MBE Advance Access published August 6, 2015

Article

Title: The pseudoautosomal regions of the U/V sex chromosomes of the brown alga *Ectocarpus* exhibit unusual features

Authors: Luthringer R[1§], Lipinska AP[1§], Roze D[3], Cormier A[1], Macaisne N[1], Peters AF[2], Cock JM[1], Coelho SM[1*]

Affiliations: [1]Sorbonne Université, UPMC Univ Paris 06, CNRS, Algal Genetics Group, UMR 8227, Integrative Biology of Marine Models, Station Biologique de Roscoff, CS 90074, F-29688, Roscoff, France. [2]Bezhin Rosko, 29250 Santec, France [3]UMI 3614, Evolutionary Biology and Ecology of Algae, CNRS, Sorbonne Universités, UPMC, PUCCh, UACH, Station Biologique de Roscoff, CS 90074, 29688 Roscoff, France

*Correspondance: coelho@sb-roscoff.fr

[§]equal contribution

Abstract

The recombining regions of sex chromosomes (pseudoautosomal regions, PAR) are predicted to exhibit unusual features due to their being genetically linked to the non-recombining, sex-determining region. This phenomenon is expected to occur in both diploid (XY, ZW) and haploid (UV) sexual systems, with slightly different consequences for UV sexual systems because of the absence of masking during the haploid phase (when sex is expressed) and because there is no homozygous sex in these systems. Despite a considerable amount of theoretical work on PAR genetics and evolution, these genomic regions have remained poorly characterized empirically. We show here that although the pseudoautosomal regions of the U/V sex chromosomes of the brown alga *Ectocarpus* recombine at a similar rate to autosomal regions of the genome, they exhibit many genomic features typical of non-recombining regions. The pseudoautosomal regions were enriched in clusters of genes that are preferentially, and often exclusively, expressed during the sporophyte generation of the life cycle, and many of these genes appear to have evolved since the Ectocarpales diverged from other brown algal lineages. A modelling-based approach was used to investigate possible evolutionary mechanisms underlying this enrichment in sporophyte-biased genes. Our results are consistent with the evolution of the PAR in haploid systems being influenced by differential selection pressures in males and females acting on alleles that are advantageous during the sporophyte generation of the life cycle.

Introduction

Sex chromosomes have commonly been found to possess strikingly distinctive features compared with autosomes, for example in terms of the content and density of genes and repeat sequences. These characteristics are thought to be a consequence of suppression of recombination between the sex chromosomes (X and Y or Z and W in diploid systems, or U and V in haploid systems; reviewed in [1]). A broadly established model of sex chromosome evolution predicts gradual expansion of the region of suppressed recombination, driven by selection for linkage between the sex-determining region (SDR) and loci at which selection differs between males and females [2,3]. Expansion of the SDR reduces the recombining portion of the sex chromosome, the so-called pseudoautosomal region (PAR). However, the recombining region is usually not lost completely and it is thought that most species retain a PAR because homologous recombination in this region plays a critical role in chromosomal pairing and segregation during meiosis [4,5]. Moreover, there are situations where sexually antagonistic forces may be too weak to drive a marked expansion of the SDR, and an extensive PAR may be preserved. This may be expected to occur, for example, in organisms with a low level of phenotypic sexual dimorphism (e.g. [6]) or where sexually-antagonistic (SA) selection has been resolved by alternative processes such as the evolution of sex-biased gene expression [7].

The evolutionary fate of PAR genes is expected to differ from that of either autosomal or fully sex-linked genes. In particular, sex differences in allele frequencies should be maintained more easily in the PAR, either due to SA polymorphisms (which are maintained under a wider range of conditions than on autosomes), or to other forms of selection, such as heterozygous advantage [1]. These effects are expected to be strongest very near the SDR, and to decay as the genetic distance from the SDR increases (the rate of decay being inversely proportional to the strength of selection maintaining polymorphism) [8,9].

There has been little empirical work on PARs. Analyses of the structure and genetic behaviour of the PAR have mainly focused on organisms that have old, well-differentiated sex chromosomes such as humans and other mammals [10,11], and more recently birds [12]. These PARs have been shown to exhibit several unusual features compared with autosomes, including higher levels of recombination [13,14,15], greater abundance of repetitive DNA [12,16] and differences in GC content [17]. These studies focused on organisms with diploid sexual chromosome systems (XY and ZW), whereas in a large number of taxa including many red, brown and green algae, land plants and fungi, sex is determined during the haploid phase of the life cycle (UV systems; [18]). Many of the theoretical predictions made for diploid sexual systems are also relevant to UV sexual systems, for example concerning the evolution of recombination suppression and the maintenance of sex differences in allele frequencies in the PAR [2]. Some effects, such as the potential of sex differences in selection to drive gene differentiation in the PAR, are expected to be stronger in UV systems because the U and V chromosomes occur only in females and males, respectively (in contrast with the X, for example, which can occur in males and females). At present however, few empirical data are available for haploid sexual systems to test these various predictions.

We have recently shown that the UV sex chromosomes of the brown alga *Ectocarpus* have a small non-recombining SDR, despite being at least 70 MY old, and that this region is bordered by two relatively large PARs [6]. Here, we show that the PARs of these chromosomes recombine at a similar rate to autosomal regions of the genome and yet exhibit many features typical of non-recombining regions. The PARs are enriched in physically linked clusters of genes that are preferentially, and often exclusively, expressed during the sporophyte generation of the life cycle and many of these genes appear to have evolved since the Ectocarpales diverged from other brown algal lineages. A model is presented that provides a possible mechanism for the accumulation of these sporophyte-biased genes on the PARs.

RESULTS

The pseudoautosomal region of the *Ectocarpus* sex chromosome exhibits unusual structural features

The PARs of the *Ectocarpus* sex chromosome (linkage group 30, LG30) represent about 2 Mbp of sequence on each side of the 1 Mbp SDR. We have previously noted that the PARs exhibit a number of structural differences compared to the autosomes. For instance, values for gene density, mean intron length, and %GC content are intermediate between those of the autosomes and the SDR [6].

Several studies [19,20,21] suggest that chromosome size should be taken into account when comparative analyses of chromosome structure are carried out. In *Ectocarpus*, TE content tends to be negatively correlated with linkage group physical size (Spearman's correlation test rho = -0.113, P = 0.598) while gene density and GC percentage increase with chromosome size (Spearman's correlation test rho = 0.303, P = 0.151 and rho = 0.284, P = 0.178, respectively). Consequently, to analyse in detail the unusual structural features of the *Ectocarpus* PARs, we compared the sex chromosome not only to the autosomal regions as a whole (all chromosomes apart from the sex chromosome) but also with one specific chromosome, linkage group 4 (LG04), which is of similar size (5.028 Mbp) to the sex chromosome. For this comparison, all genes on LG30 and LG04 were manually curated to produce high quality annotations for both chromosomes. Comparison of these two genomic regions showed that the PARs contained more transposable element (TE) sequences and lower gene density than LG04, and that GC content and the size of coding regions were significantly lower for the PAR, compared to LG04 (Fig. 1*A*-*D*). Moreover, PAR genes tended to have longer introns, and fewer and smaller exons on average than genes on LG04 (Fig. 1*E*-*H*). All of these differences were also detected when the PARs were compared with

the autosomes (Fig. 1*A-D*; Fig. S1), confirming that the PARs are unusual. Moreover, the features that distinguish the PARs from the autosomes were not confined to the regions that were close to the SDR. The PARs exhibited some structural heterogeneity along their length, with for example a significant negative correlation between TE content and gene content (Pearson's correlation test $r = -0.606$, $P < 0.01$), but we found no evidence that the features that distinguish the PARs from the autosomes (gene structure, GC content, etc.) were more marked in the vicinity of the SDR (Table S1). These unusual structural features are therefore characteristic of the entire PARs.

