

A Haploid System of Sex Determination in the Brown Alga Ectocarpus sp.

Sophia S. Ahmed, J. Mark Cock, Eugenie Pessia, Remy Luthringer, Alexandre Cormier, Marine Robuchon, Lieven Sterck, Akira F Peters, Simon M. Dittami, Erwan Corre, et al.

▶ **To cite this version:**

Sophia S. Ahmed, J. Mark Cock, Eugenie Pessia, Remy Luthringer, Alexandre Cormier, et al.. A Haploid System of Sex Determination in the Brown Alga Ectocarpus sp.. Current Biology - CB, 2014, 24 (17), pp.1945-1957. 10.1016/j.cub.2014.07.042 . hal-01132642

HAL Id: hal-01132642

https://hal.sorbonne-universite.fr/hal-01132642

Submitted on 18 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Title: A haploid system of sex determination in the brown alga *Ectocarpus* sp.

Authors: Ahmed S[1,2§], Cock JM[1§], Pessia E[3§], Luthringer R[1], Cormier A[1], Robuchon M[1,8], Sterck L[4], Peters AF[5], Dittami SM[1], Corre E[7], Valero M[8], Aury J-M[6], Roze D[8], Van de Peer Y[4,9], Bothwell J[2], Marais GAB[3], Coelho SM[1*]

Affiliations: [1]UPMC Univ Paris 6, CNRS, UMR 7139 'Marine Plants and Biomolecules', Station Biologique de Roscoff, CS 90074 29688 Roscoff, France; [2]Queens University Belfast, Medical Biology Centre, Belfast BT9 7BL, Northern Ireland, UK; [3]Université Lyon 1, Centre National de la Recherche Scientifique, UMR 5558, Laboratoire de Biométrie et Biologie Évolutive, 69622 Villeurbanne, France; [4]Department of Plant Systems Biology, VIB, Technologiepark 927, B-9052 Gent, Belgium; [5]Bezhin Rosko, 29250 Santec, France; [6]Commissariat à l'Energie Atomique (CEA), Institut de Génomique (IG), Genoscope, Evry, France; [7]Abims Platform, FR2424, Station Biologique de Roscoff, CS 90074, 29688 Roscoff, France; [8]UMI 3604, Evolutionary Biology and Ecology of Algae, CNRS, Sorbonne Universités, UPMC, PUCCh, UACH, Station Biologique de Roscoff, CS 90074, 29688 Roscoff, France.[9]Genomics Research Institute, University of Pretoria, Hatfield Campus, Pretoria 0028, South Africa.

*Correspondance: coelho@sb-roscoff.fr

[§]Equal contribution

Running title: Sex chromosome evolution in a brown alga

SUMMARY

Background: A common feature of most genetic sex-determination systems studied so far is that sex is determined by non-recombining genomic regions, which can be of various sizes depending on the species. These regions have evolved independently and repeatedly across diverse groups. A number of such sex-determining regions (SDR) have been studied in animals, plants and fungi, but very little is known about the evolution of sexes in other eukaryotic lineages.

Results: We report here the sequencing and genomic analysis of the SDR of *Ectocarpus*, a brown alga that has been evolving independently from plants, animals and fungi for over a billion years. In *Ectocarpus,* sex is expressed during the haploid phase of the life cycle, and both the female (U) and the male (V) sex chromosomes contain non-recombining regions. The U and V of this species have been diverging for more than 70 My, yet gene degeneration has been modest and the SDR is relatively small with no evidence for evolutionary strata. These features may be explained by the occurrence of strong purifying selection during the haploid phase of the life cycle and the low level of

sexual dimorphism. V was dominant over U, suggesting that femaleness may be the default state, adopted when the male haplotype is absent.

Conclusions: The *Ectocarpus* UV system has clearly had a distinct evolutionary trajectory not only to the well-studied XY and ZW systems, but also to the UV systems described so far. Nonetheless, some striking similarities exist, indicating remarkable universality of the underlying processes shaping sex chromosome evolution across distant lineages.

HIGHLIGHTS

- *Ectocarpus* U and V sex chromosomes evolved more than 70 MY ago
- The non-recombining region in the U and V is small and degeneration has been modest
- U and V are structurally similar but V is dominant over U
- Haploid selection and low sexual dimorphism may explain the sex chromosome structure

INTRODUCTION

Genetic determination of sex is mediated by sex-determining regions (SDRs) of various sizes or by sex chromosomes in a broad range of eukaryotes. Sex chromosomes have arisen independently and repeatedly across the eukaryotic tree and comparative analysis of different sex-determination systems has provided insights into how these systems originate and evolve. A typical sex chromosome pair is thought to derive from a pair of autosomes through the acquisition of genes involved in sex determination. If more than one locus involved in sex determination is located on the chromosome, recombination between loci is expected to be suppressed to avoid the production of mal-adapted individuals with a combination of male and female alleles of the sex-determining genes. This leads to the establishment of a non-recombining region on the nascent sex chromosome, with important consequences for the evolution of this region of the genome [1]. For example, as a result of the suppression of recombination within the SDR, repetitive DNA tends to accumulate, leading to an increase in SDR size and degeneration of genes within the non-recombining region. At a later stage, deletion of non-functional DNA from within the SDR may lead to a decrease in the physical size of the SDR.

There is also evidence that the non-recombining region can progressively encroach on the flanking regions of the chromosome, so that it encompasses an increasingly greater proportion of the sex chromosome. This process is thought to be driven by the recruitment of genes with differential selective benefits to the two sexes (sexually antagonistic genes) into the SDR [2] (but see [3]). Extension of the SDR in this manner can lead to the creation of "strata", which are regions of the SDR that have become non-recombining at different points in evolutionary time [4-7].

The genetic mechanism of sex determination also influences how the sex chromosomes evolve. In organisms where sex is expressed in the diploid phase, such as most animals and land plants, one sex

is heterogametic (XY or ZW) whilst the other is homogametic (XX or ZZ). In these systems only the Y or W contain non-recombining regions because the X and Z recombine in the homogametic sex. In some algae and bryophytes the male and female sexes are genetically determined after meiosis, during the haploid phase of the life cycle [8, 9]. This type of sexual system, termed UV to distinguish it from the XY and ZW systems described above [10], exhibits specific evolutionary and genetic properties that have no exact equivalent in diploid systems. In UV systems, the female and male SDR haplotypes function in independent, haploid, male and female, individuals and consequently there is no heterozygous sex comparable to XY males or ZW females. This difference between UV and XY/ZW systems should have important implications for SDR evolution [8, 9]. In particular, the female U and male V are expected to be under similar evolutionary pressures not only because they function independently in different individuals but also because neither the U nor the V SDR haplotype recombines [8, 9]. As a result, both haplotypes are expected to exhibit the effects of loss of recombination, such as gene degeneration, to a similar extent. Gene degeneration is, however, expected to be limited in both the U and the V regions provided they both contain genes that are essential during the haploid phase. It has also been suggested that changes in the size of the U or V involved principally additions of beneficial (but not essential) genes, rather than gene losses [8, 9]. Some asymmetry may be expected between the U and V, however, if sexual selection is stronger in males [11] or if one of the chromosomes plays a more active role in sex determination. These verbal predictions of the characteristics of UV systems still need to be rigorously tested empirically.

Although eukaryotic species with UV systems may be as common as those with XY and ZW systems, very few of the former have been characterised, with detailed sequence data being available for only two members of the Archeaplastidea lineage: the liverwort *Marchantia* (which has a fully sequenced V but a partially identified U chromosome) [12] and a UV pair of unknown age in the green alga *Volvox* [13], together with more fragmentary information recently obtained for the moss *Ceratodon* [14]. Clearly, additional detailed sequence information is required to fully test the predictions that have been made with respect to UV sex-determination systems, and to evaluate the generality of these predictions in a broad phylogenetic context.

