Chromatin landscape associated with sexual differentiation in a UV sex determination system

Josselin Gueno, Michael Borg, Simon Bourdareau, Guillaume Cossard, Olivier Godfroy, Agnieszka Lipinska, Leila Tirichine, J. Mark Cock, Susana Coelho

Chromatin landscape associated with sexual differentiation in a UV sex determination system

Josselin Gueno[1], Michael Borg[2], Simon Bourdareau[1], Guillaume Cossard[1], Olivier Godfroy[1], Agnieszka Lipinska[1,2], Leila Tirichine[3], J. Mark Cock[1*], Susana M. Coelho[1,2*]

[1]Sorbonne Université, UPMC Univ Paris 06, CNRS, UMR 8227, Integrative Biology of Marine Models, Station Biologique de Roscoff, CS 90074, F-29688, Roscoff, France. [2]Department of Algal Development and Evolution, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany [3] Nantes Universite, CNRS, US2B, UMR 6286, F-44000 Nantes, France

[*]Correspondence: susana.coelho@tuebingen.mpg.de; cock@sb-roscoff.fr

Abstract

In many eukaryotes, such as dioicous mosses and many algae, sex is determined by UV sex chromosomes and is expressed during the haploid phase of the life cycle. In these species, the male and female developmental programs are initiated by the presence of the U- or V-specific regions of the sex chromosomes but, as in XY and ZW systems, sexual differentiation is largely driven by autosomal sex-biased gene expression. The mechanisms underlying the regulation of sex-biased expression of genes during sexual differentiation remain elusive. Here, we investigated the extent and nature of epigenomic changes associated with UV sexual differentiation in the brown alga *Ectocarpus*, a model UV system. Six histone modifications were quantified in near-isogenic lines, leading to the identification of 16 chromatin signatures across the genome. Chromatin signatures correlated with levels of gene expression and histone PTMs changes in males versus females occurred preferentially at genes involved in sex-specific pathways. Despite the absence of chromosome scale dosage compensation and the fact that UV sex chromosomes recombine across most of their length, the chromatin landscape of these chromosomes was remarkably different to that of autosomes. Hotspots of evolutionary young genes in the pseudoautosomal regions appear to drive the exceptional chromatin features of UV sex chromosomes.

Introduction

In species that reproduce sexually, sex is often determined by a pair of sex chromosomes: X and Y chromosomes in male-heterogametic species, Z and W in female-heterogametic species or U and V in haploid sexual systems (1). Sex chromosomes originate from pairs of autosomes, but further differentiate after the sex-specific chromosome (Y, W or both the V and U) stops recombining (1–3). Males and females have distinct sex chromosome sets but the extensive phenotypic differences between males and females (sexual dimorphism) are largely caused by differences in autosomal gene expression or so-called sex-biased gene expression. The nature and extent of sex-biased gene expression has been investigated in recent years across a broad range of taxa using genome-wide transcriptional profiling. These studies have revealed that sex-biased gene expression is common in many species, although its extent may vary greatly among tissues or developmental stages (4).

Although many reports have described the nature and evolution of sex-biased genes across several taxa, the molecular mechanisms underlying the regulation of sex-biased genes during sexual differentiation remain poorly understood. One prevalent mechanism to regulate gene expression is through covalent modifications such as DNA methylation and post-translational modification (PTMs) of histone tails. DNA methylation regulates transcription in diverse eukaryotes (5), and may contribute to transcriptional differences between sexes (6), playing for instance an important role in differentiating female morphs like workers and queens in the honeybee (7). In the liverwort *Marchantia*, male and female gametes have different levels of DNA methylation and this is correlated with differences in the expression of genes involved in DNA methylation (8). Histone PTMs are another important component of transcriptional regulation, and can impact gene expression by altering chromatin structure or recruiting histone modifiers. Specific combinations of histone PTMs (so-called chromatin states) are associated with functionally distinct regions of the genome such as heterochromatic regions and regions of either permissive transcription or repression (9). The role of chromatin states in regulating gene expression patterns during development in animals and plants is well established (e.g.(10, 11). However, few studies have carried out chromatin profiling during sexual differentiation to link how chromatin is associated with sex-biased gene expression. In Drosophila, the genome-wide distribution of both active and repressive chromatin states differed between males and females but sex-specific chromatin states appeared not to explain sex-biased expression of genes (12), although differences in the chromatin landscape of males and females influenced by the Y chromosome may contribute to sex-biased gene expression (13). Yen and Kellis (2015) used a comparative epigenomic approach to contrast male versus female human samples, revealing that the X chromosome underlies epigenetic differences

between sexes, but epigenomic differences are not reflected in gene expression differences. In the tunicate *Oikopleura dioica*, distinct combinations of histone PTMs were uncovered in testis versus ovaries (14).

In organisms with XY or ZW sex determination systems, sex chromosomes often exhibit unique patterns of gene expression and unusual patterns of chromatin marks compared with autosomes (e.g. (8, 12, 15, 16). For instance, in *Drosophila* males, where the Y chromosome is transcriptionally repressed and the X chromosome is hyper-transcribed due to dosage compensation (17), both of these transcriptional modifications are correlated with changes in the chromatin configuration (18–21). Sex chromosomes are derived from autosomes, but they are governed by unique evolutionary and functional constrains (22, 23). The sex-limited chromosome (Y or W) degenerates, i.e., loses most of its ancestral gene content, accumulates repetitive DNA and evolves into a heterochromatic appearance (15, 16, 24, 25). In contrast, the homologous chromosome (X or Z) acquires dosage compensatory mechanisms by evolving a hyper-transcriptional state (dosage compensation) (26–29). In *Drosophila*, the ratio of euchromatin-to-heterochromatin is different between the two sexes, which is mainly due to the presence of the a repeat-rich Y chromosome in males (12, 13, 30). Similarly, the Z-specific region in schistosomes has a unique chromatin landscape, dominated by active histone PTMs, that are associated with dosage compensation (31).

In contrast, little is known about how chromatin impacts sexual differentiation in organisms with a UV sexual system such as mosses and algae (32–37), although recent work has analysed the patterns of histone post translational modifications during the haploid-diploid life cycle of the brown alga *Ectocarpus* (38). In UV sexual systems, sex is expressed during the haploid phase of the life cycle, where inheritance of a U or V sex chromosome at meiosis determines whether the multicellular adult will be female or male, respectively (1, 39). UV sexual systems differ markedly from XY and ZW systems (3, 39–41). For example, sexual individuals will only have a single U or V sex chromosome, so chromosome-scale dosage compensation or meiotic sex chromosome inactivation mechanisms are unlikely to exist. Moreover, because Y or W sex chromosomes often undergo genetic degeneration, their size, repeat content and gene density is markedly different to the partner X or Z chromosome. In contrast, U and V chromosomes are expected to undergo only mild degeneration (34, 40, 42) and do not exhibit such an asymmetry because each chromosome functions independently in a haploid context and therefore experiences similar evolutionary pressures (42).

The brown alga *Ectocarpus* has emerged as a powerful model organism to study UV sexual systems (reviewed in (39)). A reference genome is available for this species (43–45), including high quality assembly of its sex chromosomes (42, 43, 45–47). The *Ectocarpus* non-recombining U and V specific regions (SDRs) are relatively small compared with the

pseudoautosomal regions (PAR) (the SDR occupies 1/10th of the sex chromosome), such that the large majority of the sex chromosome recombines within the PAR (42, 46, 47). Note that the female and male SDRs on the U and V chromosomes respectively have about the same physical size and share similar number of albeit distinct genes (42). Given the lack of chromosome-scale dosage compensation, the small size of the SDR, and the fact that the SDRs displays only mild levels of degeneration, the U and V sex chromosomes in this system are not expected to present a markedly distinct chromatin landscape compared with autosomes. However, this prediction has never been tested.

The expression pattern of genes on the U and V sex chromosome differs from that of autosomal genes (39). For example, most *Ectocarpus* genes located on the U and V SDRs are upregulated in the haploid gametophyte phase of the life cycle (42, 48). Moreover, when compared with autosomes, the sex chromosome PARs harbour an excess of evolutionary young or taxonomically restricted genes (47) and are enriched in both life cycle-related genes (sporophyte-biased genes) and female-biased genes (49). Nevertheless, what chromatin states associates with this intriguing composition of genes and patterns of gene expression in a UV sexual system still remains unclear.

Here, we investigated sex-specific chromatin landscapes of autosomes and sex chromosomes in *Ectocarpus*, a model brown alga with a UV sexual system. We built on our recent chromatin profiling in *Ectocarpus* by studying six different histone PTMs – four that are associated with gene activation, namely H3K4me3, H3K9ac, H3K27ac and H3K36me3, and two that are associated with reduced gene expression, namely H4K20me3 and H3K79me2 (38). H4K20me3 was associated with repeated sequences and H3K79me2 with genomic regions that often extended over several genes (38). Note that *Ectocarpus* DNA is not methylated (44) so we have not analysed this modification here. Moreover, *Ectocarpus* lacks polycomb complexes and associated PTMs, including the repression-associated mark H3K27me3 mark (38). Similarly, H3K9me2/3 have been detected in *Ectocarpus* but at very low abundance (38). Consequently, none of these additional repression-associated methylation marks were analysed in this study.

Comparison of the profiles of these six histone PTMs with transcriptomic data showed that chromatin states were predictive of transcript abundance. The chromatin landscapes across the genomes of males and females were similar, overall. However, the chromatin signatures of genes that exhibited sex-biased expression was markedly different in males and females indicating that histone modifications may play an important role in mediating sexual differentiation. Moreover, a substantial proportion of the PAR genes presented sex-specific chromatin patterns. The U and V sex chromosomes were found to have very distinct chromatin landscapes to autosomes, despite the absence of a requirement for chromosome-

scale dosage compensation in *Ectocarpus* and the fact that the U and V chromosomes do not exhibit strong signs of genetic degeneration.

Material and Methods

Biological Material

The near-isogenic male (Ec457) and female (Ec460) *Ectocarpus* lines (Table S1) were generated by crossing brother and sister gametophytes for either five or six generations, respectively (42). The resulting male and female strains, therefore, had essentially identical genetic backgrounds apart from the non-recombining SDR (Table S2). To verify the homogeneity in terms of genetic background, we used the male and female input DNA from the ChIP-seq experiments aligned to the *Ectocarpus* reference genome to assess SNP diversity. The SNPs for male and female samples were called with bcftools mpileup and filtered for minimal mapping quality (--minQ 30), depth of coverage (--minDP 10) and missing data (--max-missing 0.9). We found 2,862,827 valid sites out of which 2,995 were variants (either SNPs or INDELS), differing from the reference genome (the SDR regions were excluded). We next compared the distribution of variant sites between males and females. Only 121 of the 2,995 variant sites were segregated between sexes, which accounts for 0.004% of all sites. Given this very low level of female/male polymorphism, it is highly likely that any differences we observed between the two strains are due to the presence of the female and male SDRs. Furthermore, note that the level of genetic diversity within the SDRs (which represent 869,870 bp and 893,800 bp in the female and male, respectively) has been shown to be extremely low (46), as is the case for the non-recombining regions for animals and plants (50). Therefore, the results presented here are likely to be representative of any male and female strain of *Ectocarpus* species, although they are based on only a single male V chromosome and a single female U chromosome.

Male and female gametophytes were cultured until near-maturity for 13 days as previously described (51) at 13°C in autoclaved natural sea water supplemented with 300 µl/L Provasoli solution, with a light:dark cycle of 12:12 h (20 µmol photons.m^{-2}.s^{-1}) using daylight-type fluorescent tubes. Note that we used between 400-600 male and female haploid individual gametophytes in each replicate, although there was only one genotype per each sex. Ten individual gametophytes were grown in each petri dish. The level of maturity of male and female individual gametophytes was assessed under the microscope to ensure synchrony in terms of developmental stage. All manipulations were performed in a laminar flow hood under sterile conditions.

Comparisons of male and female transcriptomes using RNA-seq

RNA for transcriptome analysis was extracted from the same duplicate male and female cultures as were used for the ChIP-seq analysis (see above). For each sex, total RNA was extracted from a mix of 90 gametophytes each, using the Qiagen Mini kit (http://www.qiagen.com). RNA quality and quantity were assessed using an Agilent 2100 bioanalyzer, associated with Qubit2.0 Fluorometer using the Qubit RNA BR assay kit (Invitrogen, Life Technologies, Carlsbad, CA, USA), as described previously (48, 49).

For each replicate sample, cDNA was synthesized using an oligo-dT primer. The cDNA was fragmented, cloned, and sequenced by Fasteris (CH-1228 Plan-les-Ouates, Switzerland) using an Illumina HiSeq 4000 set to generate 150-bp single-end reads. See Table S1 for RNA-seq accession numbers.

Data quality was assessed using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc; accessed May 2019). Reads were trimmed and filtered using Cutadapt (52) with a quality threshold of 33 (quality-cutoff) and a minimal size of 30 bp.

Filtered reads were mapped to version v2 of the *Ectocarpus* sp. 7 reference genome (45, 53) using TopHat2 with the Bowtie2 aligner (54). More than 85% of the sequencing reads from each library could be mapped to the reference genome (Table S1). Note that the reference genome is from a male strain but the female SDR scaffolds have been added. Consequently, male and female data were mapped to the same reference genome.

The mapped sequencing data were then processed with featureCounts (55) to obtain counts for sequencing reads mapped to genes. Gene expression levels were represented as transcripts per million (TPMs). Genes with expression values below the fifth percentile of all TPM values calculated per sample were considered not to be expressed and were removed from the analysis. This resulted in a total of 18,462 genes that were considered to be expressed.

Differential expression analysis was performed with the DESeq2 package (Bioconductor) (56). Genes were considered to be male-biased or female-biased if they exhibited at least a twofold difference (fold change; FC) in expression between sexes with a false discovery rate (FDR) < 0.05. A list of the sex-biased genes can be found in Table S5.

To calculate breadth of expression we employed the tissue-specificity index tau (57) using published expression data from nine tissues or stages of the life cycle (female and male immature and mature gametophytes, mixed male and female gametophytes, partheno-sporophytes, upright partheno-sporophyte filaments, basal partheno-sporophyte filaments, diploid sporophytes) from *Ectocarpus* (45, 47–49, 58). This allowed us to define broadly expressed (housekeeping) genes (with tau<0.25) and narrowly expressed genes (tau>0.75).