Recombination along the sex chromosome

The structural analysis described above strongly indicated that the *Ectocarpus* PARs exhibit features resembling those of the non-recombining SDR. Recombination is completely suppressed within the SDR of the *Ectocarpus* sex chromosome [6] but analysis of molecular marker segregation has confirmed that the PARs recombine during meiosis [22]. In order to build a more comprehensive recombination map of the *Ectocarpus* sex chromosome, an expanded segregating population of 280 individuals was genotyped with 23 LG30 markers. The average recombination rate for the PARs was 40 cM/Mbp whereas the average recombination rate for autosomes was 23 cM/Mbp. Comparisons of average rates between adjacent markers indicated that this difference was not significant (Mann-Whitney U-test, $P = 0.28$). However, recombination events were unevenly distributed along the sex chromosome (Fig. 2). Specifically, two regions of high recombination (one of them recombining at about ten times the genome average) were found on each side of the SDR. Recombination between markers within these peaks was significantly higher than the background recombination rate on the sex chromosome (Mann-Whitney U-test, $P = 0.0038$). When markers within these recombination peaks were excluded, the PARs had an average recombination rate of 15.3 cM/Mb, which was still not significantly different from the genome average (Mann-Whitney

U-test, P = 0.388). Globally, we found no significant correlation between recombination rate and TE or gene content (Pearson's correlation test, P >0.05) along the length of the PARs, although there was a tendency for regions that exhibited higher recombination rates to have higher gene density and lower TE density (Fig. 2).

Genetic recombination rates along the PARs were also studied in a segregating family generated from two parental strains of another *Ectocarpus* species, *E. siliculosus* lineage 1a [23], demonstrating that the PARs are also a recombining region in this sister species (Fig S2, Table S2).

Expression patterns of PAR genes during the *Ectocarpus* life cycle

The PARs contain 209 protein-coding genes. We investigated their patterns of expression, using RNA-seq, at several stages of the life cycle of *Ectocarpus*, including male and female immature and fertile gametophytes, and different tissues of the sporophyte generation. The PAR genes exhibited significantly lower mean expression levels than genes on LG04 (median 5.88 RPKM for the PARs compared with 11.16 RPKM for LG04; Mann-Whitney U-test, P = 4.50×10^{-10}). and than autosomal genes in general (median 9.88 RPKM for all autosomes; Mann-Whitney U-test, P < 1.10×10^{-07}) (Fig. 3*A*). This difference in transcript abundance was particularly marked during the gametophyte generation, and slightly less marked during the sporophyte generation.

A heatmap representing the expression levels of the PAR genes revealed a striking pattern (Fig. 3*B*, Fig. S3*A*). Several clusters of genes had coordinated patterns of expression during the life cycle, including two clusters of PAR genes that were strongly up-regulated during the sporophyte-generation, and a cluster of genes that exhibited transcription below the detection limit (RPKM<1), during both the gametophyte and the sporophyte generations. The sporophyte-biased gene clusters were localised in regions of the PAR that exhibited low levels

of recombination (in supercontigs sctg_96 and sctg_266, Fig. 2, Table S3). No other linkage group exhibited similar patterns of generation-biased gene clusters (Fig. S3*B*).

To further analyse the relationship between genomic location and life cycle expression pattern, we carried out a genome-wide analysis to identify genes that were differentially expressed during the alternation between the sporophyte and gametophyte generations of the life cycle. About 25% of the genes in the *Ectocarpus* genome were significantly differentially regulated between the generations (FC>=2, FDR<0.1), with slightly fewer sporophyte-biased genes (about 12% of the genome, 1,883 genes) than gametophyte-specific genes (about 13%, 2,083 genes). The PAR was found to be significantly enriched in genes that are up-regulated during the sporophyte generation (chi-square test, $P_{adj}=2.2 \times 10^{-7}$, Bonferroni correction) (Fig. 3*C*), while none of the autosomes exhibited a significant enrichment in sporophyte-biased genes (Fig. S3*B*, Fig. S4*C*).

To examine the relationship between level of expression and degree of generation-bias, the sporophyte-biased genes on the PARs, on LG04 and on all autosomes were grouped according to fold-change in transcript abundance between the sporophyte and gametophyte generations, and the mean expression level (RPKM) of each group was plotted (Fig. 3*D*). For LG04, and for autosomal genes in general, the degree of sporophyte-biased expression was determined by the level of expression in the gametophyte, so that their high fold difference correlated with low gametophyte expression. In contrast, all the sporophyte-biased genes on the PAR exhibited very low levels of expression in the gametophyte-generation and the degree of sporophyte-biased expression (fold change) was determined both by attenuation of expression during the gametophyte generation and by the strength of expression during the sporophyte generation.

Two types of measurement can be used to describe the expression of a gene in a multicellular organism: the level of gene expression in terms of the number of transcripts present in a

particular tissue, and the breadth of expression (τ), which measures how often the gene is expressed through the life cycle and/or in how many different tissues it is transcribed [24].

We calculated the breadth of expression of *Ectocarpus* genes using gene expression data collected for two types of tissues and at different stages of the life cycle. Globally, PAR genes exhibited greater expression specificity than either LG04 genes or autosomal genes in general (Mann-Whitney U-test, $P < 0.003$) (Fig. S5*A*). Sporophyte-biased PAR genes had τ values that were significantly higher than those of unbiased PAR or autosomal genes (Mann-Whitney U-test, $P < 6.5 \times 10^{-5}$) (Fig. S5*B*).

Fifty-one sporophyte-biased and 18 gametophyte-biased genes were identified on the PARs (Table S3). A significant proportion (ca. 50%) of the PAR sporophyte-biased genes were located in the two life cycle gene clusters mentioned above. In these clusters, nine (sctg_266) and 13 out of 19 (sctg_96) contiguous genes exhibited sporophyte-specific expression (Fig. 2, Fig. S4*A*). Clustering analysis confirmed that the distribution of sporophyte-genes on the PAR was not random (Runs test, $P < 2.2 \times 10^{-16}$). The sporophyte-biased genes in the two clusters included a duplicated pair of adjacent genes for which there was one copy in each cluster (Table S3). The regions corresponding to the clusters, which are not closely linked to the SDR (Fig. 2), exhibit lower recombination rates (on average 9 cM/Mbp) than the average PAR rate. However, genes located on the clusters did not exhibit different characteristics from other sporophyte-biased genes located outside the clusters and did not differ from unbiased PAR genes (Fig. S4B). The remaining sporophyte-biased genes were distributed along the PAR in triplets (1), pairs (5) or individually (16) (Fig. S4*A*). Neither functional domains nor orthologues in public databases were detected for most of these genes and it was therefore not possible to identify any enrichment with respect to function. However, possible roles in protein-protein interactions (leucine rich repeats, tetratricopeptide repeats or ankyrin repeats motifs) were predicted for 7 of the 51 sporophyte-biased PAR genes.

Fewer than 12% of the genes in the *Ectocarpus* genome (i.e, 1947 genes) exhibits sex-biased gene expression [6], including 31 that are located in the PAR (Fig. S4*A*, Table S3). This latter set of genes did not display any unusual structural characteristics compared with unbiased PAR genes (Fig. S4*B*). There was also no significant tendency for generation-biased genes on the PAR to be also sex-biased (chi-square test, P = 0.25). Nonetheless, 12 of the 69 generation-biased on the PAR exhibited both generation- and sex-bias and there was a marked correlation between the precise type of life cycle generation-bias and the type of sex-bias: all seven of the genes that were both gametophyte-biased and sex-biased were male-biased, whereas four out of five of the genes that were both sporophyte-biased genes and sex-biased were female-biased (Table S3).