We report here the identification and the genetic and genomic characterisation of the U and V sex-determining regions of the brown algal model *Ectocarpus* sp. (formerly included in *E. siliculosus*) [15, 16]. Brown algae belong to the Stramenopiles, a lineage very distantly related to animals, fungi and green plants (the common ancestors dating back more than a billion years). The brown algae are considered to possess sex chromosomes rather than mating-type chromosomes [17-19] for a number of reasons: 1) there is a strict correlation between gamete size and sex in anisogamous species, 2) all sexual brown algal species exhibit some form of sexual dimorphism [20, 21] and 3) heteromorphic sex chromosomes have been identified in some species [22, 23]. Previous work has shown that sex is determined by a single, Mendelian locus in *Ectocarpus* sp. [24]. During the haploid-diploid life cycle of this organism, meio-spores, produced by the sporophyte generation, develop into dioicous

(separate male and female) gametophytes, which then produce either male or female anisogametes (Figure 1A).

We show here that the *Ectocarpus* sp. UV has features typical of sex chromosomes in other systems, such as low gene density and a large amount of repeated DNA. The male and female sex-determining regions (SDRs) are extremely diverged, reflecting a long independent evolutionary history, which we estimated at more than 70 million years. Despite its age, the SDR constitutes only a fifth of the sex chromosome. A possible explanation for this observation was suggested by the low number of sex-biased genes, implying that sexual conflict may be insufficient in *Ectocarpus* sp. to drive extensive SDR expansion. Both the male and female SDR haplotypes showed signs of degeneration despite the action of purifying selection during the haploid phase of the life cycle. Analysis of expression data suggested that the genes that have escaped degeneration function during the haploid phase of the life cycle. The male SDR haplotype was dominant over the female haplotype, suggesting that the V chromosome determines maleness, with femaleness possibly being the default state when this chromosome is absent. A male-specific high mobility group (HMG) domain gene was identified as a candidate male sex-determining gene. Analysis of the *Ectocarpus* sp. SDR has underlined the universality of sex chromosome evolution across the eukaryotes and has provided important insights into sex chromosome evolution in UV sexual systems.

RESULTS

Identification and characterisation of the *Ectocarpus* sp. SDR

The initial screen for SDR sequence scaffolds used comparative genome hybridisation experiments [25] to identify three male-specific scaffolds. PCR-based markers were used to localise these scaffolds to linkage group 30 of the *Ectocarpus* sp. genetic map [26] (Figure 1B, Table S1A-C). Searches for additional male SDR scaffolds were then carried out by searching for scaffolds carrying male-specific genes using male and female transcriptomic data and by adapting the Y chromosome genome scan (YGS) method, which uses short-read sequencing and k-mer comparison to identify sex-linked sequences [27] (see the Supplementary Notes section for further details). Together, these methods allowed the identification of two large sequence scaffolds corresponding to the male SDR haplotype. Sex linkage was systematically verified by genetic mapping (Table S1B-C).

Further analysis of the segregation patterns of genetic markers corresponding to SDR scaffolds in a single family of 2000 siblings detected no recombination events (Figure 1B). The SDR therefore behaves as a discrete, non-recombining haplotype. This genetic analysis indicated that the male SDR extended over a region of approximately 920 kbp (Figure 1C, Table 1).

To characterise the female haplotype of the sex locus, we sequenced the genome of a female *Ectocarpus* sp. strain that is closely related to the sequenced male strain (Figure S1A) [16]. Several strategies were used to identify candidate female SDR scaffolds (Supplemental Material, Tables S1E-H). These included searches for female orthologues of male SDR protein sequences, a search for

scaffolds carrying female-specific genes based on male and female transcriptomic data and again by adapting the YGS method [27] but to search for female, rather than male, scaffolds. The cumulative size of the female sex-linked scaffolds was 929 kbp. Assuming that the combination of approaches used here has provided a near-complete list of male and female SDR scaffolds, this indicates that the male and female SDR haplotypes are of similar size (Figure 1C, Table 1).

To confirm cosegregation of the SDR with sexual phenotype, 34 *Ectocarpus* strains of known sex from different geographical origins and species were genotyped with several sex locus markers, corresponding to both the male and female SDR haplotypes (Table S1D). In all cases the SDR genotype correlated with sexual phenotype confirming that this region is the sex-determining locus in *Ectocarpus*.

The SDR is flanked by two large recombining regions, which we refer to as pseudoautosomal (PAR) domains. Analysis of molecular marker segregation [26] indicates that these regions recombine during meiosis, unlike the SDR (Figure 1B). The PAR had gene density, intron length, and percent GC content intermediate between those of the autosomes and the SDR (Table 1, Figure 1B). These unusual features are characteristic of the entire recombining part of the chromosome and are not restricted to the regions closest to the SDR (Figure 1B). It is currently not clear why the PAR exhibits these structural differences compared to the autosomes.

Both the male and female SDR haplotypes are rich in transposable element sequences (Figure 1B, 2A) and gene poor compared to the autosomes (Table 1), features typical of non-recombining regions [1]. With only one exception (LTR transposons in the female SDR), all TE classes were more abundant in the SDR and the PAR than in the autosomes, with the differences being particularly marked for both SDR haplotypes. When individual classes of transposable elements were considered, retrotransposons (which represent the least abundant transposon class in the *Ectocarpus* sp. genome as a whole) showed the most marked proportional enrichment in the SDR haplotypes compared to the autosomes (Figure S2A).

About 30% of the euchromatin of the male-specific (non-recombining) region of the human Y-chromosome consists of multiple, different "ampliconic sequences" which exhibit 99.9% identity within each set of repeated sequence. The identity between these sequences has been taken as evidence for a high level of gene conversion within this region [5, 28]. It was further suggested that gene conversion might "substitute" for inter-chromosomal recombination to some extent, counteracting the degenerative effects of reduced recombination within the SDR. Very little intra-haplotype sequence similarity was identified within either the male or the female *Ectocarpus* sp. SDR haplotypes (Table S1J). The total lengths of the repeated regions within the male and female SDR were only 2.5% and 3.2%, respectively. It therefore seems unlikely that mechanisms similar to those proposed for the human Y chromosome have operated in this SDR, although it should be noted that large ampliconic repeats are difficult to assemble and some sequences of this type may not have been identified, particularly for the female haplotype.

The male SDR haplotype contains 17 protein coding genes and three pseudogenes, whereas 15 protein-coding genes and seven pseudogenes were found in the female haplotype (Figure 1C, Figure 3, Table S2). Eight of the female protein coding genes and three of the pseudogenes are homologous to male SDR sequences ("gametologues"), consistent with the two SDR haplotypes having evolved from a common ancestral autosomal region. The classification of these genes as gametologues was supported by expression analysis, which showed that transcript abundances for gametologue pairs were strongly correlated (Figure S2B), and by their conserved intron/exon structures (Figure S3). This correlated expression pattern is consistent with the gametologue genes having been retained because they have non-sex-specific functions during the haploid phase of the life cycle. The genes and pseudogenes that were only found in one (male or female) haplotype may have either been acquired since the divergence of the U and V regions or been lost by the counterpart haplotype. Eighteen of the male and female genes/pseudogenes that were found in only one haplotype had homologues outside the SDR (including, in two cases, genes on linkage group 30; Figure 3, Table S2). The high similarity between some of these SDR genes and their closest autosomal homologues would be consistent with these gene pairs having arisen from recent gene duplication events (i.e. since the divergence of the U and the V) that created either the SDR or the autosomal copy. The remaining two genes that were found in only one haplotype may represent cases of gene loss in the other haplotype, but they could also have resulted from gene relocation to the SDR. Testing these hypotheses will require comparison with a homologous gene from an outgroup species.