Genome-wide detection of histone PTMs

Male versus female *Ectocarpus* sp. gametophyte ChIP-seq experiments were carried for H3K4me3, H3K9ac, H3K27ac, H3K36me3, H4K20me3, and H3K79me2 and three controls (an input control corresponding to sonicated DNA, histone H3 and immunoglobulin G monoclonal rabbit (IgG)) as in (38). RNA-seq data (see above) was generated from the same samples, to ensure that the histone PTM and gene expression data were fully compatible. For ChIP-seq, 2.8 g (corresponding to 2800 individual gametophytes) of *Ectocarpus* tissue was fixed for five minutes in seawater containing 1% formaldehyde and the formaldehyde eliminated by rapid filtering followed by incubation in PBS containing 400 mM glycine. Nuclei were isolated by grinding in liquid nitrogen and in a Tenbroeck Potter in nuclei isolation buffer (0.1% triton X-100, 125 mM sorbitol, 20 mM potassium citrate, 30 mM MgCl2, 5 mM EDTA, 5 mM β-mercaptoethanol, 55 mM HEPES at pH 7.5 with complete ULTRA protease inhibitors), filtering through Miracloth and then washing the precipitated nuclei in nuclei isolation buffer with and then without triton X-100. Chromatin was fragmented by sonicating the purified nuclei in nuclei lysis buffer (10 mM EDTA, 1% SDS, 50 mM Tris-HCl at pH 8 with cOmplete ULTRA protease inhibitors) in a Covaris M220 Focused-ultrasonicator (duty 25%, peak power 75, cycles/burst 200, duration 900 seconds at 6°C). The chromatin was incubated with an anti-histone PTM antibody (anti-H4K20me3, reference 5737S, anti-H3K4me3, reference 9751S and anti-H3K9ac, reference 9649S, Cell Signal Technology; anti-H3K27ac, reference 07360, Millipore; anti-H3K36me3, reference 9050, Abcam; anti-H3K79me2, reference D15E8, Cell Signal Technology) overnight at 4°C and the immunoprecipitation carried out using Dynabeads protein A and Dynabeads protein G. Following immunoprecipitation and washing, a reverse cross-linking step was carried out by incubating for at least six hours at 65°C in 200 mM NaCl and the samples were then digested with Proteinase K and RNAse A. Purified DNA was analysed on an Illumina HiSeq 4000 platform with a single-end sequencing primer over 50 cycles and pair-end sequencing for H3K79me2. At least 20 million reads were generated for each immunoprecipitation. The ChIP-seq dataset has been deposited in the NCBI Gene Expression Omnibus database under the accession numbers described in Table S2.

Quality control of the sequence data was carried out using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Poor quality sequences were removed and the high quality sequences trimmed with Cutadapt (52, 59). Illumina reads were mapped onto the *Ectocarpus* v2 genome (45) using Bowtie (60), which contains both male and female SDR. Duplicates were removed using samtools markdup in the Samtools package (v 1.9) (61).

Quality control of ChIP-seq data sets followed the Encode ChIP-seq guidelines and practices (62)(Table S3). ChIP-seq analysis was carried out for two biological replicates for each PTM in

both the male and female samples. Spearman correlation analysis of replicates was performed with multiBamSummary and then by plotCorrelation (v3.1.2 deepTools) (63). Replicate samples were strongly correlated (Pearson correlations >0.92, Figure S2).

To identify peaks and regions of chromatin mark enrichment in a gene-by-gene basis, each data set, after combining data for biological replicates, was analysed separately for the male and female gametophyte. Peaks corresponding to regions enriched in H3K4me3, H3K9ac and H3K27ac were identified using the MACS2 (version 2.1.1) callpeak module (minimum FDR of 0.01) and refined with the MACS2 bdgpeakcall and bdgbroadcall modules (64). H3K36me3 and H4K20me3 were analysed using SICER (v1.1) (minimum FDR of 0.01) (65, 66) with a window size of 200 bp and a gap size of 400 bp. Note that peaks associated with sex-biased, PAR and SDR genes were manually inspected to validate reproducibility between replicates. The signal was normalized using the Signal Extraction Scaling (SES) method (67).

Heatmaps, average tag graphs and coverage tracks were plotted using EaSeq (68). Chord diagrams were generated using the circlize package in R (69).

Detection of chromatin states and signatures

The six chromatin mark data sets were analyzed using ChromHMM (Ernst and Kellis, 2012) to learn a hidden Markov model and to assign chromatin states across the *Ectocarpus* genome. The bam alignment files for the six histone marks were converted into bed files with the software bedtools, option bamtobed (70). Then, ChromHMM was run using the « BinarizeBed» fonction on bed files with 200bp per bin. A single joint model was learned using data from both male and female *Ectocarpus* and using the « LearnModel» function. We started with a 17-state model we then used ChromHMM CompareModels module to compare decreasing number of states to the 16-state model. We then calculated for each of the 17 states the similarity (correlation between emission parameters) to its closets state in smaller models. A 12-state model was chosen as a point after which any further decrease in the number of states in the model resulted in states from the 17-state model being recovered with decreasing similarity.

The coverage of chromatin states in different categories of the genome was generated via the output files of ChromHMM called « Segmentation file ». An intersectBed (bedtools software) was made between states and the coordinates of all genes, allowing us to know which states overlap which genes.

Because each gene in the genome is composed of multiple emission states, we simplified our analysis by grouping similar states into five major categories based on the presence/absence of activation-associated and repression-associated marks – 'Permissive 1' for states with

mainly permissive TSS marks, 'Permissive 2' for states with mainly activation-associated mark H3K36me3, 'Silent' for states with enriched in H4K20me3 and/or H3K79me2 marks, 'Mixed' for states with a combination of permissive and repressive marks and 'Null' for absence of marks (Figure 1A). We then associated these five major categories with each gene in the *Ectocarpus* genome to assemble a series of unique combinations that we termed 'chromatin signatures'. Genes were considered to be in the null signature (S16) only when associated with no other emission state but E12. This resulted in a total of 16 distinct chromatin signatures across the *Ectocarpus* genome (S1 to S16) (Figure 1C).

Coverage for each histone PTM

The coverage for each histone PTM per chromosome was calculated using bedtools coverage where the coverage of each PTM was normalized by the size of the chromosome. The pseudoautosomal regions (PAR) and the sex-specific, non-recombining regions (SDR) of the sex chromosome were analysed separately, as in (12).

Statistical analysis

Statistical analysis was performed in R 3.6.3. Permutation tests were performed to study the differences of proportions of chromatin states in PAR and SDR genes compared to autosomal genes. We randomly subsampled 100,000 times a number of chromatin states equal to the number of PAR genes, SDR genes or both, from autosomal genes in order to perform proportion tests. We compared observed and simulated Pearson's Chi-square statistics to assess whether the observed differences in chromatin state proportions between gene sets (autosomal, SDR, PAR, SDR+PAR) were statistically due to chance. A significant p-value indicates that the observed difference in proportion is not due to chance. In order to eliminate any possible effect of transposable element (TEs) prevalence (which is different between PAR, SDR and autosomal genes) we also performed these tests using a randomized set of autosomal genes that displayed exact the same TE prevalence. Similarly, we performed permutation tests (100,000 permutations) to determine whether the distribution of chromatin states of evolutionary young genes was significantly different from that of evolutionary conserved autosomal genes with similar expression levels (within 25% of the median), separately in males and females. We then compared Pearson's Chi-square statistics between observed and simulated datasets.

We performed linear models of log2(TPM+1) as a function of chromatin signatures and the interaction between chromatin state and genomic location (i.e., autosome or PAR). We report significant p-values in bold when states significantly influence the level of expression (state S1

used as reference level). Interaction term is significant when the effect of chromatin state of expression level is significantly different for an autosomal gene compared with a PAR gene.

We used the list of evolutionary young genes identified in (47, 58). In brief, evolutionary young genes, i.e. genes that are taxonomically restricted to *Ectocarpus*, were defined as genes present in the genome of only *Ectocarpus* and having no BLASTp match (10^{-4}e value cutoff) with a range of other stramenopile genome-wide proteomes from public databases (indicating that they are likely to have evolved since the split from the most recent common ancestor): the brown algae *Cladosiphon okamuranus*, *Macrocystis pyrifera*, *Saccharina japonica*, *Scytosiphon lomentaria*, the eustigmatophyte *Nannochloropsis gaditana*, the pelagophyte *Aureococcus anophagefferene*, and the diatom *Thalassiosira pseudonana.*

GO-term analysis

Gene set enrichment analysis (GSEA) was carried out separately for each sex, grouping genes with either activation-associated (S1-S5) or repression associated (S13-S16) signatures. We used Fisher's exact Test implemented in the R package TopGO (71) to identify significantly enriched terms in biological processes. Top 30 categories (p-value<0.01) were plotted using ggplot2 package for R (Wickham, 2016) investigated enrichment in terms of molecular function ontology and report significant GO-terms with p-value < 0.01.

Results

Identification of chromatin states in males and females of *Ectocarpus*

Our previous work associated H3K4me3, H3K9ac and H3K27ac to the transcription start sites (TSS) of active genes, whereas H3K36me3 was associated with gene bodies (38). H4K20me3 was associated with repeated sequences, particularly transposons, whereas H3K79me2 peaks often covered several kilobases and included multiple genes. The presence of H3K79me2 and H4K20me3 on gene bodies correlated with decreased transcript abundance (38)). Thus, H3K4me3, H3K9ac, H3K27ac and H3K36me3 may be considered activation-associated marks and H3K79me2 and H4K20me3 repression-associated marks in *Ectocarpus*. Together, these six histone PTMs are therefore expected to provide a broad overview of the chromatin landscape in male and female *Ectocarpus*.

Near-isogenic male and female gametophyte (haploid) lines (Table S1, Figure S1) were used to generate sex-specific ChIP-seq profiles for the six histone PTMs (Table S2-3). The male and female haploid lines were generated by inbreeding over five and six generations respectively, and were virtually identical genetically except for the sex-specific region (SDR) of the sex chromosome (see methods, Table S2). We profiled at least 400 individual, clonal

gametophytes for each male and female replicate line and confirmed high reproducibility between our ChIP-seq replicates (Figure S2).

We used the ChromHMM algorithm to define twelve representative chromatin states common to males and females based on distinct combinatorial patterns of the six histone PTMs (Figure 1A, Figure S3). Emission states E1-E3 consisted of combinations of the TSS-enriched permissive marks H3K4me3, H3K9ac and H3K27ac (designated as group 'permissive 1') while emission states E4-S5 (group 'permissive 2') corresponded to regions enriched in HK36me3 (Figure 1A). Emission states E6-E8 corresponded to mixed states that all included H4K20me3 or H3K79me2 together with one or more of the activation-associated marks (group 'mixed'). Finally, silent emission states (E9-E11) were all enriched with H3K79me2 and/or H4K20me3 and H3K36me3 (group 'silent') while emission state E12 corresponded to a 'null' state that was devoid of the histone PTMs we assayed (group 'Null'). An example of the histone PTM coverage over a 602 kbp region of the *Ectocarpus* genome is shown in Figure 1B.

While ChromHMM provided a broad overview of chromatin states across the *Ectocarpus* genome, our aim was to focus on chromatin changes between females and males at the gene level. Because more than one emission state could be present over the length of a gene in a multitude of combinations, we determined which of the five above groups of emission states were represented at each gene in the *Ectocarpus* genome and then defined a total of 16 distinct combination of these groups that we herein refer to as chromatin signatures (see methods) (Figure 1C). Based on the predominant histone PTMs represented in each signature, the signatures were then classed into three groups: activation-associated, mixed and repression-associated.

Chromatin signatures of different categories of *Ectocarpus* genes

To elucidate the relationship between chromatin signatures and gene expression in *Ectocarpus*, we generated paired RNA-seq data using the same biological samples as those used for the ChIP-seq analysis (see methods). Together with previously published datasets (48, 49, 72), we defined four categories of genes based on their expression patterns: transcribed genes (TPM≥1), silent genes (TPM<1), housekeeping genes with broad expression patterns in multiple tissues and life cycle stages (tau < 0.25; see methods) and narrowly expressed genes (NEGs; tau > 0.75; see methods).

The most common chromatin signature for the transcribed genes (38.4% and 38.2% in females and males respectively) was S3, which corresponds to co-localisation of three or four of the activation-associated histone PTMs (H3K36me3, H3K27ac, H3K9ac, H3K4me3; Figure 1D, Table S4). At 'silent' genes, signature S15 containing H3K79me2 and H4K20me3 was the most common (21.7% and 21.9% in females and males, respectively; Figure 1D, Table S4).

Consistent with their ubiquitous expression, the majority of the housekeeping genes in *Ectocarpus* (49% and 48.8% in males and females, respectively) were associated with activation-associated signature S3. In contrast, NEGs were associated with a much larger proportion of mixed and repression-associated signatures compared with housekeeping genes, consistent with the restricted expression of NEGs (Figure 1D, Table S4). Finally, the relative proportion of chromatin signatures was broadly similar between males and females for each of the four gene categories (transcribed, silent, housekeeping and NEG) (Figure 1D, Table S4). Together, these data support our categorisation of activation-associated or repression-associated chromatin at *Ectocarpus* genes across the genome and suggests that the chromatin landscape remains largely stable during sexual differentiation.

Identification of histone PTMs associated with gene activation and gene repression

To further investigate the relationship between the observed chromatin signatures and gene expression, we assessed the transcript abundances of genes corresponding to each chromatin signature. Consistently, genes that were assigned to an activation-associated signature had higher transcript levels than genes with mixed signatures, whereas genes assigned to a repression-associated signature exhibited the lowest levels of expression overall (Figure 2A). A clear trend towards increasingly higher transcript abundance was correlated with the gradual acquisition of more activation-associated marks (H3K9ac, H3K27ac, H3K4me3 and H3K36me3; Figure 2A; Table S5). These observations support the proposed association of these four histone PTMs with gene activation (38) and further validate our assignment of chromatin signatures.

Conversely, genes corresponding to chromatin signatures with H4K20me3 and/or H3K79me2 consistently exhibited lower transcript levels than genes with equivalent chromatin states without H4K20me3 and H3K79me2 (Figure 2A; Table S5). For example, transcript abundance for genes with signature S13 was significantly lower than genes with signature S3 (Wilcox test, p-value < 2.22E10-16 Figure 2A; Table S5). These results are consistent with H4K20me3 and H3K79me2 being associated with repressed gene expression in *Ectocarpus*. Note however that because H4K20me3 is frequently associated with transposons (38), the observed association with reduced gene expression could also be indirect through the silencing of intronic transposon sequences. Finally, *Ectocarpus* genes associated with null signature S16, which corresponded to regions devoid of any of the assayed histone PTMs, exhibited very low transcript abundance (Figure 2A, Table S5).

Next, we compared the expression level of genes with activation-associated signatures (S1-S5) in males and females to genes with repression-associated signatures (S13-S16). Note that we did not include genes with mixed signatures in this analysis because they exhibited a

combination of both activation-associated and repression-associated marks and because they were expressed at intermediate levels (Figure 2A). As expected, genes marked with active signatures were expressed at higher levels in both sexes than those that were associated with repressive-associated signatures (Figure 2B; pair-wise Wilcox test, p-value<2.2E-16). Importantly, levels of gene expression in males and females were also significantly different for genes marked with active signatures in one sex but with repression-associated signatures in the other (Figure 2B; pair-wise Wilcox test, p-value=0.004 and p-value=0.02). Thus, despite the lack of global changes in chromatin landscape between males and females, our profiles illustrate localised chromatin signature changes associated with sex-specific gene expression.

Gene ontology (GO) term enrichment analysis showed that genes with activation-associated (S1-S5) signatures were enriched in functions related to metabolic process, whereas genes with repressive signatures (S13-S16) were enriched in functions related to signalling (Figure 2C). Interestingly, GO term enrichment appeared more stable between sexes with repressive-associated chromatin signatures, whereas sex-specific GO term enrichment was more apparent for genes with active signatures (Figure 2C). This difference was not due to repression-associated signatures being more stably maintained at genes in males and females compared to activation-associated signatures because similar proportions of genes exhibited stable maintenance of either activation-associated or repression-associated signatures in females versus males (83.1% and 81.6%; respectively; Figure 2D). Therefore, it appears that genes with activation-associated chromatin signatures exhibit more sex-specific functions than those with repressive signatures.