The *Ectocarpus* PAR is enriched in young genes

Recently evolved genes (referred to as "orphan" genes) tend to exhibit similar features to those that we observed for the PAR genes, including shorter coding regions, fewer exons, lower expression and weaker codon bias compared with older genes [25,26]. We therefore investigated whether gene age might be one of the factors underlying the unusual features of the PAR. Complete genome resources are currently insufficient to identify orphan genes, which are defined as having evolved within a species or group of species in *Ectocarpus,* but we were able to distinguish "young genes" from "old genes" by carrying out Blastp comparisons with other complete stramenopile genomes, including the recently published *Saccharina japonica* genome [27], and sequences in the public databases. Young genes were defined as having no Blastp match (10^{-4} E-value cut-off) with any of these other genomes (indicating that they are likely to have evolved since the split from the most recent common ancestor, about 100 MYA [28]). The PAR was significantly enriched in young genes compared to all the autosomal linkage groups (34% vs. 10%, chi-square test with Bonferroni correction, P =1.5x 10^{-14}). On average, young genes tended to be smaller and to have higher

tissue specificity than old genes and their coding regions were smaller with lower CAI and GC3 (Mann-Whitney U-test, Table S4). When gene age was factored out by comparing only old genes or only young genes between the PAR and the autosomes, the PAR genes still exhibited higher percentage TE, lower GC content, longer gene size, shorter coding regions (significant for old genes only), shorter exons (significant for old genes only) and longer introns (Mann-Whitney U-test, Table S4). Taken together, these analyses indicated that the unusual features of the PAR genes could be explained in part by enrichment in young genes. However, when age is corrected for, PAR genes still exhibit markedly different features to autosomal genes. Interestingly, the proportion of young genes that showed generation-bias expression patterns was higher on the PAR than on the autosomes (52% vs 28%, chi-square test, $P = 4.18 \times 10^{-7}$).

Evolution of the PAR genes

The rate and pattern of evolution of *Ectocarpus* genes was analysed by comparing sequences from the reference strain (*Ectocarpus* sp. lineage 1c Peru) with orthologous sequences from another *Ectocarpus* species (*Ectocarpus siliculosus* lineage 1a). Compared with a set of 88 genes from LG04, the 84 PAR genes that were analysed displayed, on average, significantly elevated values for non-synonymous to synonymous substitution ratios (dN/dS) (Mann-Whitney U-test, $P < 0.001$). However, when the generation-biased genes (40 genes) were removed from the data set, no significant difference in mean dN/dS ratios was detected between the PAR and autosomal gene sets (Fig. 4*A*). Moreover, the sporophyte-biased PAR genes showed dN/dS ratios that were significantly higher than sporophyte-biased genes on LG04 (Mann-Whitney U-test, $P = 2.268 \times 10^{-5}$), indicating that the increased evolutionary rates were related to the fact that these generation-biased genes were located on the PAR. The faster rate of evolution of the sporophyte-biased PAR genes was due to an increase in the rate of non-synonymous substitutions (dN) and not to a decrease in the rate of synonymous

substitution (dS) (Fig. 4*B*,*C*) (Mann-Whitney U-test, $P < 0.01$). Finally, note that although the average dN/dS ratio for unbiased PAR genes was similar to that of the autosomal gene set, the average values for both dN and dS were significantly greater than for the autosomal genes (Mann-Whitney U-test, $P < 0.01$).

Interestingly, there was a weak, negative correlation between expression breadth and dN/dS for the PAR genes (Spearman's rho=0.206, $P = 0.0526$). In other words, PAR genes with higher dN/dS tended to exhibit a lower breadth of expression.

Of the 40 sporophyte-biased PAR genes analysed, 24 had dN/dS ratios that were greater than 0.5, which could be an indication of adaptive evolution [29]. To perform a maximum likelihood analysis of positive selection (PAML), we searched for orthologues of the sporophyte-biased genes using transcriptome data for two additional *Ectocarpus* species (*E. fasciculatus* lineage 5b and *Ectocarpus* sp. lineage 1c Greenland; Table S5). Complete sets of four orthologues from the four species were obtained for only seven of the sporophyte-biased PAR genes and the PAML analysis was therefore carried out using these sets. For one of these comparisons both pairs of models (M1a-M2a, M7-M8) suggested positive selection (Esi0096_0082, $\omega = 0.86$, $P < 0.05$).

Codon-usage bias has been observed in almost all genomes and is thought to result from selection for efficient and accurate translation of highly expressed genes [30]. Optimal codons have been described for *Ectocarpus* [6] and a weak but significant correlation was noted between codon usage bias and gene expression levels [31]. In accordance with these findings, the genes on the PARs, which were expressed at a lower level, on average, than autosomal genes (Fig. 3*A*; Mann-Whitney U-test, $P = 6.06 \times 10^{-5}$), showed significantly lower frequency of optimal codons (CAI) compared with autosomal genes (Mann-Whitney U-test, $P = 2.0 \times 10^{-5}$). Interestingly, we found that genes in the regions close to the SDR tended to have higher

CAI than more distal genes, although the significance is borderline (Spearman's rho = -0.15, P = 0.028).

However, when codon usage analysis was carried out specifically for the groups of sporophyte-biased and unbiased genes, the codon adaptation indexes were significantly lower only for the sporophyte-biased genes on the PAR, compared with all other genes (Mann-Whitney U-test , $P < 0.004$) (Fig S6). Analysis of the *Drosophila* genome identified a positive correlation between codon bias and recombination rate [32,33]. *Ectocarpus* PAR genes located in regions with low recombination rates had significantly lower CAI (Mann-Whitney U-test, P = 0.01879), but we found no significant difference in CAI for sporophyte-biased genes located in PAR with low versus average-to-high recombination rates. Therefore, the local recombination rate does not explain the low codon usage bias of the sporophyte-biased PAR genes (Fig. S7).

A model for the spread of generation-biased alleles located in the PAR

In XY or ZW systems, it has been argued that the excess of sex-biased genes often observed on X (or Z) chromosomes may result from sexually-antagonistic selection (e.g. [34]). For example in XY systems, alleles with recessive or partially recessive effects that increase male fitness at a cost to female fitness are expected to spread more easily on the X than on autosomes; modifiers that decrease the expression of these genes in females may then spread, leading to an excess of male-biased genes on the X. We developed a theoretical model to explore whether a similar scenario (involving generation-antagonistic rather than sexually-antagonistic selection) could explain the excess of sporophyte-biased genes observed on the PARs. This would imply that alleles increasing the fitness of sporophytes but with a fitness cost to gametophytes would spread more easily in the PAR than on autosomes.

The model (detailed in the *SI*) considers a selected locus located at a recombination distance r from the SDR of a UV sex-determination system, at which two alleles (denoted A and a) have different effects on the fitness of sporophytes, female gametophytes and male gametophytes. The different events of the life cycle are diploid selection, meiosis (recombination), haploid selection (within each sex) and fertilization (random union of gametes); the fitnesses of the different genotypes are given in Table 1 (note that the results only depend on relative fitnesses within each ploidy phase and sex, as we assume that selection takes place independently among females, males and sporophytes). The model is similar to that recently proposed by Immler and Otto [2] but, while these authors explored conditions under which selection favours decreased recombination between a PAR locus and the SDR, we focus on the conditions for the spread of a rare allele (say allele a) at the selected locus, as a function of r, and the fitness effect of the allele on sporophytes (s_d) and on female (s_f) and male (s_m) gametophytes. We focus on generation-antagonistic alleles (s_d and $s_h = (s_f + s_m)/2$ have opposite signs), since the spread of such alleles may result in an increase in the frequency of genes that are differentially expressed in sporophytes and gametophytes (generation-biased genes) (Table 1).