Genomic degeneration of the SDR region

Suppression of recombination across the SDR is expected to lead to genetic degeneration unless there is strong selection on gene function to counteract this effect. There are several indications that genetic degradation has occurred, at least to some degree, in the *Ectocarpus* sp. SDR. We identified a set of optimal codons for *Ectocarpus* sp. (Figure S2C and S2D). Selection on codon usage is known to be of weak intensity and particularly sensitive to loss of recombination [29, 30]. The coding sequences of SDR genes exhibited significant under-representation of optimal codons (Figure 2B). This suggests maladapted codon usage (although we cannot exclude that the under-representation is due, at least in part, to reduced rates of biased gene conversion [31] due to the loss of recombination within the SDR). In addition, transcripts of SDR genes tended to be less abundant on average than transcripts of autosomal genes, although note that codon usage and expression level are likely to be correlated so these two parameters are not necessarily independent. Reduced transcript abundance was particularly marked for SDR genes that were exclusively present in one of the haplotypes (Figure 2C), and may reflect degradation of the promoter and cis-regulatory sequences of these SDR genes. The same tendency was observed for the *Volvox* mating locus, where haplotype-specific genes were expressed at lower levels than genes that are part of a gametologue pair [13], suggesting that genetic degeneration of haplotype-specific SDR genes may be a general phenomenon. Note that expression analysis of the *Ectocarpus* sp. gametologue genes did not provide any evidence that these genes are degenerating.

SDR genes were found to be much longer on average than genes elsewhere in the genome, due principally to the presence of longer introns (Table 1). This difference was partly explained by the presence of a larger amount of inserted transposable element DNA (Figures 2A and S2E), which is typical of non-recombining regions.

Although these various analyses provided some evidence for genomic degeneration in the SDR, the overall degree of degeneration was modest compared to previously characterised systems [32], perhaps because both the U and V SDR haplotypes have essential functions during the haploid phase and are constantly exposed to selection (in contrast to Y or W chromosome genes which are always heterozygous). An analysis of SDR gene expression supported this hypothesis: transcripts of SDR genes were consistently present during the haploid phase of the life cycle (Figure 4). Another potential explanation for the limited degree of degeneration is that the SDR is small compared to most previously characterised systems and this may have limited the potential for Hill-Robertson interference among selected sites [33-35].

Predicted functions of SDR genes

Of the 10 genes that were found in the male but not the female SDR haplotype, one was of particular interest because it was predicted to encode a HMG domain protein (Figure S4A, Table S4A). This family of proteins has been implicated in sex or mating type determination in both vertebrates and fungi [36, 37]. The SDR of the green alga *Volvox* also contains a HMG gene [13]. In addition, several of the genes that were found in both the male and female SDR haplotypes (gametologues) were predicted to encode potential signal transduction proteins (including a Ste20-like kinase, a casein kinase, a GTPase, a RING zinc finger protein and a MEMO domain protein; Table S2) and could potentially be involved in the regulation of sex determination.

An ancient sex-determining region

At the sequence level, the male and female haplotypes are extremely divergent. No large blocks of sequence similarity were found and the only regions with a high level of similarity corresponded to gametologue exons (Figure S3). This divergence suggests that the male and female haplotypes have been evolving independently over a long period. Two phylogenetic trees were constructed based on sequences of either an SDR or an autosomal sequence from three *Ectocarpus* lineages and three distantly related brown algal species *Scytosiphon lomentaria*, *Sphaerotrichia firma* and *Laminaria digitata*. The topology of the phylogenetic tree based on the autosomal region was consistent with sequential speciation, with sequences from male and female strains of the same lineage grouping together (Figure 5A). In contrast, in the phylogenetic tree based on the SDR gene, sequences grouped together according to gender (Figure 5B). Note that we were not able to obtain sequence for this gene from female *L. digitata* individuals, suggesting that they may have lost the female gametologue. These data suggest that the SDR originated at least 70 My ago and may be substantially older. The rate of synonymous substitution (dS) in the coding regions of the 11 male and female gametologue pairs (Figure 5C) was used to independently evaluate the age of the SDR.

The dS values for these gene pairs were compared with values for orthologous, autosomal gene pairs across twelve brown algal and diatom species for which divergence times had been estimated (Supplemental Material). The dS values for the SDR genes were remarkably high (mean value of 1.7, with most genes having dS >1) and comparisons with values obtained for the pairs of autosomal orthologues indicated that the male and female haplotypes of the SDR stopped recombining more than 100 Mya (Figure S5). Note however that the estimations based on genetic divergence are approximate because of saturation of synonymous site mutations at the evolutionary distances measured. These analyses suggest that the *Ectocarpus* sp. UV SDR is an old system comparable to the *Drosophila* (60 MY) [32] and mammalian (180 MY) [39, 40] XY systems.

When dS values were calculated on an exon-by-exon basis, individual exons with a markedly lower dS value than those of the other exons within the gametologue gene pair were identified for three of the 11 gametologue pairs (Figure S3). The presence of these rare variant exon pairs suggests that gene conversion events affecting individual exons or small gene regions may have occurred since the divergence of the male and female SDR haplotypes, but more detailed studies are needed to address this possibility.

Limited expansion of the *Ectocarpus* sp. SDR

Given its age, and the prediction that an SDR should progressively enlarge over time to encompass a large part of its chromosome [1, 41], it is remarkable that the *Ectocarpus* sp. SDR accounts for only about one fifth of linkage group 30 and extending over less than a Mbp. It is possible that the small size of the SDR is related to the low level of sexual dimorphism in *Ectocarpus* sp. as the recruitment of sexually antagonistic genes is believed to be an important driver of SDR expansion [1, 41]. Moreover, sexually antagonistic polymorphisms are predicted to be less stable in haploid systems than in diploid systems because dominance effects in XX (or ZZ) individuals are expected to favour allele maintenance in the latter [42, 43]. This effect may also limit expansion of the SDR by reducing the number of genes with sexually antagonistic polymorphisms available for recruitment into the SDR. Consistent with these hypotheses, comparison of the transcriptomes of male and female gametophytes indicated that only about 4% of *Ectocarpus* sp. genes showed sex-biased expression at the mature sexual stage of the life cycle (compared with up to 50-75 % in *Drosophila* for example; [44, 45]; Table S4C).

SDR gene expression and dominance

Quantitative PCR was used to measure the abundance of SDR gene transcripts in near-isogenic male and female strains (Figure 4) at different stages of the life cycle (Figure 1A). While no clear pattern was observed for the female SDR genes, transcripts of two thirds of the male SDR genes that were analysed were most abundant in mature gametophytes (Figure 4) suggesting that these genes have a role in fertility. Interestingly, the transcript of the male gene that is predicted to encode a HMG domain protein was more than ten-fold more abundant in mature gametophytes than at the other stages assayed (Figure 4). The other fertility-induced genes included both additional male-specific

genes (encoding conserved unknown proteins) and several gametologue pairs (predicted to encode a GTPase, a MEMO-like domain protein, a nucleotide transferase and a homoaconitate hydratase, for example) (Table S2).

Diploid gametophytes bearing both the male and the female SDR haplotypes (UV) can be generated artificially, and these individuals are always phenotypically male, indicating that the male haplotype is dominant [24, 46]. This dominance relationship would be consistent with the existence of a master regulatory gene that determines maleness, carried by the V chromosome. To determine whether the dominance of the male haplotype is dose dependent, we used the life cycle mutant *ouroboros* [46] to construct 13 independent triploid (UUV) and tetraploid (UUUV) gametophytes (Figure S1A, Table S1I). All tested polyploids produced male gametes (as determined by genetic crosses with tester lines). Measurements of transcript abundances for 11 female SDR genes did not detect a marked down-regulation of these genes in diploid heterozygous gametophytes compared to haploid gametophytes (Figure S4B and S4C). This suggests that the male haplotype does not silence female gene expression in this heterozygous context (although it was not possible to rule out that the expression of specific female haplotype genes was suppressed). It is likely, therefore, that gametophytes adopt the female developmental program by default, when the male SDR haplotype is absent.