Chromatin signatures and sex-biased gene expression in *Ectocarpus*

To investigate the role of histone PTMs in sexual differentiation, we examined the chromatin signatures of genes with sex-biased expression. A comparison of gene expression patterns in the two near-isogenic male and female lines (Figure S1), based on RNA-seq data generated using the same biological samples as were used for the ChIP-seq analysis, identified a total of 268 genes that exhibited sex-biased expression (adjusted p-value < 0.05, fold change > 2, TPM > 1; Table S5, S6).

The presence of the active signatures was associated with higher transcript abundance for sex-biased genes in both males and females (Figure S4, S5). Sex-biased genes therefore display a similar association between activation-associated chromatin and increased gene expression levels as that observed genome-wide (Figure 2A). Interestingly, 38.2% of male-biased genes (MBGs) and 37.5% of female-biased genes (FBGs) had a different chromatin signature between males and females (Table S5), suggesting that chromatin dynamics underlie sex-biased gene expression in males and females.

Sex-biased genes tend to have narrow expression patterns (49, 73) so we compared their chromatin patterns with that of NEGs. Overall, the proportions of the different chromatin signatures were significantly different compared with NEGs suggesting that their chromatin landscape is not related to their narrower pattern of expression (Chi-square test, p-value = 4.937E-15 and p-value = 0.01608 in FBGs vs NEGs in females and males, respectively, and p-value = 5.627E-4 and p-value = 3.333E-6 for MBGs vs NEGs in females and males, respectively; Figure 1D and Figure 3A).

For the MBGs, there was a difference between the relative proportions of the different chromatin signatures in males compared to females. In males, repression-associated chromatin signatures were less frequent than in females, whereas signatures that included activation-associated marks (H3K9ac, H3K27ac, H3K4me3 and/or H3K36me3) were more common (Chi-square test p-value=0.05; Figure 3A-B; Table S4). A subset of MBG (14.5%) changed from a repression-associated or mixed signature in females to an active signature in males (Figure 3C; Table S7). A similar situation, albeit less clear, was observed for FBG, where active and mixed signatures were more frequent in females (47.4%) compared with males (38.9%), although not significantly (Chi-square test p=0.093). Conversely, a larger proportion (61%) of female-biased genes were in a repressive configuration in males compared with females (52.6%; Figure 3A-B; Table S4) but, again, not significantly (Chi-square test p=0.097). Like MBG, 12.5% of the FBG had repression-associated or mixed signatures in males whilst the same genes become associated with activation-associated in females (Figure 3C, Table S7).

Figure 3D-E shows genome browser tracks at representative FBG and MBG genes illustrating histone PTM changes during sexual differentiation. The MBG Ec12_002810, which encodes a conserved protein of unknown function, had an activation-associated chromatin signature in males (Table S5), but accumulated both H3K79me2 and H3K20me3 in females where it also exhibited decreased expression (Figure 3D). Conversely, the reduced male expression of the FBG Ec-05_003380, which encodes a peroxidase enzyme, was associated with the augmentation of an H3K79me2 domain downstream of the gene (Figure 3E). This observation suggests that H3K79me2 might undergo differential deposition between males and females. Indeed, we noted that sex-specific domains of H3K79me2 were present at 12.1% (632) and 9.1% (457) of genes in males and females, respectively (Table S8). However, the majority of these loci (97.3% and 97.2%) were not differentially transcribed between males and females, suggesting that H3K79me2 dynamics might have an indirect impact on sex-biased gene expression.

We also noticed more frequent sex-specific deposition of H3K36me3, H3K79me2 and H3K20me3 (Table S8) compared with H3K27ac H3K4me3 and H3K9ac, suggesting that the former may drive the changes in chromatin signatures observed in males versus females.

In conclusion, our analysis revealed chromatin signatures modifications that were concomitant with changes in sex-biased gene expression between males and females, with MBGs undergoing more histone PTM transitions during sexual differentiation compared with FBGs.

The chromatin landscape of the *Ectocarpus* sex chromosomes

In organisms with diploid sexual systems (XY or ZW), sex chromosomes exhibit different patterns of histone PTMs to autosomes (12, 26, 31, 74, 75). Given the nature of the *Ectocarpus* UV system, where most of the U and V sex chromosome recombines at the PAR, a markedly different chromatin landscape in sex chromosomes compared with autosomes is not expected. We investigated the chromatin signature of genes in the PAR, SDR and autosomes of *Ectocarpus* to test this hypothesis.

Surprisingly, a marked difference between sex chromosomes and autosomes was observed in *Ectocarpus* (Figure 4A, Table S5, S9-12; Figure S7, S8, S9). While the relative proportion of the 16 chromatin signatures showed some variance between autosomes, genes on the sex chromosomes exhibited a strikingly different pattern compared to autosomal genes (Figure 4A). In particular, there was a significant underrepresentation of genes with activation-associated chromatin signatures (specifically S3 and S4) on the sex chromosomes compared to the autosomes (Table S12). Furthermore, the sex chromosome was significantly enriched in signatures that included the histone PTMs H4K20me3 and H3K79me2 compared with autosomes (Figure 4A-C, Table S12).

The significantly distinct chromatin patterns between the sex chromosome and the autosomes were also observed when only the PAR was taken into account (Chi-square test p-value <2.2E-16; Figure 4A-B). For example, 52.1% and 46.7% of the PAR genes in females and males, respectively, were associated with repressive signatures compared with 25.3% and 23.8% of autosomal genes for females and males, respectively (Table S5, Table S10). Interestingly, 32% of the genes located in the PAR were found to be associated with different chromatin signatures in males and females (Table S5), indicating that a substantial proportion of the PAR genes display sex-dependent chromatin signature transitions. Note that only 11 of the 430 PAR genes were classed as sex-biased genes (Table S5), so these sex-related chromatin patterns on the PAR do not appear to be correlated with sex-biased PAR gene expression.

Analysis of the sex-determining regions of the U and V chromosomes showed that the vast majority (85%) of the genes within the female SDR (i.e., U-specific genes) were associated with repression-associated signatures, whereas this proportion was significantly less prevalent for the male SDR (where 53% of genes were associated with repressive signatures) (Chi-square

test p-value = 0.02; Figure 4B; Table S11). Therefore, male and female SDRs have distinct chromatin landscapes.

Sex chromosome features and chromatin patterns

Previous work has shown that the *Ectocarpus* PAR is enriched in transposons compared with autosomes (42, 47). Considering that in *Ectocarpus* H4K20me3 co-localizes with transposon sequences in *Ectocarpus* (38), we asked if the presence of transposons in PAR genes could explain the observed chromatin state distribution patterns. More PAR genes contained a transposon sequence compared to autosomal genes (80% versus 36%, respectively) but this did not correlate with an increased proportion of PAR genes marked with H4K20me3 (28-29% for the PAR versus 25-27% for autosomes) (Table S13). Moreover, permutation tests using subsets of autosomal genes, in which 80% of the genes were selected to contain transposons (i.e., a similar proportion of genes with transposons to that observed for the PAR) indicated that the unusual pattern of chromatin signatures in the PAR was not due simply due to the presence of additional genes with transposon insertions (Table S13).

Overall, transcript abundances for genes located in the PAR were significantly lower than for genes located on autosomes (Wilcoxon p-value<2.22E-16; Table S4, Figure 4C). This difference in expression level may potentially be explained by the different chromatin environment of the PAR and the autosomes. To test this hypothesis, we selected a subset of autosomal genes that had similar transcript levels to those of the PAR genes (Table S14). Interestingly, the distribution of chromatin signatures for this set of autosomal genes was different to that of the PAR genes with similar expression levels (Figure 4E, Figure S6) indicating that gene expression level was not the cause of the difference in chromatin signature patterns between the PAR genes and the autosomes. Moreover, the lower transcript abundance for PAR genes was consistent with a higher proportion of genes in repressive configurations compared with autosomal genes (52.9% and 46.7% for the PAR compared with 25.2% and 23.8% for the autosomes, in females and males respectively; Table S10, Table S5). Note however that even PAR genes with a permissive chromatin signature had significantly lower expression levels compared to autosomal genes with similar signatures (pairwise Wilcoxon test, p-value=1.9E-8, p-value=4.5E-10 for female and male respectively; Figure 4D). Thus, our results indicate that transcription level is not the sole cause for the striking chromatin differences between PAR genes and autosomal genes.

The PAR of *Ectocarpus* is enriched in evolutionarily young, *Ectocarpus*-restricted genes ('young' genes, see methods) (47). Young genes have unusual structural characteristics including shorter coding regions, fewer exons, lower expression levels and weaker codon bias compared with older genes (47, 76, 77). We thus asked if the presence of evolutionary young

genes might explain the distinctive chromatin configuration in the PAR compared with autosomes. Genome-wide analysis of the 4534 *Ectocarpus* young genes (58) revealed a significantly different distribution of chromatin signatures compared with more conserved genes with similar expression levels (Table S15). Almost half of the PAR genes were classed as young genes (235 out of 440), which is a significant enrichment compared to autosomes (Chi-square test p-value<2.2E-16). Moreover, the distribution of chromatin signatures at young genes within the PAR was significantly different to that of conserved genes (Chi-square test, p-value=1.57E-8 and p-value=4.26E-11 in females and males, respectively) (Figure 4E, S6; Table S10). When only evolutionarily conserved genes are included in the analysis the differences in chromatin distribution between the PAR and autosomes was considerably less marked (Table S15). The enrichment of young genes on the PAR may therefore contribute to its unique chromatin distribution.

Taken together, our observations suggest that the sex chromosome exhibits significantly different features in terms of its chromatin landscape to the autosomes, not only at the level of the non-recombining SDR region but also for the PAR. The distinct chromatin features of the PAR are not explained by the preponderance of intragenic transposons nor by lower levels of gene expression but rather by the increased incidence of evolutionarily young genes.

Chromatin signatures and expression of sex chromosome genes

Gene expression levels and deposition of chromatin marks were highly correlated for the complete set of *Ectocarpus* genes (see above, Figure 2A). For example, genes with an S3 signature enriched for all four activation-associated marks had significantly higher expression than genes with mixed signature S7 that was distinguished by the added presence of H3K79me2 and H4K20me3. However, when we analysed the association between expression level and chromatin state for genes located on the PAR of the sex chromosomes the situation was different. For example, the two chromatin signatures S4 and S6 in females had a significantly weaker correlation with expression on the PAR compared with autosomes. In males, a weaker correlation with gene expression was also observed for the PAR compared to the autosomes for signatures S4, S6, S7 and S12 (Table S16, Figure S10). In other words, depending on the location (PAR or autosomes) the correlation between chromatin signature and gene expression level was not the same.

With respect to SDR genes, the only significant correlation between expression level and chromatin signature was observed for chromatin signatures S3 and S4 in males and S15 and S16 in females (Table S17), although the small sample size of SDR genes decreases the power of our statistical testing. Note that activation-associated mark H3K36me3 was more often present at male SDR genes (17/30) than at female SDR genes (4/22), with H3K36me3 coverage

also higher on the male SDR than for the female SDR (Figure S7, S8, Table S9). We also noticed that median transcript levels for male SDR genes were higher than that of female SDR genes, although the difference was not significant (Figure 4G). In conclusion, our chromatin profiles suggest a different relationship between chromatin signature and expression levels for genes on the sex chromosomes compared with autosomes, further highlighting the unique chromatin configuration of UV sex chromosomes.

Discussion

Chromatin regulation in a haploid UV sexual system

Three types of genetic sex determination system exist in nature: XX/XY, ZZ/ZW systems and U/V systems (1, 32). For UV systems, studies have focused on understanding sex determination and sex-biased gene expression (e.g. (34, 35, 78) but we know considerably less about chromatin patterns in males compared to females. Our study provides the first overview of sex-specific differences in chromatin landscape in a haploid UV system and its relationship with sex-biased gene expression, whilst also revealing the chromatin configuration of the U and V sex chromosome.

Analysis of six histone PTMs in *Ectocarpus* males and females resulted in the definition of 16 distinct chromatin signatures associated with genes. Chromatin signatures that included different combinations of H3K4me3, H3K9ac, H3K27ac and H3K36me3 were associated with transcriptionally-permissive genes, whereas chromatin signatures that included H3K79me2 and/or H4K20me3 were associated with decrease gene expression compared to equivalent states lacking these marks. Signatures with H3K36me3 were associated with broadly expressed genes but were less prevalent on genes with narrow or tissue-specific expression, which could be related to a lower sensitivity in detecting H3K36me3 accumulation in a restricted subset of cells. The difference between housekeeping and NEG genes was considerably more marked for H3K36me3-containing signatures than for those with TSS-located marks (Table S4, S5), perhaps indicating a stronger association of this mark with gene transcription. A similar association of H3K36me3 with broadly expressed genes has been described for *Drosophila* (12, 79), indicating that this correlation has been conserved across distantly related lineages. Overall, the *Ectocarpus* chromatin patterns described here are consistent with H3K4me3, H3K9ac, H3K27ac and H3K36me3 having similar roles in brown algae, land plants and animals (38, 80–82). The role of H4K20me3, in contrast, appears to be less conserved across eukaryotic supergroups, being associated with low transcriptional levels in both animals and brown algae but with euchromatin and transcriptional activation in land plants (83, 84).

Transcriptional reprogramming during life cycle transitions often involves the loss of repressive chromatin marks (85). Our recent profiling of the *Ectocarpus* sporophyte and gametophyte indicated that H3K79me2-enriched domains, which are associated with repressed genes in *Ectocarpus*, are stably maintained between generations (38). Interestingly, here we identified several examples of genes that were associated with H3K79me2 in one sex but not in another. However, the differential loss of this mark between sexes was not correlated systematically with changes in sex-biased gene expression, suggesting that H3K79me2 reprogramming might have only an indirect impact on transcription and sexual differentiation.

Chromatin signatures of *Ectocarpus* sex-biased genes

When considered genome-wide, the proportion of genes associated with each chromatin signature did not differ substantially in males compared with females. However, when individual genes were compared, a considerable fraction was associated with different chromatin states in the two sexes, including genes that did not exhibit sex-biased expression patterns. Despite a lack of transcriptional changes in many cases, the strong correlation between chromatin signatures and gene expression argues that the chromatin reconfiguration we describe is biologically significant. One hypothesis is that the sex-specific alterations in chromatin manifest prior to any significant sex differences in transcription and phenotypic differentiation. In other words, differences in chromatin state may forecast sex-biased differences in gene expression during later stages of development, as reported for mammalian foetal germ cells (86). A more refined study using several stages during male and female gametophyte development would be needed to gain further insights into this matter.

In males, FBG were more often associated with repressive signatures than in females. Similarly, more MBGs in females were marked with repression-associated signatures compared with males, whereas FBGs were more often in permissive or mixed signatures. About 37% of sex-biased genes had different chromatin signatures in males or females, which is significantly more than chromatin changes occurring at unbiased genes. These observations support a link, at least partial, between chromatin signature, expression pattern and role of sex-biased genes during sexual differentiation in *Ectocarpus*. Sex-specific chromatin states appear not to explain the sex-biased expression patterns in Drosophila (12), and mouse (87). It appears therefore that there is no absolute correlation between SBG and chromatin landscape in animals, and sex-biased gene expression may be regulated by other mechanisms, such as distal regulatory sites (87) or involving accessibility and 3D structure of chromatin (88). Future work focusing on these alternative mechanisms in *Ectocarpus* may help to further understand the regulation of sex-biased gene expression in the brown algae.

