherland, J. E., Lindstrom, S. C., Nelson, W. A., Brodie, J., Lynch, M. D. J., Hwang, M. S., Choi, H. G., Miyata, M., Kikuchi, N., Oliveira, M. C., Farr, T., Neefus, C., Mols-Mortensen, A., Milstein, D. & Müller, K. M. 2011. A new look at an ancient order: Generic revision of the Bangiales (Rhodophyta). *J. Phycol.* 47:1131-51.

Winston, F., & Carlson, M. 1992. Yeast SNF SWI transcriptional activators and the SPT SIN chromatin connection. *Trends Genet.* 8:387-91.

Yang, C., Hager, P. W. & Stiller, J. W. 2014. The identification of putative RNA polymerase II C-terminal domain associated proteins in red and green algae. *Transcription* 5:e970944.

Yang, E. C., Boo, S. M., Bhattacharya, D., Saunders, G. W., Knoll, A. H., Fredericq, S., Graf, L. & Yoon, H. S. 2016. Divergence time estimates and the evolution of major lineages in the florideophyte red algae. *Sci. Rep.* 6:21361.

Yoon, H. S., Müller, K. M., Sheath, R. G., Ott, F. D. & Bhattacharya, D. 2006. Defining the major lineages of red algae (Rhodophyta). *J. Phycol.* 42: 482-92.

Figure Legends

Figure 1. Bootstrapped maximum-likelihood trees of red algal SWI/SNF core sequences and diverse eukaryotic outgroups. Only bootstrap values ≥ 50% are included on the trees. Starred sequences from *Cyanidioschyzon* are those included in qRT-PCR analyses. **A.** SNF2 subfamily of SNF2 ATPase helicases, highlighting a conserved, ancestral gene duplication in red algae, and a second duplication unique to bangiophytes. **B.** SWP73 tree of the single sequences present in each red algal genome. **C.** SWI3 tree highlighting a gene duplication, thus far apparent only in florideophytes, but that likely occurred in the common ancestor of florideophytes and bangiophytes. **D**. SNF5 tree showing another gene duplication, apparently in the ancestor of the bangiophyte/florideophyte clade, followed by loss of the more conserved copy of SNF5 from bangiophytes.

Figure 2. Expression of core SWI/SNF subunits from *Cyanidioschyzon* based on qRT-PCR. Expression is graphed relative to the general transcription factor, TATA-binding protein (TBP). Each bar represents the mean expression based on triplicate qPCR reactions from a single, independent mRNA extraction. Shades of bars (see inset key) indicate replicated mRNA extractions from material harvested during exponential growth and from stationary cultures. SNF2-1: (XP_005536139), SNF2-2 (XP_005537173), SNF5 (XP_005539508), SWI3 (XP_005536719), SWP73 (XP_005539221).

Figure 3. Model of the evolution of SWI/SNF chromatin remodeling in red algae. Solid lines between SWI/SNF subunits suggest likely interactions based on experimentally validated interactions in green plants, animals and yeast, combined with observations of evolutionary conservation and relative expression levels in *Cyanidioschyzon*, *Chondrus* and *Porphyra* (see discussion for details). Dotted lines indicate possible interactions that lack any observational support in our analyses. Solid curved lines associated with SWI3 proteins reflect its dimerization in canonical SWI/SNF complexes, an interaction presumably conserved in unicellular reds. If functional SWI/SNF complexes do, in fact form, in *Chondrus*, the significant divergence between SWI3A and SWI3B suggest that each likely

form a homodimer in distinct SWI/SNF complexes rather than SWI3A-SWI3B heterodimers in the same complex. Tree topology is based upon Qiu et al. (2015) and Brawley et al. (2017), which indicate the major rhodophyte multicellular lineages, the Bangiophyceae and Florideophyceae, are descended from a common ancestor.

Figure S1. Phylogenetic relationships and domain architecture of red algal SWI3 sequences. ML bootstrap values (1000) iterations are shown on respective tree nodes. All rhodophyte sequences contain the core RSC8 domain conserved across SWI3 proteins, and phylogenetic analysis employed only the most conserved RSC8 regions; however, florideophyte SWI3B and *Porphyra* SWI3, which group together with moderate bootstrap support, all are missing a C-terminal SmarccC domain otherwise conserved in all red algal SWI3 genes. Florideophyte SWI3B sequences also share more distally situated BRCT motifs in their N-terminal Smarcc domains. The SmarccN domain is missing entirely from *Porphyra* SWI3.

Figure S2. Alignment of conserved regions of red algal SWI3 sequences used in phylogenetic analyses that yielded the tree shown in Figure S1.

Figure S3. Comparative expression of *Porphyra* SWI/SNF genes based on number of RNA-seq reads generated from vegetative blade centers (NCBI, SRX115569), developing neutral spores (SRX115570) and mature spores (SRX115571) and mapped by BLASTN search to full-length gene model transcripts found in the complete *Porphrya* genome. Expression levels were normalized to full transcript lengths and the total number of reads present in each SRX accession. Expression is graphed relative to the mean expression (1.0) of mRNA for TATA-binding protein, which is expected to be present for the transcription of most if not all protein-encoding genes.

Figure S4. Expression patterns of *Chondrus crispus* SWI/SNF sequences from thalli harvested at different temporal and environmental under conditions. 10 h, 13 h, 15 h are the time of day algae were harvested in the natural environment, L represents a reduction of 60% of total daylight, S is double normal seawater salinity, and LS is a combination of the two environmental factors. The three physiological treatments were performed during two hours before harvest at 13 h. Data are plotted on a log2 scale (1=2X, 2=4X compared to average of all data points) with three replicates for each condition.

Table S1. Information on genomes and EST data examined.

Table S2. Primers used in qRT-PCR of *Cyanidioschyzon* sequences.

Table S3. Comparative similarities of SWI3 sequences in *Chondrus* and *Porphyra* to each other, and to homologs from other red algae and model eukaryotes.