DISCUSSION

This study has demonstrated that sex is determined during the haploid phase of the brown alga *Ectocarpus* sp. by a non-recombining region on linkage group 30 that extends over almost 1 Mbp. The male and female haplotypes of the SDR were of similar size but were highly diverged, the only significant similarity being the presence of 11 gametologues, three of which were predicted to be pseudogenes in the female. Based on comparisons of these shared genes across diverse brown algal species, the SDR was estimated to be more than 100 million years old. Compared with previously characterised systems [47], the *Ectocarpus* sp. UV chromosomes can clearly be classed as an ancient (as opposed to a recently evolved) sex-determining system.

The brown algae belong to the Stramenopiles, which diverged from the lineages that led to green plants and animals more than a billion years ago [48]. This study therefore confirms that SDRs from diverse eukaryote groups share a number of fundamental features such as stable maintenance of pairs of functional alleles (gametologues) over long periods of evolutionary time, suppressed recombination within the SDR, low gene density and accumulation of transposable elements. The presence of 11 gametologue pairs provided unambiguous evidence that the *Ectocarpus* sp. UV pair is derived from an ancestral pair of autosomes, as has been observed for XY and ZW systems in animals and plants [1, 7, 41].

Analysis of the *Ectocarpus* sp. SDR has also allowed a number of predictions that specifically concern UV sexual systems [8, 9] to be tested. UV systems are not expected to exhibit the asymmetrical degeneracy of the sexual chromosomes (degeneracy of the Y and W chromosomes) observed in XY

and ZW systems [32] and this supposition is supported by the similar estimated sizes of the male and female SDR haplotypes in *Ectocarpus sp.*. Based on parameters such as transcript abundance and frequency of optimal codons, the *Ectocarpus* sp. SDR genes exhibit evidence of degeneration but the degree of degeneration is modest compared to that observed for Y-located genes in XY systems of comparable age [32]. As transcripts of all the SDR genes were detected in the gametophyte generation, the modest degree of degeneration is consistent with purifying selection acting to maintain gene functionality during the haploid phase, when the U and V chromosomes are found in separate, male and female, organisms. Selection is indeed expected to be stronger during the haploid phase and to limit degeneration, as suggested for the V chromosome of *Marchantia* [12], another UV system, and by the low dN/dS ratios observed for sex-linked pollen-expressed genes in *Silene latifolia*, a plant with XY chromosomes [49]. The detection of modest levels of gene degeneration indicates that UV SDRs are nonetheless subject to the degenerating effects of suppressed recombination to some degree. Expression analysis indicated that in *Ectocarpus* sp., the SDR genes that escape degeneration belong principally to gametologues pairs, which presumably play a role during the haploid phase, or are male haplotype-specific genes, which are presumably required for male fertility. The *Ectocarpus* sp. SDR contains a large proportion of sex-specific genes (20 male and female sex-specific genes compared with only 11 gametologue pairs). This situation contrasts markedly with the UV system of *Volvox*, where the vast majority of the mating region genes are shared between haplotypes [13]. This difference in gene composition suggests that these two UV systems have had different evolutionary histories, perhaps having been affected in different ways by gene gain and gene loss events. Bull predicted that changes in the sizes of the U and V SDR haplotypes should be due to gain of genes beneficial to the gametophyte rather than gene loss [8, 9]. The presence of a large proportion of haplotype-specific genes in the *Ectocarpus* sp. SDR, relative to the gametologues, and the expression patterns of many haplotype-specific genes, which indicate a role during fertility, would be consistent with his prediction. However, as there is an autosomal paralogue for most of these haplotype-specific genes, it is also possible that functional redundancy of SDR genes and their autosomal paralogues allowed gene loss to occur. Future analysis of additional related SDRs, together with an outgroup species in which the region homologous to the *Ectocarpus* sp. SDR is autosomal, may help to trace changes in SDR gene content over evolutionary time and determine the relative importance of gene gain and gene loss during the emergence of this system.

Despite being ancient, the *Ectocarpus* sp. SDR is quite small. Given the low level of sexual dimorphism in *Ectocarpus* sp. and the small number of genes that show sex-biased expression, both of which suggest that there is limited scope for sexual conflict, the small size of the SDR is consistent with the view that SDR expansion is driven by the evolution of genes with sexually antagonistic effects [1, 50]. In a number of sex chromosome systems, the expansion of the non-recombining region of the Y (or W) has been shown to have proceeded through several events of recombination suppression, which have formed regions with different degrees of X-Y (or Z-W) divergence (evolutionary strata) [4, 51] (reviewed in [1, 47]). The lack of detectable strata is consistent with the

conclusion that this region has experienced limited expansion. However, given that strata may be extremely difficult to detect in ancient haploid systems (as both U and V can accumulate rearrangements) we cannot totally rule out the absence of these events. Indeed, recent evidence suggests the possible existence of at least two recombination suppression events in the UV system of the bryophyte *Ceratodon* [14], and therefore that UV systems may acquire evolutionary strata in some cases. Note also that the *Ectocarpus* sp. system provides independent evidence that the age of an SDR does not necessarily correlate perfectly either with its size or with the degree of heteromorphy (e.g. [52, 53]).

In *Ectocarpus* sp., the male SDR haplotype was dominant over the female haplotype, even when three copies of the female haplotype were present. It is therefore possible that femaleness may simply be the default state, adopted when the male haplotype is absent. This situation is comparable to that observed in diverse animal, fungal and land plant sex-determination systems but differs from that observed with the UV systems of some mosses. In the latter, the male and female factors are co-dominant, leading to monoicy when both the male and female SDR haplotypes are present in the same gametophyte [54]. Functional differences can therefore be observed between different sex determination systems independent of the genetic nature of the system (XY, ZW or UV).

The male-specific HMG gene is a good candidate for the gene that determines maleness in *Ectocarpus* sp.. If this can be confirmed experimentally it will raise important questions about the evolution of sex and mating-type-determination gene networks across the eukaryote tree, suggesting shared or convergent mechanisms in brown algae, fungi and animals.

EXPERIMENTAL PROCEDURES

The raw sequence data generated in this study (Supplemental Information) have been submitted to Genbank with the study ID ERP002539 and to SRA with the experiment accession ID SRX468696 and SRX468697.

Ectocarpus culture

Ectocarpus strains were cultured as described [55].

RNAseq transcriptome data

RNA-seq analysis was carried out to compare the abundances of gene transcripts in male and female mature gametophytes. Synchronous cultures of gametophytes of the near-isogenic male and female lines Ec603 and Ec602 (see Table S1A, Figure S1) were prepared under standard conditions [55] and frozen at maturity. Total RNA was extracted from 2 bulks of 400 male individuals and 2 bulks of 400 female individuals (2 biological replicates for each sex) using the Qiagen Mini kit (http://www.qiagen.com) as previously described [46]. For each replicate, RNAs were quantified, cDNAs for transcriptome analysis were dT primed, fragmented, cloned, and sequenced by Fasteris (CH-1228 Plan-les-Ouates, Switzerland). We used both *de novo* assembly (Trinity) (r2012-01-25) [56]

and TopHat (v2.0.8) [57, 58] and Cufflinks (v2.1.1) [58, 59] algorithms. Statistical testing for sex biased gene expression was performed using DEseq [60].

Identification and mapping of the male SDR

A comparative genome hybridisation approach [25] identified several regions of the genome exhibiting polymorphisms between male (Ec32) and female (Ec568) strains. Primers were developed for these putative sex-linked regions and mapping was performed by genotyping the 60 individuals of the mapping population [26]. Details of the PCR conditions are given in the Supplemental Information. The approaches used to improve the assembly of the male SDR and verification of the completeness of the male SDR using both an RNA-seq-based method and an approach based on the Y chromosome Genome Scan (YGS) method developed by Carvalho and Clark [27] are described in detail in the Supplemental Information section.