Unique chromatin organisation features of the U and V sex chromosome

In organisms with UV sexual systems, the sex-specific SDRs of the U and V chromosome are both non-recombining, exhibit relatively similar structural features and appear to have been subjected to similar evolutionary pressures (42, 78, 78, 89–92). In *Ectocarpus*, the SDR is relatively small, these chromosomes do not exhibit strong signs of degeneration and there is no chromosome-scale dosage compensation. Consequently, we did not expect the chromatin landscape of sex chromosomes to be substantially different to that of autosomes. Surprisingly, our results provided evidence that, contrary to this prediction, the *Ectocarpus* U and V sex chromosomes have a strikingly different chromatin environment to the autosomes.

Genes in the male SDR exhibited different patterns of chromatin signatures to genes in the female SDR, with H3K36me3 and H3K79me2 in particular being enriched on the male compared with the female SDR. Interestingly, H3K36me3 deposition is usually enriched on X chromosomes in animals where it plays a key role in dosage compensation (93). Deposition of H3K36me3 is known to be associated with increased transcript abundance in plants and animals (94, 95), and, accordingly, we found that genes on the *Ectocarpus* male SDR exhibited higher expression levels than female SDR genes.

The *Ectocarpus* PAR has been shown to have unusual structural and gene expression features compared to autosomes (46, 47), and we found here unusual patterns of chromatin signatures in this genomic region. This chromatin signature configuration is not explained by reduced gene expression levels in this region, nor by a greater prevalence of transposon insertions in PAR genes. Rather, our observations suggest that evolutionarily young genes, which are enriched in the PAR compared to autosomes, shape the chromatin environment of the sex chromosome. As has been observed for young genes in animals (96), evolutionary young genes in *Ectocarpus* exhibited markedly different chromatin patterns compared with evolutionary conserved genes. It is currently unclear why young genes are more abundant in the PAR compared with other genomic regions. One possible cause is the presence of higher amounts of transposons in the PARs, which may play a role in the emergence of new genes (76). This hypothesis is supported by the fact that young PAR genes often share homology with elements in the repeated fraction of the *Ectocarpus* genome (47).

Moreover, sex-specific differences in chromatin signature were prominent on the PAR of the U and V sex chromosome, where a large proportion of genes (32%) displayed different chromatin signatures between the two sexes. Our observations emphasise the unique features of the PAR of the *Ectocarpus* UV sex chromosomes compared to autosomes, and suggests that transcription levels may depend on the genomic location of genes rather than solely the enrichment of histone PTMs, since the same chromatin signatures were transcribed

differently on autosomes compared with sex chromosomes. It is possible that the expression of genes on the U and V sex chromosomes is regulated by different chromatin processes than those that regulate autosomal gene expression, perhaps involving histone PTMs we did not assay in this study. Note that the PAR features are unlikely to be caused by linkage disequilibrium with the SDR because the PAR is considerably large compared with the SDR and recombines extensively (47). Further investigation of the chromatin landscape of UV chromosomes in a diploid stage, and in several developmental stages during the haploid-diploid life cycle, together with further profiling in other species with a UV sexual system, promises to reveal more extraordinary features of these prevalent sex chromosomes.

Figure legends

Figure 1. Histone PTMs and chromatin signatures of female and male *Ectocarpus*. A) A model of prevalent chromatin emission states found in *Ectocarpus* using ChromHMM. Permissive 1 and Permissive 2, activation-associated states; Mixed, States that mix activation-associated and repression-associated chromatin PTMs; Silent, repression-associated chromatin states; Null, absence of assayed histone PTMs. B) Representative region of the chromosome 19 showing profiles of mapped ChIP-seq reads for the six histone PTMs in females and males. Coverage is represented as the ratio of IP DNA relative to H3 for H3K36me3, H4K20me3 and H3K79me2 and input for TSS marks (H3K4me3, H3K9ac, H3K27ac). C) Chromatin signatures assigned to genes based on ChromHMM states (see methods). Percentages of the total gene set associated with each chromatin signature in males (M) and females (F) are shown to the right. D) Proportions of transcribed (TPM≥1TPM), silent (TPM<1TPM), housekeeping (tau<0.25) and narrowly expressed genes (tau>0.75) associated with each chromatin signature in males and females.

Figure 2. Gene expression and chromatin states. A) Transcript abundances for genes associated with different chromatin signatures in males and females. The colour code is the same as that used in Figure 1A.B) Transcript abundances for genes exhibiting either activation-associated (S1 to S5) or repression-associated (S13 or S16) chromatin signatures in females (dark pink) and males (grey). C) GO term enrichment for genes marked with activation-associated or repression-associated chromatin signatures in males and females. D) Venn diagrams representing the proportion of genes marked with activation-associated or repression-associated chromatin signatures in males and females.

Figure 3. Histone PTM patterns at sex-biased genes in *Ectocarpus* males and females. A) Proportions of the 16 chromatin signatures for female-biased, male-biased and unbiased genes in females (left) and males (right). The number of genes in each category are inside

brackets. Chromatin signature colour codes correspond to Figure 1C. B) Proportions of genes associated with each of the 16 chromatin signatures for female-biased (FBG) and male-biased (MBG) genes in females (left) and males (right). The intensity of the grey squares is proportional to the number of genes corresponding to each signature. Coloured squares represent the different chromatin signatures (see Figure 1A). C) Chord diagrams comparing chromatin signatures associated with female-biased (left) and male-biased (right) genes in females and males. The colour code for the chromatin states is the same as that used in Figure 1C. Each chord represents a sex-biased gene and illustrates whether a gene changes from one signature to another in the opposing sex. D) Representative chromatin profiles for a male-biased gene on chromosome 12 in females and males. E) Representative chromatin profiles for a female-biased gene in chromosome 5 in females and males.

Figure 4. Chromatin landscape of the U and V sex chromosomes compared with the autosomes. A) Chromatin signature distribution for each autosome and for the SDR and PAR regions of the sex chromosome in females (left panel) and in males (right panel). B) Proportions of genes associated with each of the 16 chromatin signatures for all autosomes and for the PAR and SDR regions of the sex chromosome in females and in males. The intensity of the grey is proportional to the number of genes in each signature. The colour code for the chromatin states is the same as that used in Figure 1A. C) Transcript abundances, measured as log2(TPM+1), for autosomal and for PAR genes in males and females. Significant differences were assessed using pairwise Wilcoxon rank sum test. D) Transcript abundances for autosomal and PAR genes associated with different chromatin signatures: activation-associated signatures (S1-S5) and repression-associated signatures (S13-S16). Significant differences were assessed using pairwise Wilcoxon rank sum test. E) Chromatin state distribution of evolutionary young genes compared with autosomal conserved genes with similar expression patterns. See also Table S13. F) Transcript abundances, measured as log2(TPM+1), for individual genes located in the female and male sex determining regions (SDRs). Coloured plots represent chromatin signatures corresponding to the colour code indicated in Figure 1C (see also Table S10). G) Transcript abundances of genes located within the female and male sex-specific regions (SDRs). Significant differences were assessed using pairwise Wilcoxon rank sum test.

Acknowledgements

We thank Carl Herrmann and Swann Floc'hlay for advice about ChIP-seq analysis, Thomas Broquet for discussions on the statistical analysis and Remy Luthringer for assistance in the

preparation of the Ectocarpus cultures. We thank the Institut Français de Bioinformatique and the Roscoff Analysis and Bioinformatics for Marine Science platform ABiMS (http://abims.sb-roscoff.fr) for providing computing and data storage resources. This work was supported by the ERC (projects 638240 and 864038), the CNRS, Sorbonne Université, the Max Planck Society and the Agence Nationale de la Recherche project Epicycle (ANR-19-CE20-0028-01).

Contributions

S.C. conceived of the project. J.G. performed the main experimental work and bioinformatic analyses. S.B. and L.T. developed protocols. M.B generated H3K79me2 ChIP-seq profiles. J.G. and M.B. defined chromatin signatures. A.L. performed GO-enrichment and RNA-seq analyses. G.C. performed statistical analyses. J.G., M.B., J.M.C. and S.C. interpreted the data. M.B. re-plotted and re-assembled the main figures. S.C. and J.M.C supervised J.G. S.C. wrote the manuscript with input from M.B. All authors read and commented the manuscript.

References

1. Bachtrog,D., Mank,J.E., Peichel,C.L., Kirkpatrick,M., Otto,S.P., Ashman,T.-L., Hahn,M.W., Kitano,J., Mayrose,I., Ming,R., *et al.* (2014) Sex determination: why so many ways of doing it? *PLoS Biol*, **12**, e1001899.

2. Charlesworth,D. (2017) Evolution of recombination rates between sex chromosomes. *Philos Trans R Soc Lond B Biol Sci*, **372**.

3. Umen,J. and Coelho,S. (2019) Algal Sex Determination and the Evolution of Anisogamy. *Annu. Rev. Microbiol.*, 10.1146/annurev-micro-020518-120011.

4. Grath,S. and Parsch,J. (2016) Sex-Biased Gene Expression. *Annu Rev Genet*, **50**, 29–44.

5. Jones,P.A. (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*, **13**, 484–492.

6. Nugent,B.M., Wright,C.L., Shetty,A.C., Hodes,G.E., Lenz,K.M., Mahurkar,A., Russo,S.J., Devine,S.E. and McCarthy,M.M. (2015) Brain feminization requires active repression of masculinization via DNA methylation. *Nat Neurosci*, **18**, 690–697.

7. Elango,N., Hunt,B.G., Goodisman,M.A.D. and Yi,S.V. (2009) DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, Apis mellifera. *Proc Natl Acad Sci U S A*, **106**, 11206–11211.

8. Schmid,M.W., Giraldo-Fonseca,A., Rovekamp,M., Smetanin,D., Bowman,J.L. and Grossniklaus,U. (2018) Extensive epigenetic reprogramming during the life cycle of Marchantia polymorpha. *Genome Biol*, **19**, 9.

9. Kouzarides,T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.

10. Lindeman,L.C., Winata,C.L., Aanes,H., Mathavan,S., Alestrom,P. and Collas,P. (2010) Chromatin states of developmentally-regulated genes revealed by DNA and histone methylation patterns in zebrafish embryos. *Int J Dev Biol*, **54**, 803–813.

11. Srivastava,S., Mishra,R.K. and Dhawan,J. (2010) Regulation of cellular chromatin state: insights from quiescence and differentiation. *Organogenesis*, **6**, 37–47.

12. Brown,E.J. and Bachtrog,D. (2014) The chromatin landscape of Drosophila: comparisons between species, sexes, and chromosomes. *Genome Res*, **24**, 1125–1137.

13. Brown,E.J., Nguyen,A.H. and Bachtrog,D. (2020) The Drosophila Y Chromosome Affects Heterochromatin Integrity Genome-Wide. *Mol Biol Evol*, **37**, 2808–2824.

14. Navratilova,P., Danks,G.B., Long,A., Butcher,S., Manak,J.R. and Thompson,E.M. (2017) Sex-specific chromatin landscapes in an ultra-compact chordate genome. *Epigenetics Chromatin*, **10**, 3.

15. Zhou,Q. and Bachtrog,D. (2015) Ancestral Chromatin Configuration Constrains Chromatin Evolution on Differentiating Sex Chromosomes in Drosophila. *PLoS Genet*, **11**, e1005331.

16. Zhou,Q., Ellison,C.E., Kaiser,V.B., Alekseyenko,A.A., Gorchakov,A.A. and Bachtrog,D. (2013) The epigenome of evolving Drosophila neo-sex chromosomes: dosage compensation and heterochromatin formation. *PLoS Biol*, **11**, e1001711.

17. Baker,B.S., Gorman,M. and Marin,I. (1994) Dosage compensation in Drosophila. *Annu Rev Genet*, **28**, 491–521.

18. Gelbart,M.E. and Kuroda,M.I. (2009) Drosophila dosage compensation: a complex voyage to the X chromosome. *Development*, **136**, 1399–1410.

19. Girton,J.R. and Johansen,K.M. (2008) Chromatin structure and the regulation of gene expression: the lessons of PEV in Drosophila. *Adv Genet*, **61**, 1–43.

20. Lemos,B., Branco,A.T. and Hartl,D.L. (2010) Epigenetic effects of polymorphic Y chromosomes modulate chromatin components, immune response, and sexual conflict. *Proc Natl Acad Sci U S A*, **107**, 15826–15831.

21. Straub,T. and Becker,P.B. (2007) Dosage compensation: the beginning and end of generalization. *Nat Rev Genet*, **8**, 47–57.

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

23. Furman,B.L.S., Metzger,D.C.H., Darolti,I., Wright,A.E., Sandkam,B.A., Almeida,P., Shu,J.J. and Mank,J.E. (2020) Sex Chromosome Evolution: So Many Exceptions to the Rules. *Genome Biol Evol*, **12**, 750–763.

24. Bachtrog,D. (2013) Y chromosome evolution: emerging insights into processes of Y chromosome degeneration. *Nature reviews. Genetics*, **14**, 113–124.

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

26. Lucchesi,J.C., Kelly,W.G. and Panning,B. (2005) Chromatin remodeling in dosage compensation. *Annu Rev Genet*, **39**, 615–651.

27. Picard,M.A.L., Cosseau,C., Ferre,S., Quack,T., Grevelding,C.G., Coute,Y. and Vicoso,B. (2018) Evolution of gene dosage on the Z-chromosome of schistosome parasites. *Elife*, **7**.

28. Vicoso,B. and Charlesworth,B. (2009) Effective population size and the faster-X effect: an extended model. *Evolution*, **63**, 2413–2426.

29. Chandler,C.H. (2017) When and why does sex chromosome dosage compensation evolve? *Ann N Y Acad Sci*, **1389**, 37–51.

30. Yasuhara,J.C. and Wakimoto,B.T. (2008) Molecular landscape of modified histones in Drosophila heterochromatic genes and euchromatin-heterochromatin transition zones. *PLoS Genet*, **4**, e16.

31. Picard,M.A.L., Vicoso,B., Roquis,D., Bulla,I., Augusto,R.C., Arancibia,N., Grunau,C., Boissier,J. and Cosseau,C. (2019) Dosage Compensation throughout the Schistosoma mansoni Lifecycle: Specific Chromatin Landscape of the Z Chromosome. *Genome Biol Evol*, **11**, 1909–1922.

32. Coelho,S.M., Gueno,J., Lipinska,A.P., Cock,J.M. and Umen,J.G. (2018) UV chromosomes and haploid sexual systems. *Trends Plant Sci*, **23**, 794–807.

33. Coelho,S.M. and Cock,J.M. (2020) Brown Algal Model Organisms. *Annu. Rev. Genet.*, **54**, 71–92.

34. Carey,S.B., Jenkins,J., Lovell,J.T., Maumus,F., Sreedasyam,A., Payton,A.C., Shu,S., Tiley,G.P., Fernandez-Pozo,N., Healey,A., *et al.* (2021) Gene-rich UV sex chromosomes harbor conserved regulators of sexual development. *Sci Adv*, **7**.