Overall, our analysis (Fig. 5) shows that genomic localisation has little effect on the spread of alleles when selection is similar in both sexes ($s_f \approx s_m$). However, when selection differs between the sexes (and in particular when the gametophyte-deleterious allele is neutral or slightly beneficial in one of the sexes), the model indicates that a sporophyte-beneficial allele benefits from sex-linkage, as this allows the allele to avoid being inherited by the sex where it is disfavoured. Linkage to the SDR is also predicted to benefit the gametophyte-beneficial allele, but to a lesser extent since this allele still suffers from a fitness cost in the sporophytic generation. This can be seen on Figures 5B and 5C: reducing the recombination rate between the selected locus and the sex-determining locus (from to solid curves for $r = 0.5$ to the dotted

curves for $r = 0.01$) increases the parameter regions where alleles increase in frequency when rare, this effect being more marked for the sporophyte-beneficial allele (blue curves) than for the gametophyte-beneficial allele (red curves; note that the scale of the x-axis is logarithmic). Therefore, taking into account the possibility of sex-differences in selection, being in the PAR expands the range of parameters allowing generation-antagonistic mutations to spread, but more so for sporophyte-beneficial, gametophyte-deleterious alleles than for gametophyte-beneficial, sporophyte-deleterious alleles. Again, this effect is generated by the fraction of generation-antagonistic mutations that is differentially selected in males and females. This model could thus explain the observed excess of sporophyte-biased gene expression in the PAR, assuming that reduced expression in gametophytes would have evolved secondarily to prevent the expression of alleles that are deleterious in at least one sex (note that complete linkage to the SDR would be another means to resolve this conflict).

Discussion

The *Ectocarpus* PAR does not exhibit an increased recombination rate compared with autosomes but does exhibit local peaks of recombination

PARs play a critical role in successful progression through meiosis in the heterogametic sex of most plant and animal species because at least one crossover is required for correct segregation of the sex chromosomes (e.g. [35,36]), generating a strong selective force to maintain recombination in the PAR. Accordingly, in human males, PAR1 has a crossover rate that is 17-fold greater than the genome-wide average. In contrast, the recombination rate in females, where recombination is between homologous X chromosomes, is comparable to the genome-wide average [11,37]. In UV systems, meiosis occurs in the sporophyte and, consequently, there is no male or female meiosis and all meiotic events involve pairs of U and V chromosomes in which recombination can only occur in the PARs. This feature of UV

systems might be expected to further increase overall recombination rates in the PAR, but measurement of the recombination rate along the *Ectocarpus* PAR indicated a mean rate that was not significantly different from that of the rest of the genome. The absence of a detectable increase in recombination rate is probably explained by the large relative size of the PAR in *Ectocarpus*, which occupies approximately 80% of the sex chromosome [6]. Similarly, the PAR of the blood fluke, and the PAR of the emu, which represent a high proportion (57% and 75% respectively) of their sex chromosomes, both exhibit average recombination rates that are similar to those of autosomes [38,39]. Therefore, there appears to be a general tendency for PARs that constitute a large proportion of the physical size of the sex chromosome not to exhibit increased recombination rates compared with autosomes.

Although the mean recombination rate along the *Ectocarpus* PAR was comparable to that measured for autosomes, recombination mapping identified two peaks of elevated recombination rates flanking the SDR. Fine scale mapping of recombination rates along all the *Ectocarpus* linkage groups will be required to determine whether this type of recombination peak is a specific feature of the sex chromosome or if such peaks occur in autosomes (for example, surrounding regions of reduced recombination such as the centromeres). Recombination hotspots at borders of SDRs have been described for species with XY or ZW sexual systems, including humans [11], mice [14], blood flukes [38], medaka fish [15], emu [39], flycatcher birds [12], *Populus* [40] and papaya [36]. A similar phenomenon has also been observed in fungal mating type chromosomes [41].

The PARs recombine at similar levels to the rest of the genome but exhibit structural characteristics typical of non-recombining regions

The *Ectocarpus* PARs exhibit a number of features that are typical of genomic regions with reduced levels of genetic recombination [42], including increased TE content, decreased gene

density, smaller average CDS size, larger average intron size, higher gene GC content, higher rates of both synonymous and non-synonymous substitution (and higher dN/dS ratios) and lower average gene expression levels compared to autosomes. Paradoxically, despite these features, the mean recombination rate measured for the PAR was not significantly different from that of the autosomal part of the genome. Moreover, we found no evidence that the majority of the PAR genes (excluding sporophyte-biased genes) contained higher levels of sub-optimal codons than autosomal genes. However, note that PAR gene coding regions are significantly shorter than those of autosomal genes and this might counteract any tendency for sub-optimal codons to accumulate, because selective pressures on codon usage are typically stronger for genes that encode short proteins [43]. The PAR was found to be enriched in young genes compared to autosomes but, whilst the presence of these genes contributes to some extent to the unusual features of this region, this enrichment alone does not explain all the unusual structural features of the PAR.

We considered possible evolutionary mechanisms that might explain these unusual structural and functional features of the PAR and its constituent genes. Genetic linkage to the SDR is expected to influence the evolution of the PAR, but the effect should be limited to regions of the PAR that are very close to the SDR, unless selection is very strong [9]. This was not the case for the *Ectocarpus* PAR, as the unusual structural features were characteristic of the entire PAR and were not limited to regions adjacent to the SDR. To date, no mechanisms have been proposed which would allow the SDR to influence the evolution of linked, recombining regions over the distances observed here. It is not clear at present, therefore, whether the unusual structural features of the *Ectocarpus* PAR are related in some way to the presence of the SDR on the same chromosome or if they indicate that the evolutionary history of the PAR has been different from that of the other autosomes.

Preferential accumulation of sporophyte-biased genes on the PAR

The *Ectocarpus* PAR is enriched in sporophyte-biased genes compared with the autosomes and these sporophyte-biased genes appear to be evolving in a different manner to the other genes on the PAR. PAR genes in general showed elevated levels of both synonymous and non-synonymous mutations compared to autosomal (LG04) genes whereas the sporophyte-biased PAR genes showed highly elevated rates of non-synonymous mutations but a similar synonymous mutation rate to unbiased autosomal (LG04) genes. The elevated rate of non-synonymous substitutions could be indicative of adaptive evolution, and indeed a signature of positive selection was detected for one out of the seven sporophyte-biased PAR genes that could be analysed for this feature. However, whilst positive selection may be driving the evolution of some of the sporophyte-biased genes, this is unlikely to be the case for all of them. The set of sporophyte-biased PAR genes had a reduced content of optimal codons compared to an autosomal gene set, suggesting that the majority of these genes are under relaxed purifying selection. One possible explanation for the accumulation of non-optimal codons in these genes is that they may escape haploid purifying selection [44,45,46], as they are completely silent during the gametophyte generation. Consequently, alleles with sub-optimal codons will be masked in diploid heterozygous individuals and will not be selected against during the haploid phase.

Another possibility is that the lack of expression of the sporophyte-biased PAR genes during the gametophyte generation leads to relaxed selection due to the reduced breadth of expression of these genes. Breadth of expression, i.e. the degree of tissue or developmental stage specificity, is known to effect non-synonymous substitution rates [47]. However, this hypothesis alone is not sufficient to explain the higher evolutionary rates of sporophyte-biased genes, because gametophyte-biased PAR genes, which also have a reduced breath of expression, had similar non-synonymous mutation rates to an average PAR gene.