Recombination analysis

Recombination between sex locus markers was analysed using a large segregating family of 2000 meiotic individuals (Figure S1) derived from a cross between the male line Ec494 [46] and the female outcrossing line Ec568 [26].

Sequencing of a female strain and identification and assembly of the female SDR

The genome of the female strain Ec597 (Table S1A, Figure S1A) was sequenced using a whole genome shotgun strategy that involved the implementation of both Illumina HiSeq 2000 technology and Roche 454 pyrosequencing. Velvet (version 1.1.05) was used to run several assemblies during the sequencing process, including the V3 assembly (which used all the pair-end reads and reads from one of the mate-pair libraries) and the final V4 assembly with the complete read dataset (Table S1E). An independent *de novo* assembly was also carried out with the CLC assembler (http://www.clcbio.com/products/clc-assembly-cell) using only the pair-end Illumina data.

Female SDR scaffolds were identified using three different approaches. First the deduced protein sequences of male SDR genes (all annotated genes on the two male SDR scaffolds sctg_68 and sctg_285and439) were blasted against the female genome assembly. Fourteen candidate female SDR scaffolds were identified in the V4 assembly using this approach. The second approach employed RNA-seq transcriptome data. Finally, we also adapted the YGS method [27] to identify female-linked sequences. These approaches are described in detail in the supplemental Information section. All putative female specific scaffolds were verified by PCR using between eight and 57 individuals. Several approaches were used to improve the assembly of the female SDR. Details are given in the Supplemental Information section.

Annotation of SDR scaffolds

The male SDR scaffolds had been annotated as part of the *Ectocarpus* sp. genome project [16] but the gene models were considerably improved by integrating transcript information derived from the RNA-seq analysis carried out as part of this study and using comparisons of male and female

gametologue gene models. The updated gene models can be accessed at the OrcAE database (http://bioinformatics.psb.ugent.be/orcae/overview/Ectsi) [61]. The female SDR scaffolds were annotated *de novo* by running the gene prediction program EuGène [62], which incorporated the signal prediction program SpliceMachine [63], using the optimised Markov models and SpliceMachine splice site predictions derived previously for the male genome sequence [16]. Gene prediction incorporated extrinsic information from mapping of the RNA-seq data onto the female-specific scaffolds. Both male and female SDR gene models were manually curated using the raw, mapped RNA-seq data, Cufflinks and Trinity transcript predictions and comparisons between the male and female haplotypes.

Pseudogenes were identified manually by comparing SDR sequences with genes in the public databases. An additional screen for pseudogenes was carried out by blasting male protein sequences against the genomic sequence of the female SDR and vice versa. All sequences that had been annotated as "gene" or "TE" were excluded from this latter analysis using Maskseq and RepeatMasker respectively.

Homologous genes present in both the male and female haplotypes of the SDR were considered to be gametologues if they were detected as matches in a reciprocal Blastp search against the SDR scaffolds (E value cutoff: 1 e^{-4}). The same criterion was used to identify homologues of SDR genes located outside the SDR (Table S2).

Identification of transposons and other repeated sequences in the SDR

An *Ectocarpus*-specific TE-library (described in [16]), which had been compiled with REPET [64], was used to annotate SDR transposons. TEs were also annotated by running the *de novo* annotation software Repclass [65] with default parameters. See the Supplemental information section for details.

Intra-haplotype sequence similarity

Analyses of sequence similarity within the male and female SDR haplotypes were performed using a custom Perl code [5]. By default, the threshold for sequence identity was fixed to 97%. When the threshold was reduced to 50%, the same result was obtained.

Quantitative reverse transcriptase PCR analysis of SDR gene transcript abundances during the *Ectocarpus* sp. life cycle

The abundance of male and female SDR gene transcripts during the *Ectocarpus* sp. life cycle was assessed by RT-QPCR. Primer pairs were designed to amplify regions of the 3'UTR or the most 3' exon of the gene to be analysed (Table S4D). *In silico* virtual PCR amplifications were carried out using the e-PCR program [66] and both the male and female genome sequences to check the specificity of oligonucleotide pairs. RT-QPCR analysis was carried out for 13 male SDR genes and 11 female SDR genes (Figure S4A and S4B). The remaining SDR genes could not be analysed either because they had

very small exons, which posed a problem for primer design, or it was not possible to obtain a single amplification product. RNA extraction and RT-QPCR were performed as previously [46].

Construction of phylogenetic trees for an SDR and an autosomal gene

Exon sequences from an SDR and an autosomal sequence were amplified from three *Ectocarpus* lineages, S. *firma* (E. Gepp) Zinova and *S. lomentaria* (Lyngbye), distantly related brown alga within the order Ectocarpales, and for the kelp *L. digitata* (Hudson) J.V. Lamouroux. For the SDR gene, an exon region was amplified for the gametologue pair Esi0068_0003 (male) and FeV4scaf15_1 (female). Alignable sequence data from the ITS2 nuclear autosomal region and adjacent LSU was obtained for the same strains. Sequences were edited using the Codon Code sequence aligner and aligned with Muscle in the program Seaview [67]. Evolutionary history was inferred using both the Neighbour-joining (Figure 5B-C) and PhyML method implemented in MEGA5 [68] with the same topology resolved by both methods. The strains and lineages used are described in Table S1A and the primers are described in Table S3.

Synonymous divergence

Pairwise alignments of the deduced protein sequences of gametologue gene pairs were performed in Seaview using Muscle with default parameters. Regions with poor alignments were further analysed with Gblocks [69]. The aligned protein sequences were then back-translated to coding sequence and synonymous divergence (dS) was calculated using Codeml within the suite of programs in PAML version 4 [70].

Estimating the age of the *Ectocarpus* sp. SDR

Coding sequence data from 65 stramenopile species including two diatoms were obtained from the Hogenom database version 6 and from Genbank [71]. Homologous genes were identified using a clustering approach. Orthologous sequences were identified and checked using phylogenetic information (described in Supporting Information). Coding sequences from other Phaeophyceae species were added to the cluster data and further data cleaning was carried out so that only orthologous sequences were retained, as described in Supporting Information. A pairwise alignment of the *Ectocarpus* sp. genes with all of the identified orthologous genes from each cluster was then carried out using Prank [72], and alignments were improved using Gblocks [69, 70]. The programs Codeml and Yn00 from PAML version 4 [70] were then run on each gene pair in order to calculate pairwise dS values. The resulting dS values were plotted against the divergence times estimated by Silberfeld *et al*. [38] and Brown and Sorhannus [73].

Codon usage analysis

A set of 27 optimal codons was identified by comparing the codon usage of highly expressed genes (ribosomal genes) with the rest of the genome using the multivariate approach described in Charif *et al*. [74]. Fop values were correlated with RNA-seq expression levels (Figure S2C and S2D).

Sex-determination in strains carrying different numbers of U and V chromosomes

Polyploid gametophytes were constructed using the *ouroboros* mutant [46] (Figure S1A). Details of genetic crosses and ploidy verification are given in Supporting Information.

ACKNOWLEDGMENTS

The authors wish to thank Thomas Broquet, Veronique Storm and Sylvain Mousset for advice on the statistical analysis, Aurélie Kapusta for help with Repclass, Emmanuelle Lerat for explanations about TE librairies, Thomas Bigot and Florent Lassalle for help with TPMS, Catherine Leblanc, Florian Weinberger, Gareth Pearson and Olivier de Clerk for sharing unpublished RNA-seq data and Helen Skaletsky for help with intra-chromosomal similarity analyses.

This work was supported by the Centre National de la Recherche Scientifique, the Agence Nationale de la Recherche (Project Sexseaweed), the University Pierre and Marie Curie (Emergence program), the Interreg program France (Channel)-England (project Marinexus) and the Interreg IVB EnAlgae project. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

TABLES

Table 1. Statistics for several features of the male and female *Ectocarpus* sp. SDR compared with the PAR and the complete genome.