35. Iwasaki,M., Kajiwara,T., Yasui,Y., Yoshitake,Y., Miyazaki,M., Kawamura,S., Suetsugu,N., Nishihama,R., Yamaoka,S., Wanke,D., *et al.* (2021) Identification of the sex-determining factor in the liverwort Marchantia polymorpha reveals unique evolution of sex chromosomes in a haploid system. *Curr Biol*, 10.1016/j.cub.2021.10.023.

36. Montgomery,S.A., Tanizawa,Y., Galik,B., Wang,N., Ito,T., Mochizuki,T., Akimcheva,S., Bowman,J.L., Cognat,V., Maréchal-Drouard,L., *et al.* (2020) Chromatin Organization in Early Land Plants Reveals an Ancestral Association between H3K27me3, Transposons, and Constitutive Heterochromatin. *Curr Biol*, **30**, 573-588.e7.

37. Bowman,J.L., Kohchi,T., Yamato,K.T., Jenkins,J., Shu,S., Ishizaki,K., Yamaoka,S., Nishihama,R., Nakamura,Y., Berger,F., *et al.* (2017) Insights into Land Plant Evolution Garnered from the Marchantia polymorpha Genome. *Cell*, **171**, 287-304.e15.

38. Bourdareau,S., Tirichine,L., Lombard,B., Loew,D., Scornet,D., Wu,Y., Coelho,S.M. and Cock,J.M. (2021) Histone modifications during the life cycle of the brown alga Ectocarpus. *Genome Biology*, **22**, 12.

39. Coelho,S.M., Mignerot,L. and Cock,J.M. (2019) Origin and evolution of sex-determination systems in the brown algae. *New Phytologist*, **10.1111/nph.15694**.

40. Bull,J.J. (1978) Sex Chromosomes in Haploid Dioecy: A Unique Contrast to Muller's Theory for Diploid Dioecy. *The American Naturalist*, **112**, 245–250.

41. Immler,S. and Otto,S.P. (2015) The evolution of sex chromosomes in organisms with separate haploid sexes. *Evolution*, **69**, 694–708.

42. Ahmed,S., Cock,J.M., Pessia,E., Luthringer,R., Cormier,A., Robuchon,M., Sterck,L., Peters,A.F., Dittami,S.M., Corre,E., *et al.* (2014) A haploid system of sex determination in the brown alga *Ectocarpus* sp. *Curr. Biol.*, **24**, 1945–1957.

43. Baudry,L., Guiglielmoni,N., Marie-Nelly,H., Cormier,A., Marbouty,M., Avia,K., Mie,Y.L., Godfroy,O., Sterck,L., Cock,J.M., *et al.* (2020) instaGRAAL: chromosome-level quality scaffolding of genomes using a proximity ligation-based scaffolder. *Genome Biol*, **21**, 148.

44. Cock,J.M., Sterck,L., Rouzé,P., Scornet,D., Allen,A.E., Amoutzias,G., Anthouard,V., Artiguenave,F., Aury,J.-M., Badger,J.H., *et al.* (2010) The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. *Nature*, **465**, 617–621.

45. Cormier,A., Avia,K., Sterck,L., Derrien,T., Wucher,V., Andres,G., Monsoor,M., Godfroy,O., Lipinska,A., Perrineau,M.-M., *et al.* (2017) Re-annotation, improved large-scale assembly and establishment of a catalogue of noncoding loci for the genome of the model brown alga Ectocarpus. *New Phytol*, **214**, 219–232.

46. Avia,K., Lipinska,A.P., Mignerot,L., Montecinos,A.E., Jamy,M., Ahmed,S., Valero,M., Peters,A.F., Cock,J.M., Roze,D., *et al.* (2018) Genetic diversity in the UV sex chromosomes of the brown alga *Ectocarpus*. *Genes (Basel)*, **9**.

47. Luthringer,R., Lipinska,A.P., Roze,D., Cormier,A., Macaisne,N., Peters,A.F., Cock,J.M. and Coelho,S.M. (2015) The pseudoautosomal regions of the U/V sex chromosomes of the brown alga *Ectocarpus* exhibit unusual features. *Mol Biol Evol*, **32**, 2973–2985.

48. Lipinska,A.P., Toda,N.R.T., Heesch,S., Peters,A.F., Cock,J.M. and Coelho,S.M. (2017) Multiple gene movements into and out of haploid sex chromosomes. *Genome Biol*, **18**, 104.

49. Lipinska, Cormier,A., Luthringer,R., Peters,A.F., Corre,E., Gachon,C.M.M., Cock,J.M. and Coelho,S.M. (2015) Sexual dimorphism and the evolution of sex-biased gene expression in the brown alga *Ectocarpus*. *Mol Biol Evol*, **32**, 1581–1597.

50. Wilson Sayres,M.A. (2018) Genetic Diversity on the Sex Chromosomes. *Genome Biology and Evolution*, **10**, 1064–1078.

51. 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.

52. Martin,M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet*, https://doi.org/10.14806/ej.17.1.200.

53. Montecinos,A., Valero,M., Guillemin,M.-L., Peters,A., Desrut,A. and Couceiro,L. (2017) Species delimitation and phylogeographic analyses in the Ectocarpus subgroup siliculosi (Ectocarpales, Phaeophyceae). *Journal of Phycology*, **53**, 17–31.

54. Kim,D., Pertea,G., Trapnell,C., Pimentel,H., Kelley,R. and Salzberg,S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.*, **14**, R36.

55. Liao,Y., Smyth,G.K. and Shi,W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930.

56. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, **15**, 550.

57. Yanai,I., Benjamin,H., Shmoish,M., Chalifa-Caspi,V., Shklar,M., Ophir,R., Bar-Even,A., Horn-Saban,S., Safran,M., Domany,E., *et al.* (2005) Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics*, **21**, 650–659.

58. Lipinska,A.P., Serrano-Serrano,M.L., Cormier,A., Peters,A.F., Kogame,K., Cock,J.M. and Coelho,S.M. (2019) Rapid turnover of life-cycle-related genes in the brown algae. *Genome Biol*, **20**, 35.

59. Hansen,P., Hecht,J., Ibn-Salem,J., Menkuec,B.S., Roskosch,S., Truss,M. and Robinson,P.N. (2016) Q-nexus: a comprehensive and efficient analysis pipeline designed for ChIP-nexus. *BMC Genomics*, **17**, 873.

60. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*, **10**, R25.

61. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., Durbin,R., and 1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

62. Landt,S.G., Marinov,G.K., Kundaje,A., Kheradpour,P., Pauli,F., Batzoglou,S., Bernstein,B.E., Bickel,P., Brown,J.B., Cayting,P., *et al.* (2012) ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Research*, **22**, 1813–1831.

63. Ramirez,F., Dundar,F., Diehl,S., Gruning,B.A. and Manke,T. (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res*, **42**, W187-191.

64. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W., *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.

65. Xu,S., Grullon,S., Ge,K. and Peng,W. (2014) Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells. *Methods Mol Biol*, **1150**, 97–111.

66. Zang,C., Schones,D.E., Zeng,C., Cui,K., Zhao,K. and Peng,W. (2009) A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics*, **25**, 1952–1958.

67. Diaz,A., Park,K., Lim,D.A. and Song,J.S. (2012) Normalization, bias correction, and peak calling for ChIP-seq. *Stat Appl Genet Mol Biol*, **11**, Article 9.

68. Lerdrup,M., Johansen,J.V., Agrawal-Singh,S. and Hansen,K. (2016) An interactive environment for agile analysis and visualization of ChIP-sequencing data. *Nat. Struct. Mol. Biol.*, **23**, 349–357.

69. Krzywinski,M., Schein,J., Birol,I., Connors,J., Gascoyne,R., Horsman,D., Jones,S.J. and Marra,M.A. (2009) Circos: an information aesthetic for comparative genomics. *Genome Res.*, **19**, 1639–1645.

70. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.

71. Alexa,A. and Rahnenfuhrer,J. (2020) Enrichment Analysis for Gene Ontology.

72. Lipinska,A.P., D'hondt,S., Van Damme,E.J. and De Clerck,O. (2013) Uncovering the genetic basis for early isogamete differentiation: a case study of *Ectocarpus siliculosus*. *BMC Genomics*, **14**, 909–909.

73. Meisel,R.P., Malone,J.H. and Clark,A.G. (2012) Disentangling the relationship between sex-biased gene expression and X-linkage. *Genome Res*, **22**, 1255–1265.

74. Bisoni,L., Batlle-Morera,L., Bird,A.P., Suzuki,M. and McQueen,H.A. (2005) Female-specific hyperacetylation of histone H4 in the chicken Z chromosome. *Chromosome Res*, **13**, 205–214.

75. Brockdorff,N. and Turner,B.M. (2015) Dosage compensation in mammals. *Cold Spring Harb Perspect Biol*, **7**, a019406.

76. Arendsee,Z.W., Li,L. and Wurtele,E.S. (2014) Coming of age: orphan genes in plants. *Trends Plant Sci*, **19**, 698–708.

77. Palmieri,N., Kosiol,C. and Schlötterer,C. (2014) The life cycle of Drosophila orphan genes. *Elife*, **3**, e01311.

78. McDaniel,S.F., Neubig,K.M., Payton,A.C., Quatrano,R.S. and Cove,D.J. (2013) Recent gene-capture on the UV sex chromosomes of the moss Ceratodon purpureus. *Evolution*, **67**, 2811–2822.

79. Filion,G.J., van Bemmel,J.G., Braunschweig,U., Talhout,W., Kind,J., Ward,L.D., Brugman,W., de Castro,I.J., Kerkhoven,R.M., Bussemaker,H.J., *et al.* (2010) Systematic

protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell*, **143**, 212–224.

80. Baroux,C., Raissig,M.T. and Grossniklaus,U. (2011) Epigenetic regulation and reprogramming during gamete formation in plants. *Current Opinion in Genetics & Development*, **21**, 124–133.

81. Margueron,R. and Reinberg,D. (2010) Chromatin structure and the inheritance of epigenetic information. *Nat Rev Genet*, **11**, 285–296.

82. She,W. and Baroux,C. (2015) Chromatin dynamics in pollen mother cells underpin a common scenario at the somatic-to-reproductive fate transition of both the male and female lineages in Arabidopsis. *Front Plant Sci*, **6**, 294.

83. de la Paz Sanchez,M. and Gutierrez,C. (2009) *Arabidopsis* ORC1 is a PHD-containing H3K4me3 effector that regulates transcription. *Proc Natl Acad Sci USA*, **106**, 2065.

84. Fischer,A., Hofmann,I., Naumann,K. and Reuter,G. (2006) Heterochromatin proteins and the control of heterochromatic gene silencing in Arabidopsis. *J Plant Physiol*, **163**, 358–368.

85. Vigneau,J. and Borg,M. (2021) The epigenetic origin of life history transitions in plants and algae. *Plant Reprod*, **34**, 267–285.

86. Lesch,B.J. and Page,D.C. (2013) Sex-specific chromatin states in mammalian fetal germ cells. *Epigenetics Chromatin*, **6**, P45–P45.

87. Sugathan,A. and Waxman,D.J. (2013) Genome-wide analysis of chromatin states reveals distinct mechanisms of sex-dependent gene regulation in male and female mouse liver. *Mol Cell Biol*, **33**, 3594–3610.

88. Kukurba,K.R., Parsana,P., Balliu,B., Smith,K.S., Zappala,Z., Knowles,D.A., Favé,M.-J., Davis,J.R., Li,X., Zhu,X., *et al.* (2016) Impact of the X Chromosome and sex on regulatory variation. *Genome Res*, **26**, 768–777.

89. Mignerot,L. and Coelho,S.M. (2016) The origin and evolution of the sexes: Novel insights from a distant eukaryotic linage. *C R Biol*, **339**, 252–257.

90. Ferris,P., Olson,B.J., De Hoff,P.L., Douglass,S., Diaz-Cano,D.C., Prochnik,S., Geng,S., Rai,R., Grimwood,J., Schmutz,J., *et al.* (2010) Evolution of an Expanded Sex Determining Locus in Volvox. *Science (New York, N.Y.)*, **328**, 351–354.

91. 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.

92. Okada,S., Sone,T., Fujisawa,M., Nakayama,S., Takenaka,M., Ishizaki,K., Kono,K., Shimizu-Ueda,Y., Hanajiri,T., Yamato,K.T., *et al.* (2001) The Y chromosome in the liverwort

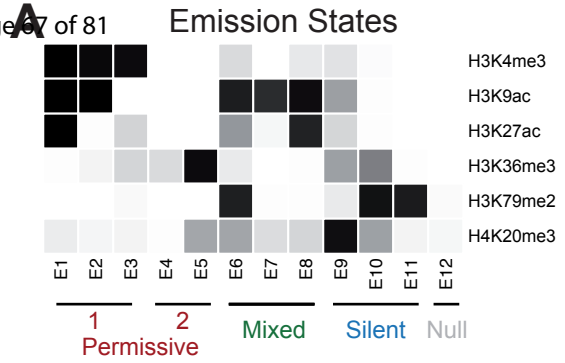
Marchantia polymorpha has accumulated unique repeat sequences harboring a male-specific gene. *Proceedings of the National Academy of Sciences*, **98**, 9454–9459.

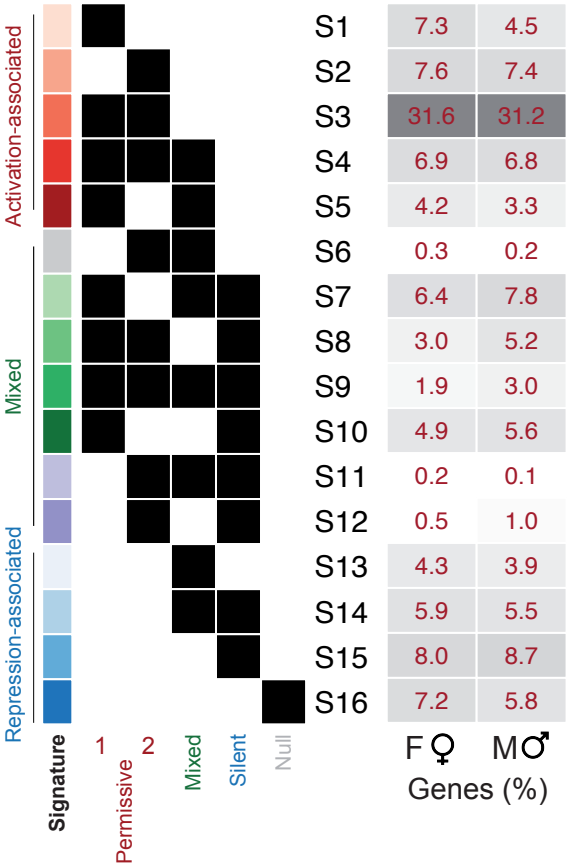
93. Bell,O., Conrad,T., Kind,J., Wirbelauer,C., Akhtar,A. and Schubeler,D. (2008) Transcription-coupled methylation of histone H3 at lysine 36 regulates dosage compensation by enhancing recruitment of the MSL complex in Drosophila melanogaster. *Mol Cell Biol*, **28**, 3401–3409.