Mathematical modelling was used to identify evolutionary mechanisms that might explain the preferential accumulation of sporophyte-biased genes in the PAR. Consistent with a recent model proposed by Immler and Otto [2], we show that generation-antagonistic alleles spread more easily on the PAR than on autosomes if selection differs between males and females. The model presented here may explain our empirical observations that generation-biased genes accumulate preferentially on the PAR, provided that differences in expression between generations result from generation-antagonistic selection. However, note that there is evidence that the relationship between sex-biased gene expression and sexually antagonistic selection is complex [48,49] and this is likely also to be the case for the relationship between generation-biased gene expression and generation antagonistic selection. Generation-biased gene expression may therefore only be an approximate proxy for generation-antagonism.

Our model also predicts that sporophyte-beneficial, gametophyte-detrimental alleles tend to benefit more from linkage to the SDR than gametophyte-beneficial, sporophyte-detrimental alleles, in situations where selection is much weaker in one sex than in the other. Such a process might explain the prevalence of sporophyte-specific genes in the *Ectocarpus* PAR. Although this phenomenon should occur predominantly in regions that are tightly linked to the SDR (as the influence of the SDR is predicted to decrease rapidly with genetic distance), it may extend over a larger proportion of the PAR if the reproductive system involves partial clonality or inbreeding (thereby reducing effective recombination rates). Note that a recent field study identified both sexual populations and populations that were reproducing asexually [50], consistent with significant levels of asexual reproduction occurring under some conditions.

Young genes are three-fold more abundant in the PARs than in autosomes. This enrichment is likely to be due to a combination of factors. As new genes are often derived from TEs [26,51] the higher density of TEs in the PARs may play a role by permitting a higher rate of creation

of new transcribed loci. This hypothesis is supported by the fact that, compared with the young autosomal genes, a greater proportion of the young genes in the PARs share homology with elements in the repeated fraction of the *Ectocarpus* genome (46.3% compared with 31.7%, Mann-Whitney U-test, P = 0.038). Note, however, that additional factors are likely to be operating because this mechanism does not explain why the young PAR genes are enriched two-fold in sporophyte-biased genes compared to young autosomal genes. Novel, transcribed loci are thought to arise at a high frequency in the genome but most of these loci are thought to be subsequently lost unless they are stabilised by selective forces [26,51]. It is possible that the mechanism considered in our model (where the excess of sporophyte-biased genes on the PAR results from the spread of sporophyte-beneficial, gametophyte-detrimental alleles) promotes the emergence of new genes with sporophyte-biased expression in the PAR. However, this mechanism alone does not seem sufficient to explain the high proportion of young PAR genes that are generation-biased (52%), as it seems unlikely that such a high proportion of the selectively advantageous new genes have generation-antagonistic effects.

Sporophyte-biased genes in the PAR occur in clusters

Almost half of the sporophyte-biased PAR genes are located in two gene clusters that are highly enriched in sporophyte-biased genes. Clustering of genes with related functions does occur in eukaryotic genomes, although to a lesser extent than in prokaryotes [52,53], but the *Ectocarpus* genome as a whole does not exhibit unusually high levels of functional clustering [54]. At present it is not clear what mechanisms led to the formation of these gene clusters on the PAR. Gene duplication has not played a major role in the evolution of these clusters although there are paralogous pairs of two genes across the two clusters. The model presented in this manuscript provides a possible mechanism for the accumulation of sporophyte-biased genes near the SDR and this could lead to clustering. However, neither cluster is adjacent to

the SDR, although it is possible that the clusters have translocated to their current positions as a result of sex chromosome rearrangements.

Conclusion

We provide the first detailed analysis of the structural and evolutionary features of the pseudoautosomal region of a pair of UV sex chromosomes. We show that this PAR recombines at a rate that is not different from any other region of the genome, but remarkably, exhibits a number of structural and evolutionary features that are typically associated with regions of suppressed recombination. The PAR has significantly accumulated clusters of genes that are differentially expressed during the sporophyte versus gametophyte generation of the life cycle, and these generation-specific genes exhibit clear signs of accelerated evolution. We propose a mechanism that may explain some of the exceptional evolutionary features of these regions compared with autosomes.

METHODS

Ectocarpus strains and culture conditions

Ectocarpus strains were cultured as described [55] and details are provided in *SI*. Table S5 describes the *Ectocarpus* species used in this study. Note that, currently only three species are recognised within the genus *Ectocarpus* (*E. siliculosus, E. fasciculatus* and *E. crouanorium*) [56] but there is increasing evidence that the taxa *E. siliculosus* represents a complex of several species. As the type specimen for *E. siliculosus* was isolated in England, we refer to the non-European strains related to *E. siliculosus* (such as the Peruvian and Greenland strains) as "*Ectocarpus* sp.". The *E.* sp. lineage 1c Peru is the reference species of *Ectocarpus* used for the genome sequencing project and genetic map [54,57]. To study PAR recombination in an

additional *Ectocarpus* species we used *E. siliculosus* lineage 1a. *E.* sp. lineage 1c Greenland and *E. fasciculatus* lineage 5b were used to evaluation of rates of gene evolution.

Generation of a fine recombination map

A segregating population of 60 individuals that had been used for the genetic map [22] and additional 220 individuals from a segregating population derived from a cross between strains Ec494 (male) and Ec568 (female) [6] were used to more precisely estimate recombination frequencies across the pseudoautosomal region. Simple sequence repeat (SSR) markers for each of the 23 supercontigs of the sex chromosome (LG30) have been described previously [6,57], and additional markers are described in Table S2 and in *SI*.

RNA-seq

RNA-seq analysis was carried out to compare the relative abundances of PAR gene transcripts at several different developmental stages of the life cycle (immature and fertile male and female gametophytes and two tissues of the sporophyte generation, namely basal filaments and upright filaments). The RNA extractions and processing of sequenced reads were performed as previously described in [6,24] (see Table S6 for the sequencing and mapping statistics and *SI* for details of the methods).

Differential expression analysis between male and female gametophytes, as well as between gametophyte (male and female libraries as replicates) and sporophyte was performed with the DESeq package (Bioconductor) [58] using an adjusted p-value cutoff of 0.1 and a minimal fold-change of 2 (see SI for more details). The PAR was also analysed for the presence of duplicated genes. Duplicated gene pairs were detected as described in [54].

Evaluation of rates of gene evolution

To estimate evolutionary rates of PAR genes we searched *E. siliculosus* lineage 1a genomic data for orthologues of *E.* sp. lineage 1c Peru genes by retaining best reciprocal Blastn matches with a minimum e-value of 10×10^{-10}. Sequences that produced a gapless alignment that exceeded 100 bp were retained for pairwise dN/dS (ω) analysis using PAML (codeml, F3x4 model, runmode=-2). To detect PAR genes under positive selection we used transcriptomic and genomic data from four different *Ectocarpus* species (detailed in *SI* and Table S5). Nucleotide alignments (with a minimum length of 100 bp) were constructed using ClustalW implemented in Mega6 [59,60], curated manually when necessary and transformed to the PAML4 required format using perl fasta manipulation scripts (provided by Naoki Takebayashi, University Alaska Fairbanks). Nonsynonymous (dN) and synonymous (dS) rates were estimated by the maximum likelihood method available in CODEML program (PAML4 package). Effective Number of Codons (ENC) and Codon Adaptation Index (CAI) were calculated using CAIcal server (http://genomes.urv.es/CAIcal/) [61].