	Male SDR	Female SDR	PAR	Genome
Total sequence (Mbp)	0.92	0.93	4.08	205.27
Genes (incl. pseudogenes)	21	24	228	15779
Average gene length (bp)	25710	18836	8188	6974
Average CDS length (bp)	1373	1050	1217	1607
Average intron length (bp)	3605	3691	1062	702
Average n. introns/gene	6.67	4.81	6.28	7.14
Gene density (genes/Mbp)	22.82	23.66	55.88	76.87
GC%	51.29	44.74	52.20	54.02

FIGURES

Figure 1. The UV sex-determination system of the brown alga *Ectocarpus* sp.. (A) Life cycle of *Ectocarpus* sp. in culture. The sexual cycle (left side of panel) involves an alternation between the diploid sporophyte and haploid, dioicous (male and female) gametophytes. The sporophyte produces meio-spores through meiosis in unilocular sporangia (single-chambered, spore-bearing structures; Uni). The meio-spores are released and develop as gametophytes (each containing either a U or a V sex chromosome), which then produce gametes in plurilocular gametangia (multiple-chambered,

gamete-bearing structures; Plr). Fusion of male and female gametes produces a zygote (containing both the U and the V sex chromosomes), which develops as a diploid sporophyte, completing the sexual cycle. Unfertilised gametes can enter an asexual parthenogenetic cycle by germinating without fusion to produce a partheno-sporophyte (right side of panel). The partheno-sporophyte produces spores through meiosis in unilocular sporangia and these develop as gametophytes, completing the parthenogenetic cycle. Note that the haploid partheno-sporophytes and the diploid sporophytes do not express sex. The parthenogenetic cycle is only shown for a female but male gametes can also develop parthenogenetically in some *Ectocarpus* lineages. Life cycle stages used for the RT-QPCR analysis of SDR gene expression are marked with an asterisk.

(B) Genetic and physical maps of the *Ectocarpus* sp. sex chromosome. The left side of the panel shows a genetic map of the *Ectocarpus* sp. sex chromosome (LG30). The positions of simple sequence repeat (SSR) markers are indicated to the right of the linkage group, with the prefix 'M' for marker followed by the number of the supercontig that contains the SSR, and finally, in some cases, a suffix to distinguish markers that originated from the same supercontig. Sex linked markers are shown in blue. Numbers to the left indicate map distances (in cM) between the intervals given by the lines that cross the vertical bar. The genetic map was generated using a segregating family of 60 individuals, except for the non-recombining region where a larger population of 2000 meiotic individuals was used. The central panel depicts the extent of recombination between markers located inside the *Ectocarpus* sp. non-recombining region. The number of meiotic siblings used to assay for recombination between each pair of markers is indicated, with the percentage of recombinants detected in brackets. Note that no recombination was detected between any of the sex locus markers. See Table S1B for the coordinate position of each marker on its respective scaffold. The right side of the panel shows a physical map of the sex chromosome, and a heatmap of the GC percent, gene density and TE density along the LG30 and along an autosome (LG06) for comparison. The heatmap was computed using a 4,000 bp sliding window.

(C) Overview of the *Ectocarpus* sp. male and female SDR haplotypes. Genes are indicated by arrows, the lighter colours corresponding to gametologues. Gene names (LocusIDs) are indicated, with pseudogenes in grey type and putative transposon remnants in grey italic. Putative transposon remnants were counted as protein coding genes but Esi0068_0068/FeV4scaf25_3 was not included in the set of gametologue pairs. The relative sizes of the male and female SDR genes are indicated but they are not drawn to the same scale as the underlying scaffolds indicated by the dotted line and the scale bar. Only female SDR scaffolds carrying genes are represented. Scaffolds are separated by double diagonal lines indicating that the relative positions of scaffolds within the SDR are unknown. Double-headed arrows indicate the estimated sizes of the SDR haplotypes. The grey bars indicate the sex chromosomes. SDR, sex-determining region; PAR, pseudoautosomal regions. See also Figure S1.

Figure 2. Comparison of genomic features of the SDR, PAR and autosomes. (A). Percentage of DNA corresponding to different classes of transposable element (TEs) in different genomic fractions. Pairwise comparisons using a Fisher exact test indicated that all of the sex chromosome compartments (PAR, male SDR, female SDR) were significantly different from the autosomal compartment (P<0.0001). (B). Median frequency of optimal codons in coding regions of autosomal, PAR and male and female SDR genes. Error bars indicate 95% confidence intervals around the median. An analysis using CAI (the Codon Adaptation Index, another codon usage index [75], which was computed using R and the seqinR package [74]), gave similar results. (C). Mean transcript abundance in sexually mature, male and female gametophytes for genes in different genome fractions, determined by RNA-seq and expressed as fragments per kb of transcript per million fragments (FPKM) mapped. The notched boxplot graph shows the means of autosomal (n=14677), PAR (n=205), male and female SDR (n=37) and SDR without the gametologues (n=16) genes. Significant adjusted p-values compared with autosomes, as calculated by Wilcoxon tests, are indicated by asterisks above each box (*p<0.01, **p<0.001).

Figure 3. Relationships between SDR genes and autosomal genes and expression patterns of the SDR genes. Schematic diagram showing homology relationships between male and female SDR genes and autosomal genes. Autosomal or PAR (i.e. non-SDR) genes are shown in green, male and female SDR genes are shown in pink and blue respectively, with putative functional genes in dark blue or dark pink and pseudogenes in light blue or light pink. Putative transposon remnants are shown in grey. A green box indicates the existence of at least one homologue outside the SDR and the number to the right of the green box indicates the number of matches outside the SDR with an e-value of less than 10^{-4}. Homology relationships were defined based on a Blastp e-value of less than 10^{-4} when predicted protein sequences were blasted against the complete set of *Ectocarpus* sp. predicted proteins. Percentage identity between predicted proteins is indicated above the arrows. The value in brackets corresponds to the length of the matched region as a percentage of the total length of the protein to the left of the arrow. Gene abbreviations are as in the following examples: for male SDR or non-SDR genes 68_16 indicates Esi0068_0016, for female SDR genes 15_1 indicates FeV4scaf15_1. Note that the order of the genes is not intended to correspond to their locations in the genome. The right side of the panel depicts transcript abundances for each of the male and female SDR genes in male and female mature gametophytes, respectively, measured by RNA-seq and expressed as FPKM. See also Figure S2.

Figure 4. SDR gene expression during the life cycle. Male and female SDR gene expression during the life cycle of *Ectocarpus* sp. measured by RT-QPCR, relative to a housekeeping gene (EF1α). SP, diploid heterozygous sporophyte; iGA, immature gametophyte; mGA, mature gametophyte; pSP, partheno-sporophyte; CHP, conserved hypothetical protein. Gene annotations are indicated in brackets (see

Table S2 for further details). Abundances of transcripts for female and male SDR genes were measured using RNA from gametophytes and partheno-sporophytes of strains carrying either the U or the V sex chromosome, respectively, and from diploid sporophytes (strains carrying both the U and the V). Bars with different letters are statistically different ($P<0.05$). Details on the statistical analysis are presented as supplementary information. The coloured dots next to gene names indicate whether the gene is a gametologue (blue and pink dots) or if it is only found in either the male or female haplotype (blue or pink dot, respectively). Graphs corresponding to gametologue pairs are linked by a horizontal line.