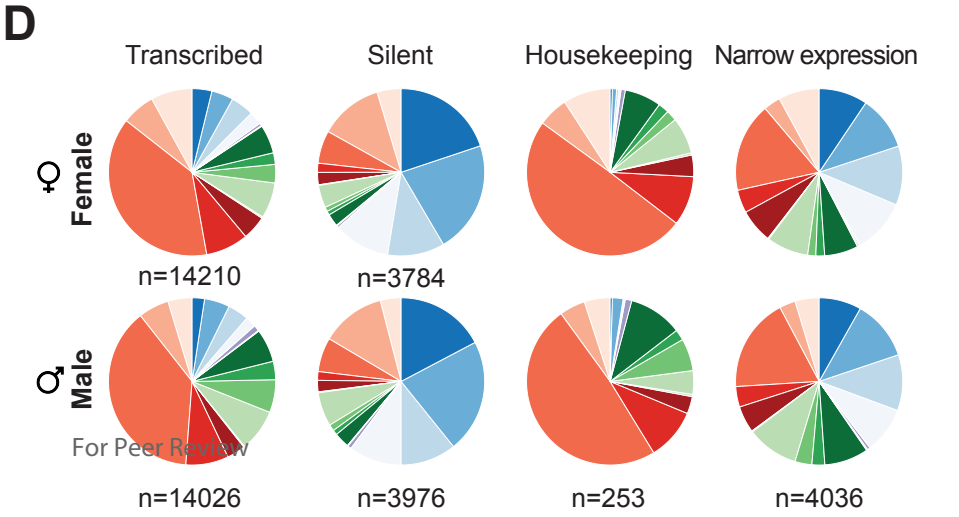
94. Roudier,F., Ahmed,I., Bérard,C., Sarazin,A., Mary-Huard,T., Cortijo,S., Bouyer,D., Caillieux,E., Duvernois-Berthet,E., Al-Shikhley,L., *et al.* (2011) Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *The EMBO Journal*, **30**, 1928–1938.

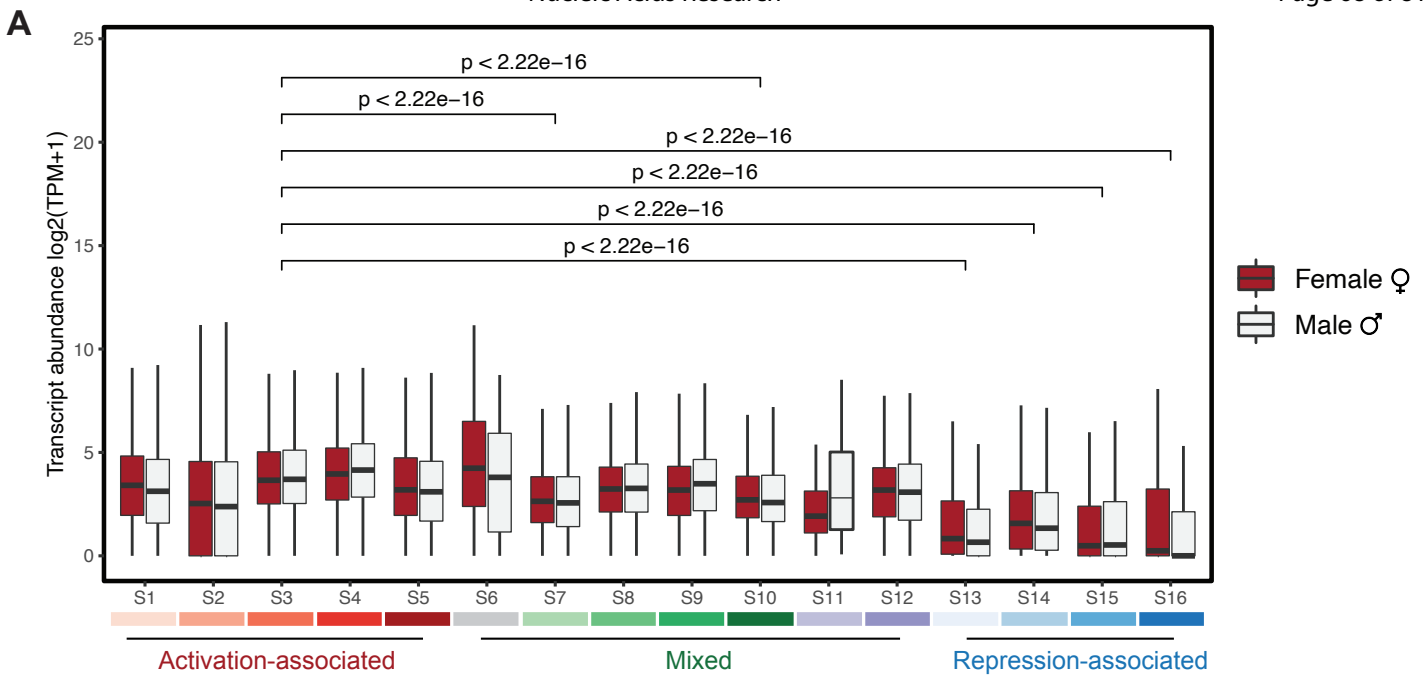
95. Shilatifard,A. (2006) Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem*, **75**, 243–269.

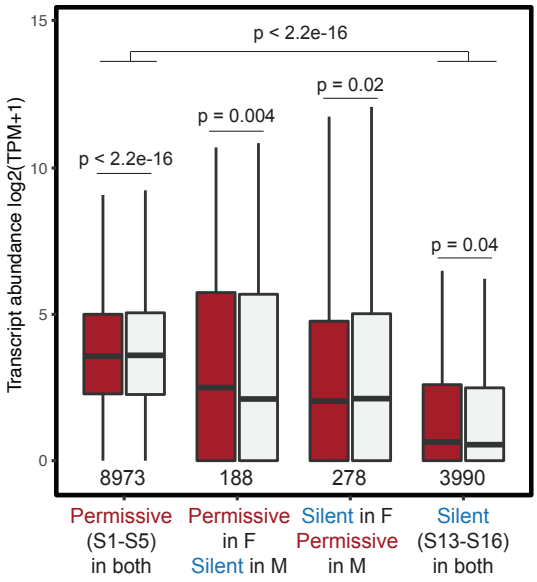
96. Werner,M.S., Sieriebriennikov,B., Prabh,N., Loschko,T., Lanz,C. and Sommer,R.J. (2018) Young genes have distinct gene structure, epigenetic profiles, and transcriptional regulation. *Genome Research*, **28**, 1675–1687.