Classification of "old" and "young" genes

To determine the effect of gene age on various structural parameters, *Ectocarpus* genes were classified as "old genes" or as "young genes" based on the presence or absence, respectively, of homologous sequences in seven complete stramenopile genomes or in the NCBI database (excluding *Ectocarpus* sequences; February 2015). For the stramenopiles, Blastp searches were carried out against the following complete deduced proteomes: *Thalassiosira pseudonana* (diatom; Thaps3 assembled and unmapped scaffolds, http://genome.jgi-psf.org/Thaps3/Thaps3.download.ftp.html)[62], *Phaeodactylum tricornutum* (diatom; Phatr2 assembled and unmapped scaffolds, http://genome.jgi-psf.org/Phatr2/Phatr2.download.ftp.html) [63], *Aureococcus anophagefferens* (Pelagophyceae; http://genome.jgi-psf.org/Auran1/Auran1.download.ftp.html) [64]; *Nannochloropsis oceanica* (Eustigmatophyceae;

https://bmb.natsci.msu.edu/BMB/assets/File/benning/genome_assembly.txt) [65], *Nannochloropsis gaditana* (Eustigmatophyceae; http://www.nature.com/ncomms/journal/v3/n2/full/ncomms1688.html) [66], *Phytophthora capsici* (oomycete; http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3551261 [67] and *Saccharina japonica* (http://124.16.129.28:8080/saccharina/) [27]. Recent estimates indicate that all these species diverged from the Ectocarpales lineage more than 100 Mya [28,68]. Genes were classified as old genes if their protein sequences detected a Blastp match with an E-value of less than 10^{-4} in any of the subject genomes.

FIGURE LEGENDS

Figure 1. Structural characteristics of the PAR compared with the SDR, LG04 and autosomes. A) % TE calculated per supercontig; B) gene density per supercontig; C) % GC per gene; D) %GC3 per coding sequence; E) gene size; F) number of exons per gene; G) total intron length per gene; H) CDS size per gene. Statistical differences were tested using pairwise Mann-Whitney U-test. Letters shared in common between the groups indicate no significant difference.

Figure 2. Recombination frequency and distribution of transposable elements, gene density in the sex chromosome of *Ectocarpus*. The x-axis indicates the physical position along the sex chromosome. Upper-panel: y-axis indicates the recombination rate (cM/Mbp). Error bars represent 95% confidence intervals. The recombination frequency around the SDR is unusually high. The average recombination between two adjacent markers on the PAR is 40.3 cM/Mbp (18.8-66.9, 95% CI), compared with 217.2 and 95.0 cM/Mbp for the peaks at the borders of the SDR. The red dashed line represents the average recombination frequency over the entire *Ectocarpus* genome (51). The black and red lines on the x axis indicate boundaries between supercontigs (sctgs) and the mid points of supercontigs, respectively. Grey

background rectangle above the upper graph indicates the distribution of generation-biased genes along the sex chromosome. Orange: sporophyte-biased genes; green: gametophyte biased genes. Horizontal bars and asterisks represent clusters of sporophyte-biased genes. See also Fig. 3. Gene and transposable element (TE) density along the *Ectocarpus* sex chromosome on the lower panel are represented by the red and blue (dashed) lines respectively. Analysis of gene and TE density was performed by calculating the % of bases on each supercontig that are part of a gene or a TE, respectively. Vertical grey dashed lines indicate the boundaries between the PARs and the SDR. Note that the existing genetic map only allowed 70% of the genome sequence to be assigned to linkage groups [22] and therefore we cannot exclude the possibility that missing scaffolds have led to an underestimation of the Mbp/cM ratio in some regions of the sex chromosome.

Figure 3. PAR gene expression during different life cycle stages. A) Average gene expression (log_2RPKM) of all autosomes, LG04 (a linkage group of similar size to the sex chromosome) and PARgenes in male and female gametophytes (immature and fertile), and sporophytes. Letters shared between groups indicate no significant difference (Mann-Whitney U-test, $P < 6.0 \times 10^{-5}$). B) Heatmap showing the expression levels of PAR genes during different life cycle stages relative to their position on the sex chromosome (the SDR is excluded). Clusters of sporophyte-biased genes (also represented in Fig. 2) are highlighted by asterisks. GA: gametophyte; SP: sporophyte C) Enrichment of sporophyte-biased genes on the PAR compared with autosomes (chi-square test with Bonferroni correction, $***P_{adj}=6.03 \times 10^{-5}$). D) Expression of sporophyte-biased genes on autosomes, LG04 and PAR measured during the sporophyte (pink) and gametophyte (green) generations. Mean gene expression levels (log_2RPKM) at several degrees of generation-bias (from FC>1 to FC>10) are shown. Error bars represent standard errors of the mean.

Figure 4. Rates of evolution of PAR (generation-biased and unbiased) genes compared with autosomal genes (LG04). Pairwise dN, dS and dN/dS ratios were calculated by comparing orthologous gene sequences from *Ectocarpus* sp. (lineage 1c Peru) and *Ectocarpus siliculosus* (lineage 1a). A) Ratio of non-synonymous to synonymous substitutions (dN/dS). B) Non-synonymous substitutions (dN). C) synonymous substitutions (dS). Letters shared between groups indicate no significant difference (Mann-Whitney U-test, $P < 0.01$).

Figure 5. Conditions for the spread of generation-antagonistic alleles. The sporophyte-beneficial allele (a) increases in frequency when rare above the blue curves, while the gametophyte-beneficial allele (A) increases in frequency when rare above the red curves. Solid curves: $r = 0.5$; dashed curves: $r = 0.1$; dotted curves: $r = 0.01$. The strength of selection in males s_m is fixed to 0.1, while the different panels correspond to different values of s_f: 0.05 (A), 0 (B), and −0.05 (C). Note that swapping s_f and s_m would yield exactly the same results, as the model assumes that both sexes are equivalent.

AKNOWLEDGEMENTS

This work was supported by the Centre National de la Recherche Scientifique, the Agence Nationale de la Recherche (Project SEXSEAWEED ANR12-JSV7-0008, project Bi-CYCLE ANR-10-BLAN-1727 and project IDEALG), the University Pierre and Marie Curie Emergence program, the Interreg program France (Channel)-England (project MARINEXUS) and the ERC (grant agreement 638240) . We thank Lynn Delgat and Anne Vanderheyden for help with the alignments for the dN/dS analysis, Claire Gachon for sharing unpublished sequence data from *E. fasciculatus* and *E.* sp. lineage 1c Greenland, and Didier Jollivet for helpful discussions. We would also like to thank three anonymous reviewers for their comments and suggestions that helped improving the manuscript.

TABLES

Table 1. Fitnesses of the different genotypes at the selected locus.

	AA	Aa	aa	A	a
Sporophyte	1	$1+hs_d$	$1+s_d$		
Female gametophyte				$1+s_f$	1
Male gametophyte				$1+s_m$	1

SUPPLEMENTARY FIGURES

Figure S1. Structural characteristics of the PAR compared with the SDR and autosomes. Small linkage groups (below 25 cM) were excluded from the analysis because they are likely to represent fragments of chromosomes. A) % TE calculated per supercontig; B) gene density per supercontig; C) % GC per gene; D) %GC3 per coding sequence; E) gene size; F) number of exons per gene; G) total intron length per gene; H) CDS size per gene.

Figure S2. Recombination map for the sex chromosome of *Ectocarpus siliculosus* lineage 1a. The positions of simple sequence repeat (SSR) markers are indicated to the right of each linkage group, with the prefix 'M' for marker followed by a first number corresponding to the sex chromosome supercontig in *Ectocarpus* sp. strain Ec32 (lineage 1c Peru), which was originally used to design the SSR marker, and a second number identifying the marker (see also Table S2). Numbers to the left indicate the map distances (in cM) between the intervals given by the lines that cross the vertical bar. SSR markers corresponding to SDR regions in Ec32 (*Ectocarpus* sp. lineage 1c Peru) are marked in red. These markers also showed sex linkage in *Ectocarpus siliculosus* lineage 1a.