Figure 5. Estimation of the age of the *Ectocarpus* sp. SDR. (A) Maximum Likelihood tree created in MEGA5 [68] based on the Kimura 2-parameter model using sequence data amplified from 453 bases of the autosomal region ITS2 and adjacent 5'-LSU. The percentage of trees in which the associated taxa clustered together (bootstrap values from 1000 resamplings) is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Joining and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with the best log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories, +*G*, parameter = 0.2094). Distinct lineages are indicated by different colours. Samples correspond to three different *Ectocarpus* lineages (*E. siliculosus* lineage 1a (E. sil 1a), *E. sp.* lineage 1c (E. sp 1c), *E. fasciculatus* lineage 5b (E. fas 5b), and three distantly related brown algae, *Sphaerotrichia firma* (S. fir), *Scytosiphon lomentaria* (S. lom) *and Laminaria digitata* (L. dig). Lineage names and sex are indicated at the branch tips. The strains used are described in Table S1A.

(B) Maximum Liklihood tree with equivalent parameters to that shown in (A) (Gamma distribution, +*G*, parameter = 0.2868) for 148 bases of the sex-linked, exonic region of one gametologue pair (Esi0068_0003/FeV4scaf15_1). Pink and blue indicate sequences from male and female individuals respectively.

(C) Plot of dS values of gametologue and PAR homologous pairs against gene distance, with gene order according to the male physical map. Blue and purple lozenges represent genes on the two male SDR scaffolds, sctg_68 and sctg_285and439, respectively. Green triangles at each end of the x-axis represent two flanking PAR genes. One-sided standard error bars represent half the standard error of the estimation. Double diagonal bars indicate that the orientation of the locus relative to the flanking PAR is not known. Dotted lines indicate mean levels of synonymous site divergence between *Ectocarpus* sp. autosomal genes and autosomal genes of species from the brown algal groups indicated. See also Figure S5.

REFERENCES

1. Charlesworth, D., Charlesworth, B., and Marais, G. (2005). Steps in the evolution of heteromorphic sex chromosomes. Heredity *95*, 118-128.
2. Jordan, C.Y., and Charlesworth, D. (2012). The potential for sexually antagonistic polymorphism in different genome regions. Evolution *66*, 505-516.
3. Ironside, J.E. (2010). No amicable divorce? Challenging the notion that sexual antagonism drives sex chromosome evolution. Bioessays *32*, 718-726.
4. Lahn, B.T., and Page, D.C. (1999). Four evolutionary strata on the human X chromosome. Science *286*, 964-967.
5. Skaletsky, H., Kuroda-Kawaguchi, T., Minx, P.J., Cordum, H.S., Hillier, L., Brown, L.G., Repping, S., Pyntikova, T., Ali, J., Bieri, T., et al. (2003). The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature *423*, 825-837.
6. Lemaitre, C., Braga, M.D., Gautier, C., Sagot, M.F., Tannier, E., and Marais, G.A. (2009). Footprints of inversions at present and past pseudoautosomal boundaries in human sex chromosomes. Genome Biol Evol *1*, 56-66.
7. Wang, J., Na, J.K., Yu, Q., Gschwend, A.R., Han, J., Zeng, F., Aryal, R., VanBuren, R., Murray, J.E., Zhang, W., et al. (2012). Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. Proc Natl Acad Sci U S A *109*, 13710-13715.
8. Bull, J.J. (1983). Evolution of Sex Determining Mechanisms, (Benjamin/ Cummings).
9. Bull, J. (1978). Sex Chromosomes in Haploid Dioecy: A Unique Contrast to Muller's Theory for Diploid Dioecy. Amer Nat *112*, 245-250.
10. Bachtrog, D., Kirkpatrick, M., Mank, J.E., McDaniel, S.F., Pires, J.C., Rice, W., and Valenzuela, N. (2011). Are all sex chromosomes created equal? Trends Genet *27*, 350-357.
11. Bachtrog, D. (2011). Plant sex chromosomes: a non-degenerated Y? Curr Biol *21*, R685-688.
12. Yamato, K.T., Ishizaki, K., Fujisawa, M., Okada, S., Nakayama, S., Fujishita, M., Bando, H., Yodoya, K., Hayashi, K., Bando, T., et al. (2007). Gene organization of the liverwort Y chromosome reveals distinct sex chromosome evolution in a haploid system. Proc Natl Acad Sci U S A *104*, 6472-6477.
13. Ferris, P., Olson, B.J., De Hoff, P.L., Douglass, S., Casero, D., Prochnik, S., Geng, S., Rai, R., Grimwood, J., Schmutz, J., et al. (2010). Evolution of an expanded sex-determining locus in *Volvox*. Science *328*, 351-354.
14. McDaniel, S.F., Neubig, K.M., Payton, A.C., Quatrano, R.S., and Cove, D.J. (2013). Recent gene-capture on the UV sex chromosome of the moss *Ceratodon purpureus*. Evolution.
15. Peters, A.F., Marie, D., Scornet, D., Kloareg, B., and Cock, J.M. (2004). Proposal of *Ectocarpus siliculosus* (Ectocarpales, Phaeophyceae) as a model organism for brown algal genetics and genomics. J Phycol *40*, 1079-1088.
16. Cock, J.M., Sterck, L., Rouzé, P., Scornet, D., Allen, A.E., Amoutzias, G., Anthouard, V., Artiguenave, F., Aury, J., Badger, J., et al. (2010). The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. Nature *465*, 617-621.
17. Billiard, S., López-Villavicencio, M., Devier, B., Hood, M.E., Fairhead, C., and Giraud, T. (2011). Having sex, yes, but with whom? Inferences from fungi on the evolution of anisogamy and mating types. Biol Rev Camb Philos Soc *86*, 421-442.
18. Hood, M.E., Petit, E., and Giraud, T. (2013). Extensive divergence between mating-type chromosomes of the anther-smut fungus. Genetics *193*, 309-315.

19. Menkis, A., Jacobson, D.J., Gustafsson, T., and Johannesson, H. (2008). The mating-type chromosome in the filamentous ascomycete Neurospora tetrasperma represents a model for early evolution of sex chromosomes. PLoS Genet *4*, e1000030.

20. Berthold, G. (1881). Die geschlechtliche Fortpflanzung der eigentlichen Phaeosporeen. Mitt Zool Stat Neapel *2*, 401-413.

21. van den Hoek, C., Mann, D.G., and Jahns, H.M. (1995). Algae: An Introduction to Phycology, (Cambridge: Cambridge University Press).

22. Evans, L.V. (1963). A large chromosome in the laminarian nucleus. Nature *198*, 215.

23. Lewis, R.J. (1996). Chromosomes of the brown algae. Phycologia *35*, 19-40.

24. Müller, D.G. (1975). Sex expression in aneuploid gametophytes of the brown alga *Ectocarpus siliculosus* (Dillw.) Lyngb. Arch. Protistenk. Bd *117*, 297-302.

25. Dittami, S.M., Proux, C., Rousvoal, S., Peters, A.F., Cock, J.M., Coppée, J.Y., Boyen, C., and Tonon, T. (2011). Microarray estimation of genomic inter-strain variability in the genus *Ectocarpus* (Phaeophyceae). BMC Mol Biol *12*, 2.

26. Heesch, S., Cho, G.Y., Peters, A.F., Le Corguillé, G., Falentin, C., Boutet, G., Coëdel, S., Jubin, C., Samson, G., Corre, E., et al. (2010). A sequence-tagged genetic map for the brown alga *Ectocarpus siliculosus* provides large-scale assembly of the genome sequence. New Phytol *188*, 42-51.

27. Carvalho, A.B., and Clark, A.G. (2013). Efficient identification of Y chromosome sequences in the human and Drosophila genomes. Genome Res *23*, 1894-1907.

28. Rozen, S., Skaletsky, H., Marszalek, J.D., Minx, P.J., Cordum, H.S., Waterston, R.H., Wilson, R.K., and Page, D.C. (2003). Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. Nature *423*, 873-876.

29. Bartolomé, C., and Charlesworth, B. (2006). Evolution of amino-acid sequences and codon usage on the Drosophila miranda neo-sex chromosomes. Genetics *174*, 2033-2044.