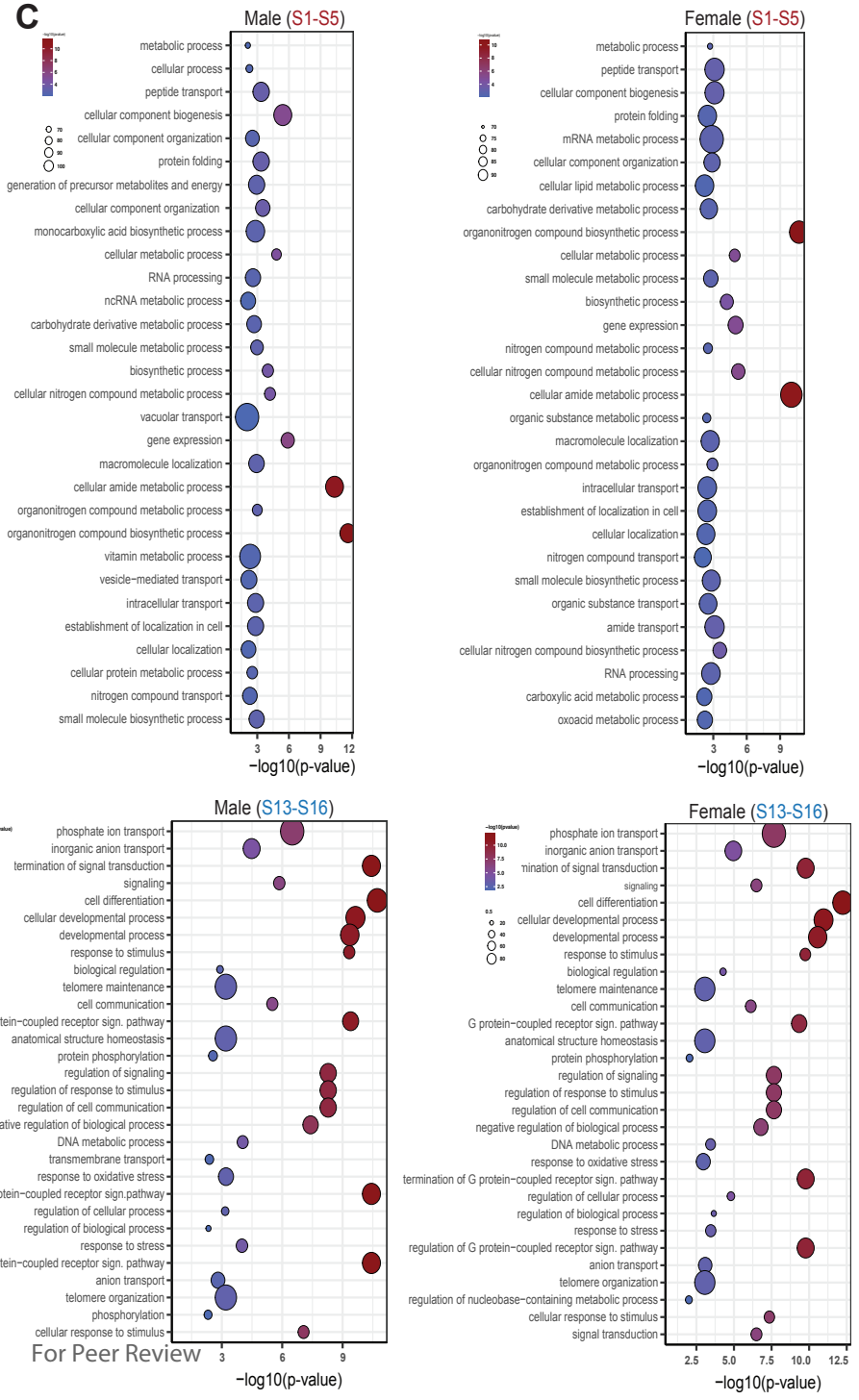


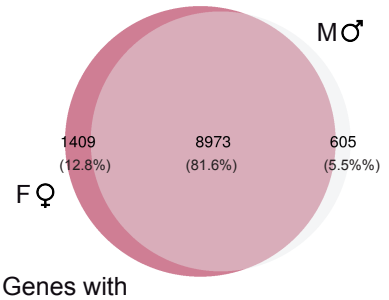




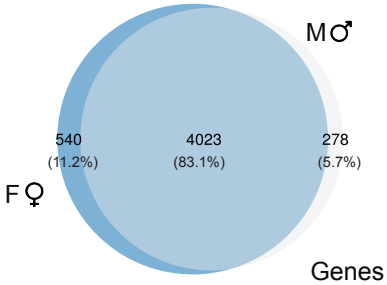




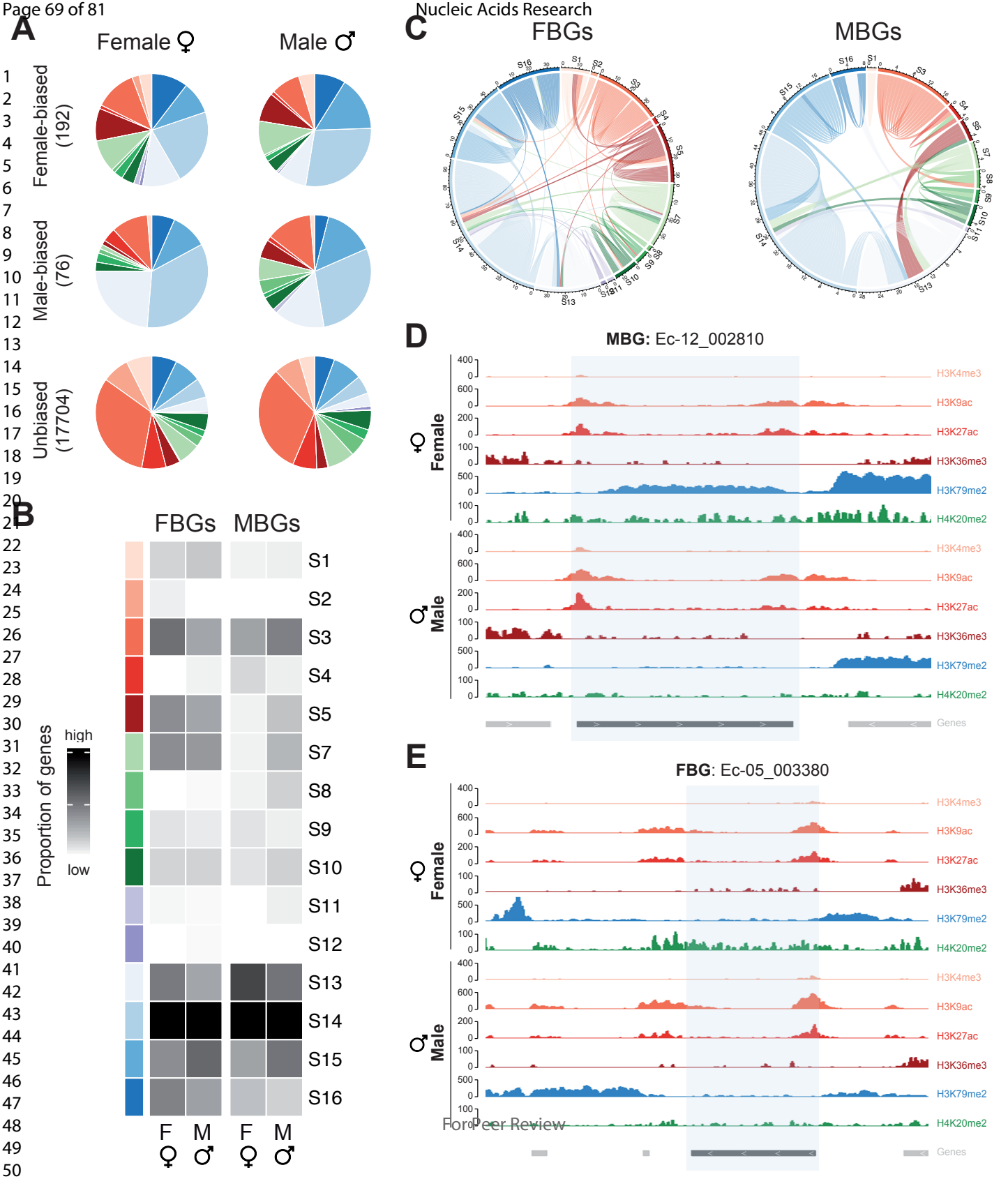


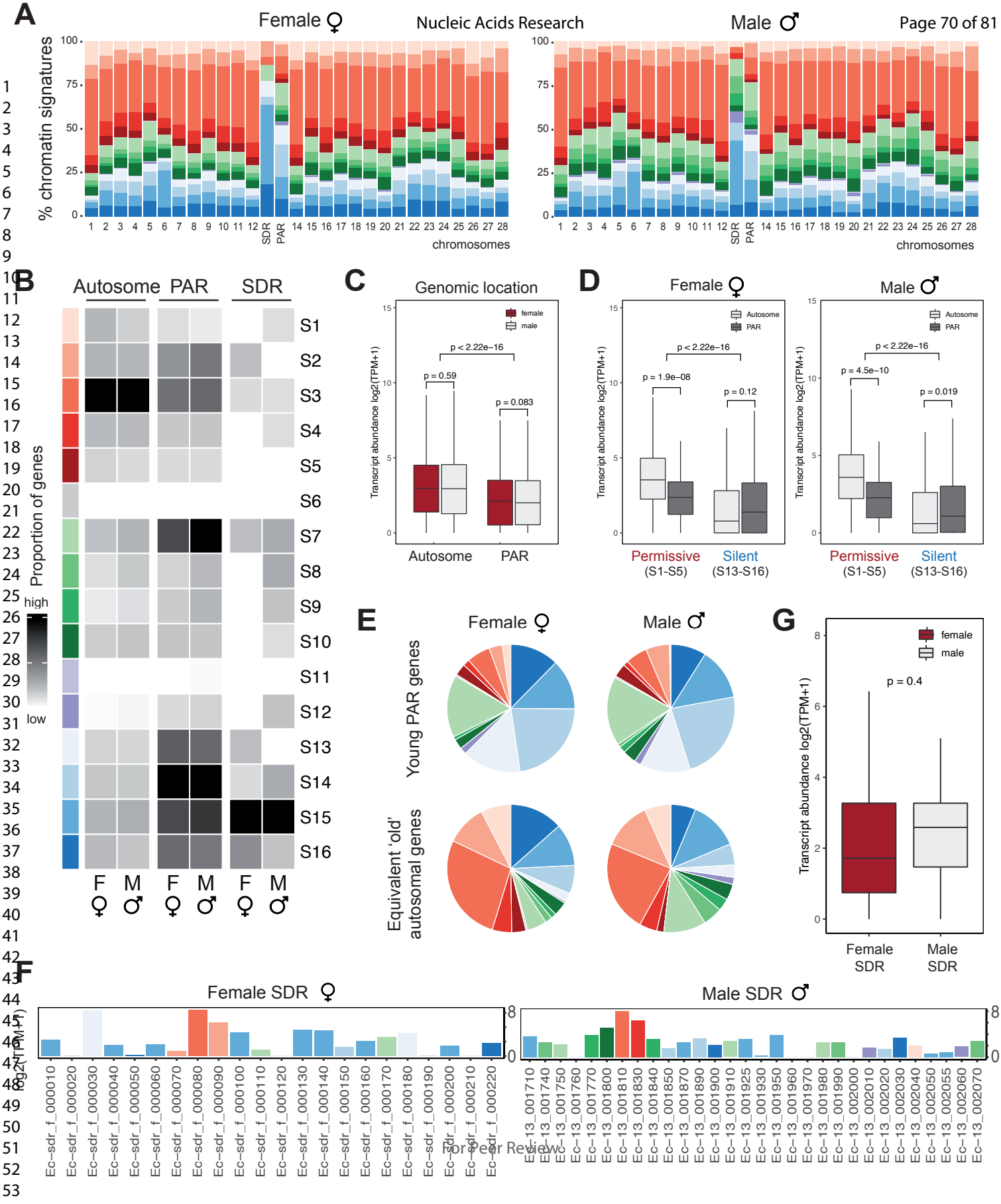


Genes with a permissive signature



Genes with a silent signature





Supplemental Figures Legends

Figure S1. Pedigree of the male and female strains used in this study. SP, sporophyte; m, male gametophyte; f, female gametophyte.

Figure S2. Spearman correlation scores for comparisons of the genomic distributions of ChIP-seq signal peaks for the six histone PTMs. Rep1, replicate 1; Rep2, replicate 2.

Figure S3. Chromatin state annotation in males and females using ChromHMM.

Figure S4. Abundances of the transcripts of sex-biased genes (SBG) marked with different chromatin signatures in females and males. Abundances of transcripts of SBGs in different chromatin signatures in females (pink) and males (blue). Values in brackets indicate the number of genes analysed.

Figure S5. Abundances of transcripts of SBGs associated with each of the different chromatin signatures in males and females. The colour code is the same as that used in Figure 1A. The total number of SBGs associated with each signature are indicated in brackets.

Figure S6. Proportions of chromatin states for PAR genes compared with the proportions of chromatin states for a set of autosomal genes with a similar pattern of expression levels to the PAR genes.

Figure S7. Percentage of coverage for specific histone PTMs for the SDRs, PAR and autosomes in male and females. Scatter plot showing the percent of coverage (in base pairs) for each of the five histone PTMs, H3K4me3, H3K9ac, H3K27ac, H3K36me3, H3K79me2 and H4K20me3. Light blue and light pink represent coverage in male and female, respectively. Dark blue and red dots correspond to coverage for the V and U sex chromosomes, respectively. Light shading indicates the two PARs and dark shading the non-recombining, sex specific region (SDR) of the sex chromosome (chromosome 13).

Figure S8. Coverage (represented as percentage of base pairs) in three different genomic regions (PAR, SDR and autosomes) marked with different histone PTMs in females (left) and males (right).

Figure S9. Distribution of ChromHMM emission states across the sex chromosome and two representative autosomes. The colour code is the same as that used in Figure 1A.

Figure S10. Transcript abundances, measured as log2(TPM+1), for PAR genes associated with different chromatin signatures in males and females. The colour code is the same as that used in Figure 1C.

Supplemental Tables Legends

Table S1. *Ectocarpus* strains used, RNA-seq sequencing statistics and SRA accession numbers.

Table S2. SNPs between male and female lines.

Table S3. Sequencing statistics for the ChIP-seq analysis and GEO reference for the dataset. N. peaks, number of peaks; FRiP, fraction of reads in peaks.

Table S4. Percentages of genes associated with each of the 16 chromatin signatures for different gene sets in males and females. Global, all genes in the genome; Transcribed genes, genes with TPM > 1; Silent genes, genes with TPM <1; Housekeeping and Narrowly-expressed, genes with tau <0.75 and tau >0.75, respectively; Unbiased, no sex-biased expression. For the chromatin signatures, refer to Figure 1A.

Table S5. Chromatin signatures (S1-S16) and transcript abundances (measured as TPM) for all *Ectocarpus* genes in males and females. FBG, female-biased gene; MBG, male-biased gene. For the chromatin states, refer to Figure 1C.

Table S6. Number of sex-biased genes in each of the chromatin signatures S1-S16 in males and females. FBG, female-biased gene, MBG, male-biased gene.

Table S7. Transitions between chromatin signatures observed for male-biased and female-biased genes in males compared with females. For chromatin signatures, refer to Figure 1C.

Table S8. Proportion of genes marked with each of the histone PTMs in males in females in different gene categories. Histone PTM peaks in each gene were analysed based on peak callers MACS2 and Sicer (see methods for details).

Table S9. Coverage of the six histone PTMs across male and female genomes. The sex chromosome (chromosome 13) is divided into PAR1 (pseudo-autosomal region 1), SDR (sex-determining region) and PAR2 (pseudo-autosomal region 2).

Table S10. Chromatin signatures of PAR genes in males and females.

Table S11. Chromatin signatures and transcript abundances (log2TPM+1) for SDR genes (see also Figure 4F).

Table S12. Permutation tests performed to determine whether the relative proportions of the different chromatin signatures were statistically different in different regions of the genome. We randomized the genomic location of autosomal genes 100,000 times and tested the difference between the observed proportions for the SDR, the PAR or the entire sex chromosome and the permuted gene sets using Chi-square statistics. Tests were performed independently for each chromatin state. Significant p-values (<0.01) are highlighted in bold.

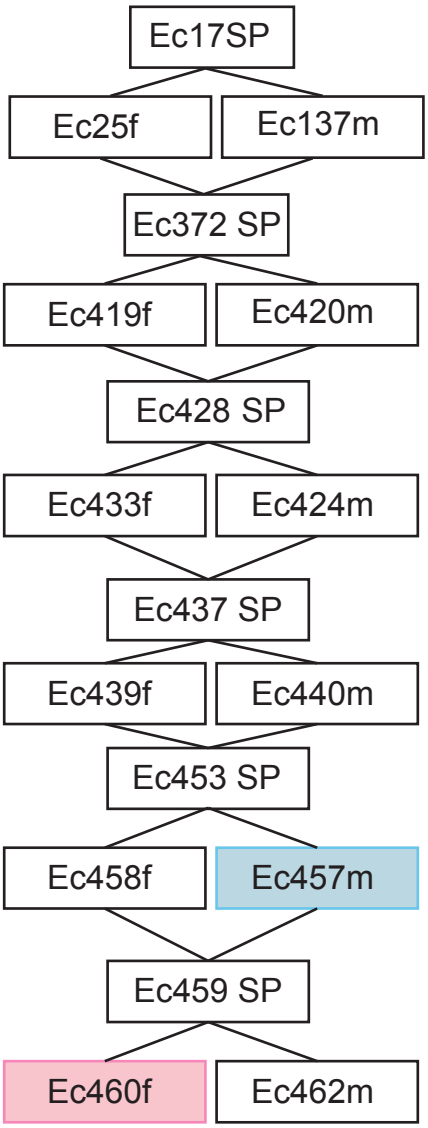
Table S13. The presence of transposon sequences in the majority (80%) of PAR genes does not explain the distinct chromatin landscape of the PAR. Correlation between the presence of transposable elements within introns and the presence of H4K20me3 in PAR genes and autosomal genes (left table). Permutation tests comparing the proportion of each chromatin signature in the PAR with the proportion of that signature in 100,000 samples of 430 autosomal genes with transposon sequences in 80% of the genes. For most chromatin signatures, the proportion on the PAR was significantly different from those of the autosomal gene samples indicating that transposon content does not explain the unusual pattern of chromatin states observed for the PAR. Significant p-values (<0.01) are highlighted in bold (right table).

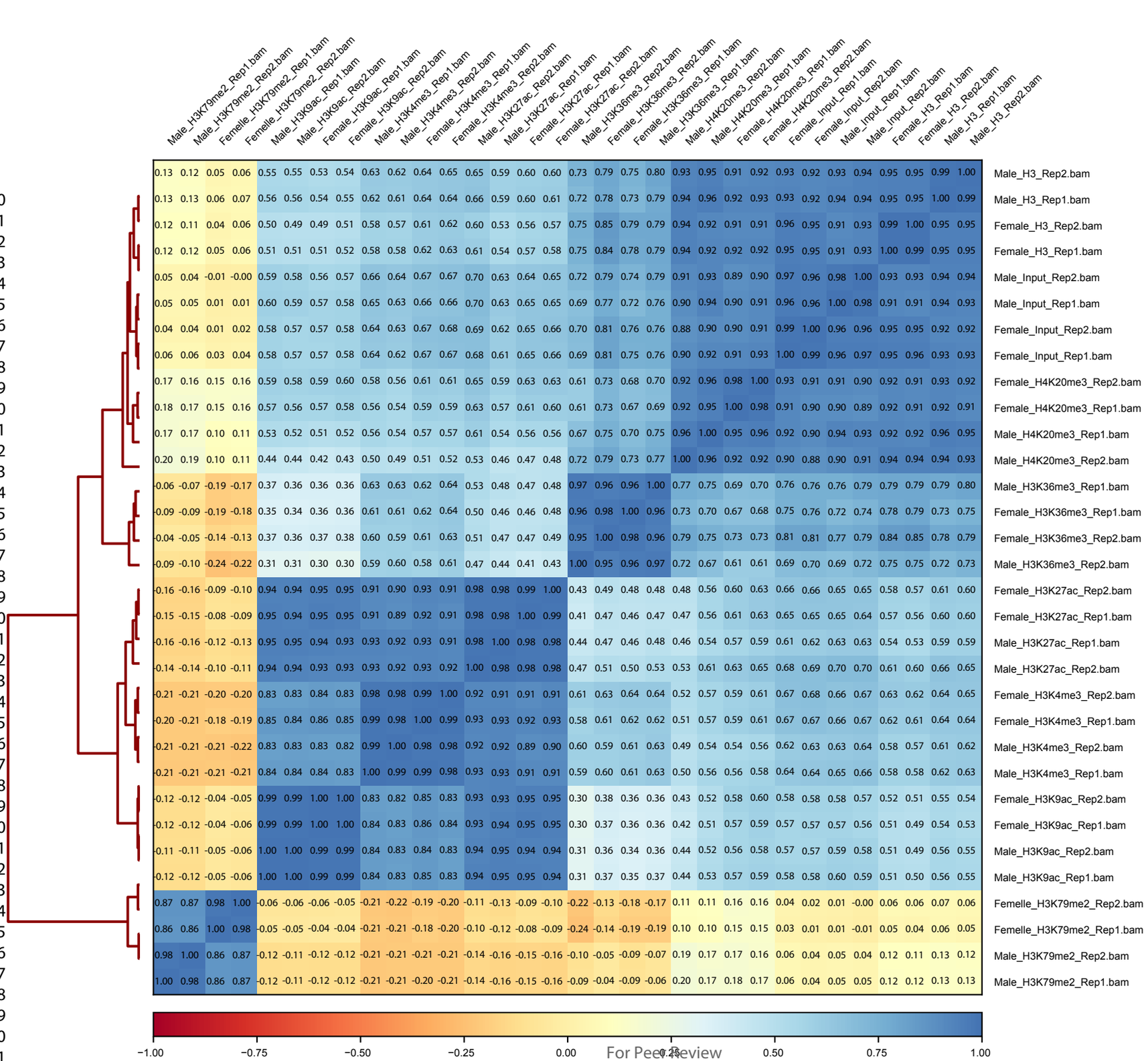
Table S14. Comparison of chromatin signatures of the PAR genes with those of a set of autosomal genes with a similar pattern of gene expression levels. To establish the autosomal gene set, for each PAR gene, the full set of autosomal genes was searched for the gene that had the most similar level of expression. When the TPM of the PAR gene was zero, an autosomal gene with a TPM of zero was selected at random. Figure S6 presents the proportions of chromatin signatures associated with the two gene sets.