Figure S3. Heatmaps representing expression levels of genes on the PAR (A) and on four autosomal linkage groups (B). Genes were clustered based either on their expression patterns (A) or on their physical location along the linkage group (B). The heatmaps shown in (B) should be compared with that shown in Figure 3B for the PARs. Gene names have not been included for clarity. GA: gametophyte; SP: sporophyte.

Figure S4. Distribution and features of sex-biased and generation-biased genes on the PAR. A) Distribution of sex-biased and generation-biased genes on the PAR. Each line corresponds to the following comparisons: 1) sporophyte (SP) versus gametophyte (GA); 2) female versus male gametophytes (immature and mature). Genes are represented by coloured bars. Pink bar, female-biased gene; blue bar, male-biased gene; orange bar, sporophyte-generation-biased gene; green bar, gametophyte-generation-biased gene. The SDR is not shown. Asterisks highlight the two main sporophyte-biased gene clusters. B) Generation- and sex-biased genes on the PAR show no significant differences in structural characteristics compared with unbiased genes (pairwise Mann-Whitney U-test, $P > 0.05$). C) Distribution of sporophyte-biased genes on the *Ectocarpus* chromosomes. Sporophyte-biased genes are significantly enriched on the sex chromosome (LG30) (chi-square test with Bonferroni correction, ***$P < 0.001$).

Figure S5: Expression breadth of PAR genes. A) Expression breadth of all autosomal, LG04 and PAR genes . Different letters above the bar plot indicate significant differences (Mann-Whitney U-test, $P < 0.01$). B) Generation-biased gene expression breadth. Different letters above the bar plot indicate significant differences (Mann-Whitney U-test, $P < 0.001$).

Figure S6. Average codon usage bias on the PAR. Codon adaptation index (CAI) are significantly different for the sporophyte-biased genes on the PAR compared to the autosomal, LG04 and other PAR genes (Mann-Whitney U-test, $P < 0.001$). Different letters above the bar plot indicate significant differences.

Figure S7: Codon adaptation index of genes located in regions of the PAR with recombination rates below, similar to or above the PAR average.

SUPPLEMENTARY TABLES

Table S1. Values for different structural parameters along the PAR do not correlate with distance from the SDR (Spearman's rho; P_{adj} with Bonferroni correction).

Table S2. Simple sequence repeat marker primers used to measure recombination rates along the sex chromosome of *E. siliculosus* lineage 1a (presented in Fig. S1).

Table S3. List of genes on the PAR with corresponding functional annotation and expression bias information.

Table S4. Structural characteristics of old versus young genes depending on their chromosomal location (LG04 or PAR).

Table S5. *Ectocarpus* species used in the study. Lineages of *Ectocarpus* are based on [23].

Table S6. RNA-seq sequencing data statistics.

Table S7. Physical and genetic size of the *Ectocarpus* linkage groups

REFERENCES

1. Otto SP, Pannell JR, Peichel CL, Ashman TL, Charlesworth D, et al. (2011) About PAR: the distinct evolutionary dynamics of the pseudoautosomal region. Trends Genet 27: 358-367.
2. Immler S, Otto SP (2015) The evolution of sex chromosomes in organisms with separate haploid sexes. Evolution.
3. Charlesworth D, Charlesworth B, Marais G (2005) Steps in the evolution of heteromorphic sex chromosomes. Heredity 95: 118-128.

4. Rouyer F, Simmler MC, Johnsson C, Vergnaud G, Cooke HJ, et al. (1986) A gradient of sex linkage in the pseudoautosomal region of the human sex chromosomes. Nature 319: 291-295.

5. Shi Q, Spriggs E, Field LL, Ko E, Barclay L, et al. (2001) Single sperm typing demonstrates that reduced recombination is associated with the production of aneuploid 24,XY human sperm. Am J Med Genet 99: 34-38.

6. Ahmed S, Cock JM, Pessia E, Luthringer R, Cormier A, et al. (2014) A haploid system of sex determination in the brown alga Ectocarpus sp. Curr Biol 24: 1945-1957.

7. Vicoso B, Kaiser VB, 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.

8. Kirkpatrick M, Guerrero RF (2014) Signatures of Sex-Antagonistic Selection on Recombining Sex Chromosomes. Genetics 197: 531-541.

9. Charlesworth B, Jordan CY, Charlesworth D (2014) The evolutionary dynamics of sexually antagonistic mutations in pseudoautosomal regions of sex chromosomes. Evolution 68: 1339-1350.

10. Raudsepp T, Das PJ, Avila F, Chowdhary BP (2012) The pseudoautosomal region and sex chromosome aneuploidies in domestic species. Sex Dev 6: 72-83.

11. Flaquer A, Rappold GA, Wienker TF, Fischer C (2008) The human pseudoautosomal regions: a review for genetic epidemiologists. Eur J Hum Genet 16: 771-779.

12. Smeds L, Kawakami T, Burri R, Bolivar P, Husby A, et al. (2014) Genomic identification and characterization of the pseudoautosomal region in highly differentiated avian sex chromosomes. Nat Commun 5: 5448.

13. Lien S, Szyda J, Schechinger B, Rappold G, Arnheim N (2000) Evidence for heterogeneity in recombination in the human pseudoautosomal region: high resolution analysis by sperm typing and radiation-hybrid mapping. Am J Hum Genet 66: 557-566.

14. Soriano P, Keitges EA, Schorderet DF, Harbers K, Gartler SM, et al. (1987) High rate of recombination and double crossovers in the mouse pseudoautosomal region during male meiosis. Proc Natl Acad Sci U S A 84: 7218-7220.

15. Kondo M, Nagao E, Mitani H, Shima A (2001) Differences in recombination frequencies during female and male meioses of the sex chromosomes of the medaka, Oryzias latipes. Genet Res 78: 23-30.

16. Hinch AG, Altemose N, Noor N, Donnelly P, Myers SR (2014) Recombination in the human Pseudoautosomal region PAR1. PLoS Genet 10: e1004503.

17. Montoya-Burgos JI, Boursot P, Galtier N (2003) Recombination explains isochores in mammalian genomes. Trends Genet 19: 128-130.

18. Bachtrog D, Kirkpatrick M, Mank JE, McDaniel SF, Pires JC, et al. (2011) Are all sex chromosomes created equal? Trends Genet 27: 350-357.

19. Burt DW (2002) Origin and evolution of avian microchromosomes. Cytogenet Genome Res 96: 97-112.

20. (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432: 695-716.

21. Jensen-Seaman MI, Furey TS, Payseur BA, Lu Y, Roskin KM, et al. (2004) Comparative recombination rates in the rat, mouse, and human genomes. Genome Res 14: 528-538.

22. Heesch S, Cho GY, Peters AF, Le Corguillé G, Falentin C, 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.

23. Stache-Crain B, Müller DG, Goff LJ (1997) Molecular systematics of *Ectocarpus* and *Kuckuckia* (Ectocarpales, Phaeophyceae) inferred from phylogenetic analysis of nuclear and plastid-encoded DNA sequences. J Phycol 33: 152-168.

24. Lipinska A CA, Luthringer R, Peters AF, Corre E, Gachon CMM, Cock JM, Coelho SM (2015) Sexual dimorphism and the evolution of sex-biased gene expression in the brown alga *Ectocarpus*. Molecular Biology and Evolution in press.