30. Bachtrog, D. (2003). Adaptation shapes patterns of genome evolution on sexual and asexual chromosomes in Drosophila. Nat Genet *34*, 215-219.

31. Pessia, E., Popa, A., Mousset, S., Rezvoy, C., Duret, L., and Marais, G.A. (2012). Evidence for widespread GC-biased gene conversion in eukaryotes. Genome Biol Evol *4*, 675-682.

32. Bachtrog, D. (2013). Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. Nat Rev Genet *14*, 113-124.

33. Hill, W.G., and Robertson, A. (1966). The effect of linkage on limits to artificial selection. Genet Res *8*, 269-294.

34. Charlesworth, B., and Charlesworth, D. (2000). The degeneration of Y chromosomes. Philos Trans R Soc Lond B Biol Sci *355*, 1563-1572.

35. Bachtrog, D. (2008). The temporal dynamics of processes underlying Y chromosome degeneration. Genetics *179*, 1513-1525.

36. Idnurm, A., Walton, F.J., Floyd, A., and Heitman, J. (2008). Identification of the sex genes in an early diverged fungus. Nature *451*, 193-196.

37. Foster, J.W., Brennan, F.E., Hampikian, G.K., Goodfellow, P.N., Sinclair, A.H., Lovell-Badge, R., Selwood, L., Renfree, M.B., Cooper, D.W., and Graves, J.A. (1992). Evolution of sex determination and the Y chromosome: SRY-related sequences in marsupials. Nature *359*, 531-533.

38. Silberfeld, T., Leigh, J.W., Verbruggen, H., Cruaud, C., de Reviers, B., and Rousseau, F. (2010). 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 *56*, 659-674.

39. Veyrunes, F., Waters, P.D., Miethke, P., Rens, W., McMillan, D., Alsop, A.E., Grützner, F., Deakin, J.E., Whittington, C.M., Schatzkamer, K., et al. (2008). Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. Genome Res *18*, 965-973.
40. Potrzebowski, L., Vinckenbosch, N., Marques, A.C., Chalmel, F., Jégou, B., and Kaessmann, H. (2008). Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. PLoS Biol *6*, e80.
41. Bergero, R., and Charlesworth, D. (2011). Preservation of the Y transcriptome in a 10-million-year-old plant sex chromosome system. Curr Biol *21*, 1470-1474.
42. Fry, J.D. (2010). The genomic location of sexually antagonistic variation: some cautionary comments. Evolution *64*, 1510-1516.
43. Rice, W.R. (1984). Sex chromosomes and the evolution of sexual dimorphism. Evolution *38*, 735-742.
44. Ellegren, H., and Parsch, J. (2007). The evolution of sex-biased genes and sex-biased gene expression. Nat Rev Genet *8*, 689-698.
45. Assis, R., Zhou, Q., and Bachtrog, D. (2012). Sex-biased transcriptome evolution in Drosophila. Genome Biol Evol *4*, 1189-1200.
46. Coelho, S.M., Godfroy, O., Arun, A., Le Corguillé, G., Peters, A.F., and Cock, J.M. (2011). *OUROBOROS* is a master regulator of the gametophyte to sporophyte life cycle transition in the brown alga *Ectocarpus*. Proc Natl Acad Sci U S A *108*, 11518-11523.
47. Bergero, R., and Charlesworth, D. (2009). The evolution of restricted recombination in sex chromosomes. Trends Ecol Evol *24*, 94-102.
48. Yoon, H.S., Hackett, J.D., Ciniglia, C., Pinto, G., and Bhattacharya, D. (2004). A molecular timeline for the origin of photosynthetic eukaryotes. Mol Biol Evol *21*, 809-818.
49. Chibalina, M.V., and Filatov, D.A. (2011). Plant Y chromosome degeneration is retarded by haploid purifying selection. Curr Biol *21*, 1475-1479.
50. Qiu, S., Bergero, R., and Charlesworth, D. (2013). Testing for the footprint of sexually antagonistic polymorphisms in the pseudoautosomal region of a plant sex chromosome pair. Genetics *194*, 663-672.
51. Ellegren, H., and Carmichael, A. (2001). Multiple and independent cessation of recombination between avian sex chromosomes. Genetics *158*, 325-331.
52. Stöck, M., Horn, A., Grossen, C., Lindtke, D., Sermier, R., Betto-Colliard, C., Dufresnes, C., Bonjour, E., Dumas, Z., Luquet, E., et al. (2011). Ever-young sex chromosomes in European tree frogs. PLoS Biol *9*, e1001062.
53. Vicoso, B., Kaiser, V.B., and Bachtrog, D. (2013). Sex-biased gene expression at homomorphic sex chromosomes in emus and its implication for sex chromosome evolution. Proc Natl Acad Sci U S A *110*, 6453-6458.
54. Allen, C.E. (1935). The genetics of bryophytes. Bot Rev *1*, 269-291.
55. Coelho, S.M., Scornet, D., Rousvoal, S., Peters, N.T., Dartevelle, L., Peters, A.F., and Cock, J.M. (2012). How to cultivate *Ectocarpus*. Cold Spring Harb Protoc *2012*, 258-261.
56. Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol *29*, 644-652.

57. Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.

58. Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc *7*, 562-578.

59. Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol *28*, 511-515.

60. Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol *11*, R106.

61. Sterck, L., Billiau, K., Abeel, T., Rouzé, P., and Van de Peer, Y. (2012). ORCAE: online resource for community annotation of eukaryotes. Nat Methods *9*, 1041.

62. Foissac, S., Gouzy, J.P., Rombauts, S., Mathé, C., Amselem, J., Sterck, L., Van de Peer, Y., Rouzé, P., and Schiex, T. (2008). Genome Annotation in Plants and Fungi: EuGene as a model platform. . Current Bioinformatics *3*, 87-97.

63. Degroeve, S., Saeys, Y., De Baets, B., Rouzé, P., and Van de Peer, Y. (2005). SpliceMachine: predicting splice sites from high-dimensional local context representations. Bioinformatics *21*, 1332-1338.

64. Flutre, T., Duprat, E., Feuillet, C., and Quesneville, H. (2011). Considering transposable element diversification in de novo annotation approaches. PLoS One *6*, e16526.

65. Feschotte, C., Keswani, U., Ranganathan, N., Guibotsy, M.L., and Levine, D. (2009). Exploring repetitive DNA landscapes using REPCLASS, a tool that automates the classification of transposable elements in eukaryotic genomes. Genome Biol Evol *1*, 205-220.

66. Schuler, G.D. (1997). Sequence mapping by electronic PCR. Genome Res *7*, 541-550.

67. Gouy, M., Guindon, S., and Gascuel, O. (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol *27*, 221-224.

68. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol *28*, 2731-2739.

69. Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol *17*, 540-552.

70. Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol *24*, 1586-1591.

71. Penel, S., Arigon, A.M., Dufayard, J.F., Sertier, A.S., Daubin, V., Duret, L., Gouy, M., and Perrière, G. (2009). Databases of homologous gene families for comparative genomics. BMC Bioinformatics *10 Suppl 6*, S3.

72. Löytynoja, A., and Goldman, N. (2005). An algorithm for progressive multiple alignment of sequences with insertions. Proc Natl Acad Sci U S A *102*, 10557-10562.

73. Brown, J.W., and Sorhannus, U. (2010). A molecular genetic timescale for the diversification of autotrophic stramenopiles (Ochrophyta): substantive underestimation of putative fossil ages. PLoS One *5*.

74. Charif, D., Thioulouse, J., Lobry, J.R., and Perrière, G. (2005). Online synonymous codon usage analyses with the ade4 and seqinR packages. Bioinformatics *21*, 545-547.

75. Sharp, P.M., and Li, W.H. (1987). The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res *15*, 1281-1295.