25. Palmieri N, Kosiol C, Schlotterer C (2014) The life cycle of Drosophila orphan genes. Elife 3: e01311.

26. Arendsee ZW, Li L, Wurtele ES (2014) Coming of age: orphan genes in plants. Trends Plant Sci 19: 698-708.

27. Ye N, Zhang X, Miao M, Fan X, Zheng Y, et al. (2015) Saccharina genomes provide novel insight into kelp biology. Nat Commun 6: 6986.

28. Silberfeld T, Leigh JW, Verbruggen H, Cruaud C, de Reviers B, et al. (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.

29. Swanson WJ, Wong A, Wolfner MF, Aquadro CF (2004) Evolutionary expressed sequence tag analysis of Drosophila female reproductive tracts identifies genes subjected to positive selection. Genetics 168: 1457-1465.

30. Kanaya S, Yamada Y, Kinouchi M, Kudo Y, Ikemura T (2001) Codon usage and tRNA genes in eukaryotes: correlation of codon usage diversity with translation efficiency and with CG-dinucleotide usage as assessed by multivariate analysis. J Mol Evol 53: 290-298.

31. Wu X, Tronholm A, Caceres EF, Tovar-Corona JM, Chen L, et al. (2013) Evidence for deep phylogenetic conservation of exonic splice-related constraints: splice-related skews at exonic ends in the brown alga Ectocarpus are common and resemble those seen in humans. Genome Biol Evol 5: 1731-1745.

32. Haddrill PR, Halligan DL, Tomaras D, Charlesworth B (2007) Reduced efficacy of selection in regions of the Drosophila genome that lack crossing over. Genome Biol 8: R18.

33. Campos JL, Charlesworth B, Haddrill PR (2012) Molecular evolution in nonrecombining regions of the Drosophila melanogaster genome. Genome Biol Evol 4: 278-288.

34. Vicoso B, Charlesworth B (2006) Evolution on the X chromosome: unusual patterns and processes. Nat Rev Genet 7: 645-653.

35. Burgoyne PS, Mahadevaiah SK, Sutcliffe MJ, Palmer SJ (1992) Fertility in mice requires X-Y pairing and a Y-chromosomal "spermiogenesis" gene mapping to the long arm. Cell 71: 391-398.

36. Wai CM, Moore PH, Paull RE, Ming R, Yu Q (2012) An integrated cytogenetic and physical map reveals unevenly distributed recombination spots along the papaya sex chromosomes. Chromosome Res 20: 753-767.

37. Page DC, Bieker K, Brown LG, Hinton S, Leppert M, et al. (1987) Linkage, physical mapping, and DNA sequence analysis of pseudoautosomal loci on the human X and Y chromosomes. Genomics 1: 243-256.

38. Criscione CD, Valentim CL, Hirai H, LoVerde PT, Anderson TJ (2009) Genomic linkage map of the human blood fluke Schistosoma mansoni. Genome Biol 10: R71.

39. Janes DE, Ezaz T, Marshall Graves JA, Edwards SV (2009) Recombination and nucleotide diversity in the sex chromosomal pseudoautosomal region of the emu, Dromaius novaehollandiae. J Hered 100: 125-136.

40. Yin T, Difazio SP, Gunter LE, Zhang X, Sewell MM, et al. (2008) Genome structure and emerging evidence of an incipient sex chromosome in Populus. Genome Res 18: 422-430.

41. Hsueh YP, Idnurm A, Heitman J (2006) Recombination hotspots flank the Cryptococcus mating-type locus: implications for the evolution of a fungal sex chromosome. PLoS Genet 2: e184.

42. Charlesworth D, Charlesworth B (2005) Sex chromosomes: evolution of the weird and wonderful. Curr Biol 15: R129-131.

43. Duret L, Mouchiroud D (1999) Expression pattern and, surprisingly, gene length shape codon usage in Caenorhabditis, Drosophila, and Arabidopsis. Proc Natl Acad Sci U S A 96: 4482-4487.

44. Lewis KR, Benson-Evans K (1960) The chromosomes of *Cryptothallbus mirabilis* (Hepaticae: Riccardiaceae). Phyton 14: 21-35.

45. Lewis KR (1961) The genetics of bryophytes. Trans Brit Bryol Soc 4: 111-130.

46. Lewis KR, John B (1968) The chromosomal basis of sex determination. Int Rev Cyto 17: 277-379.

47. Duret L, Mouchiroud D (2000) Determinants of substitution rates in mammalian genes: expression pattern affects selection intensity but not mutation rate. Mol Biol Evol 17: 68-74.

48. Innocenti P, Morrow EH (2010) The sexually antagonistic genes of Drosophila melanogaster. PLoS Biol 8: e1000335.

49. Parsch J, Ellegren H (2013) The evolutionary causes and consequences of sex-biased gene expression. Nat Rev Genet 14: 83-87.

50. Couceiro L, Peters A, Le Gac M, Hunsperger H, Mauger S, et al. The *Ectocarpus* life cycle in the field supports the hypothesis of niche separation between gametophytes and sporophytes. Evolution doi:10.1111/evo.12702

51. Tautz D, Domazet-Loso T (2011) The evolutionary origin of orphan genes. Nat Rev Genet 12: 692-702.

52. Williams EJ, Hurst LD (2002) Clustering of tissue-specific genes underlies much of the similarity in rates of protein evolution of linked genes. J Mol Evol 54: 511-518.

53. Mugford ST, Louveau T, Melton R, Qi X, Bakht S, et al. (2013) Modularity of plant metabolic gene clusters: a trio of linked genes that are collectively required for acylation of triterpenes in oat. Plant Cell 25: 1078-1092.

54. Cock JM, Sterck L, Rouzé P, Scornet D, Allen AE, et al. (2010) The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. Nature 465: 617-621.

55. Coelho SM, Scornet D, Rousvoal S, Peters NT, Dartevelle L, et al. (2012) How to cultivate *Ectocarpus*. Cold Spring Harb Protoc 2012: 258-261.

56. Peters AF, van Wijk SJ, Cho GY, Scornet D, Hanyuda T, et al. (2010) Reinstatement of *E. crouaniorum* Thuret in Le Jolis as a third common species of *Ectocarpus* (Ectocarpales, Phaeophyceae) in western Europe, and its phenology at Roscoff, Brittany. Phycol Res 58: 157-170.

57. Heesch S, Cho GY, Peters AF, Le Corguille G, Falentin C, 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.

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

59. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948.

60. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30: 2725-2729.

61. Puigbo P, Bravo IG, Garcia-Vallve S (2008) CAIcal: a combined set of tools to assess codon usage adaptation. Biol Direct 3: 38.

62. Armbrust E, Berges J, Bowler C, Green B, Martinez D, et al. (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. Science 306: 79-86.

63. Bowler C, Allen A, Badger J, Grimwood J, Jabbari K, et al. (2008) The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. Nature 456: 239-244.

64. Gobler CJ, Berry DL, Dyhrman ST, Wilhelm SW, Salamov A, et al. (2011) Niche of harmful alga *Aureococcus anophagefferens* revealed through ecogenomics. Proc Natl Acad Sci U S A 108: 4352-4357.

65. Vieler A, Wu G, Tsai CH, Bullard B, Cornish AJ, et al. (2012) Genome, functional gene annotation, and nuclear transformation of the heterokont oleaginous alga Nannochloropsis oceanica CCMP1779. PLoS Genet 8: e1003064.

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

67. Lamour KH, Mudge J, Gobena D, Hurtado-Gonzales OP, Schmutz J, et al. (2012) Genome sequencing and mapping reveal loss of heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen Phytophthora capsici. Mol Plant Microbe Interact 25: 1350-1360.

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