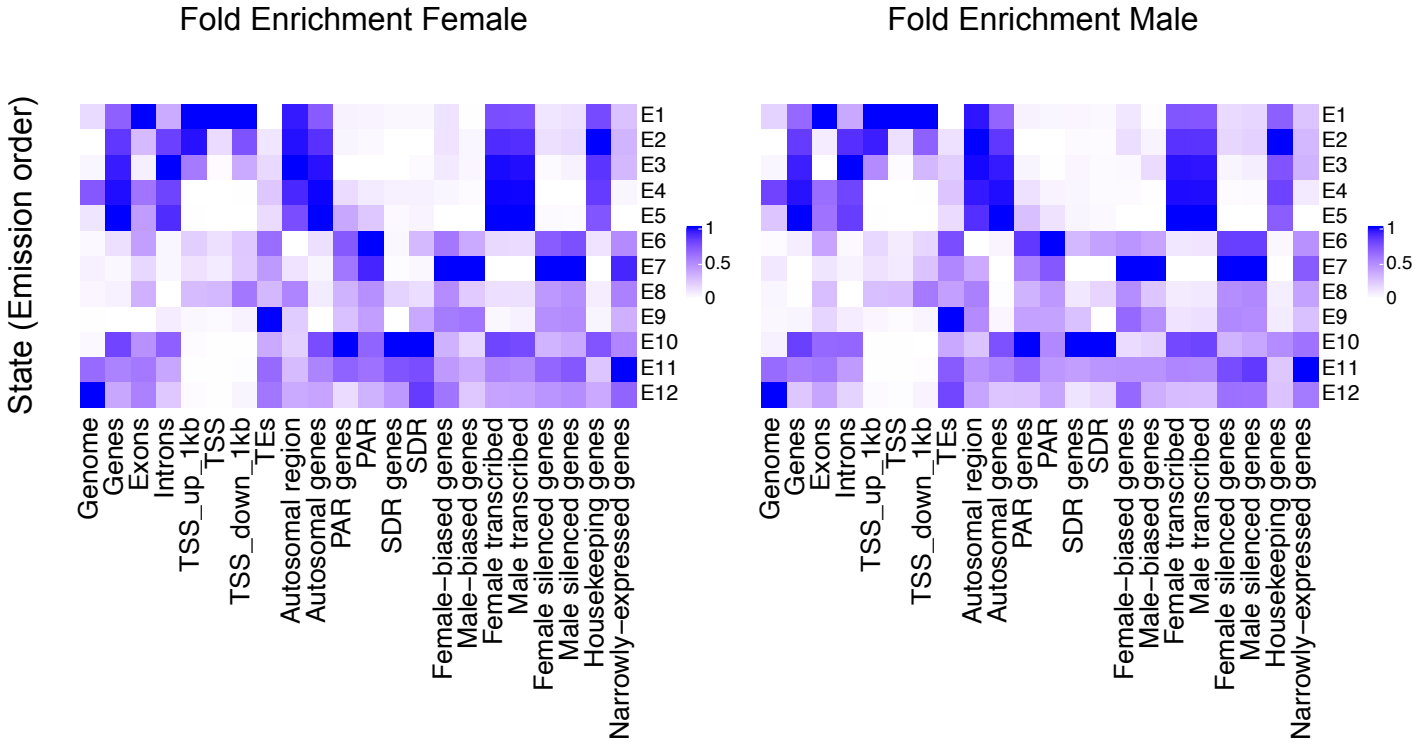
Table S15. Permutation analysis to test whether there was a significant difference between the distribution of chromatin signatures in young genes compared with evolutionary conserved genes with similar expression levels (left table) and to test whether the distribution of signatures of PAR genes that are evolutionary conserved are different from autosomal genes (right table). Significant differences are highlighted in bold (p-value <0.01).

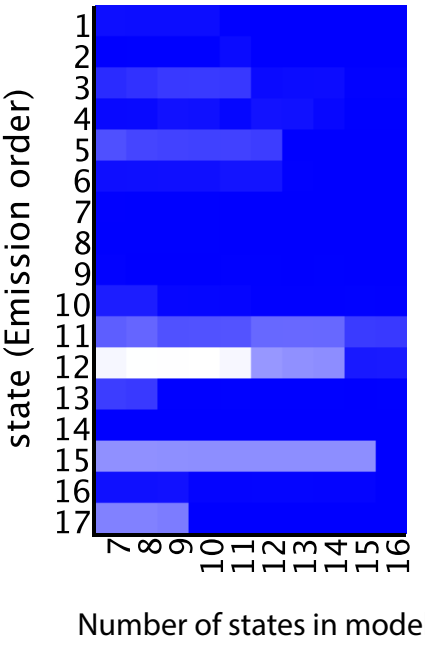
Table S16. Linear models to test whether there was a significant correlation between expression level (log2(TPM + 1)) and chromatin state (upper table) or to test whether location of a gene on the PAR or on an autosome significantly influenced the expression level associated with each chromatin signature (bottom table). Significant interaction terms, in bold, represent a significantly different effect of the chromatin signature on gene expression level in the PAR region compared to autosomal genes (p-value <0.05).

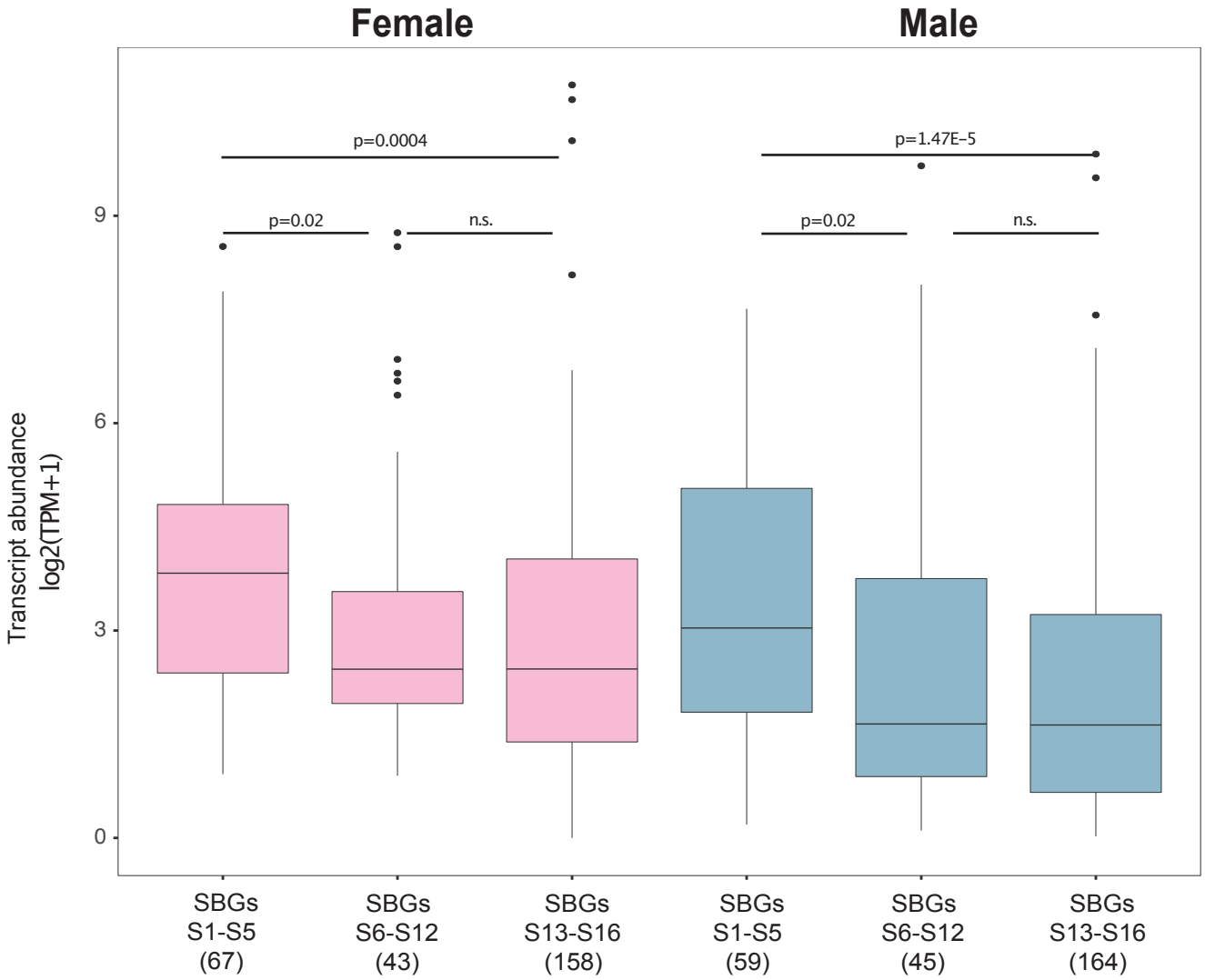
Table S17. Linear models to test whether there was a significant correlation between expression level (log2(TPM + 1)) and chromatin signature for SDR genes.

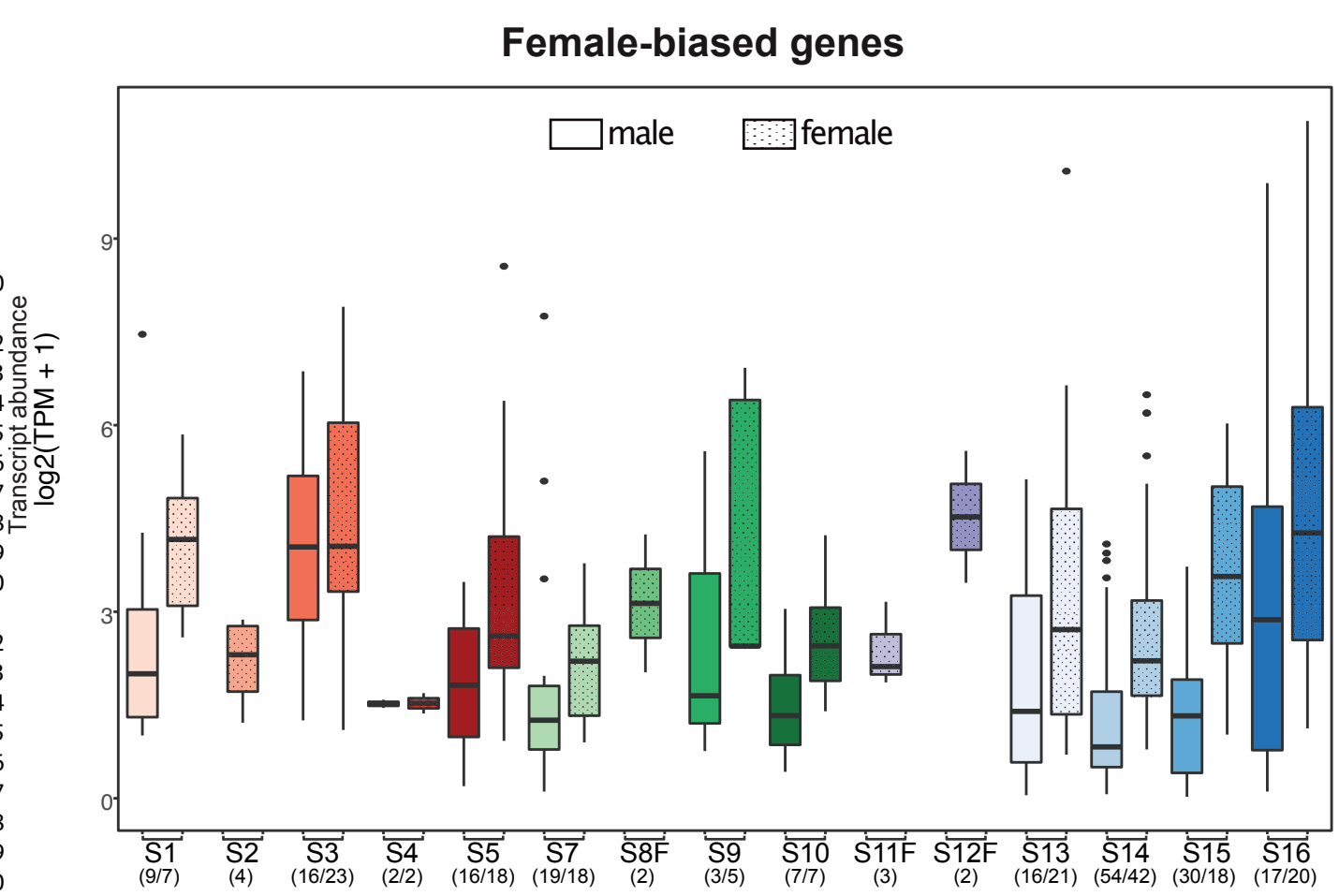

Ec17SP
Ec25f
Ec137m
Ec372 SP
Ec419f
Ec420m
Ec428 SP
Ec433f
Ec424m
Ec437 SP
Ec439f
Ec440m
Ec453 SP
Ec458f
Ec457m
Ec459 SP
Ec460f
Ec462m

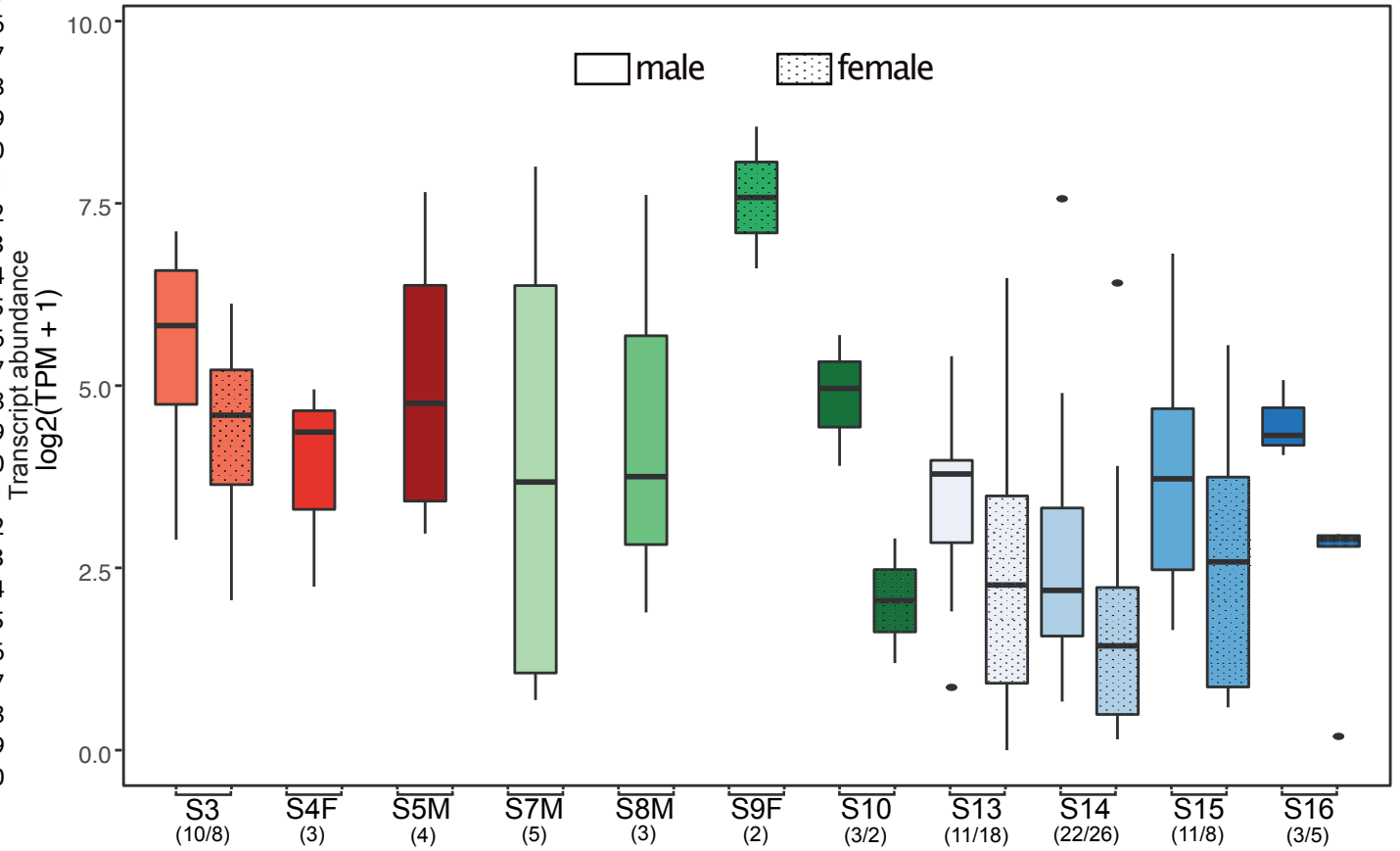


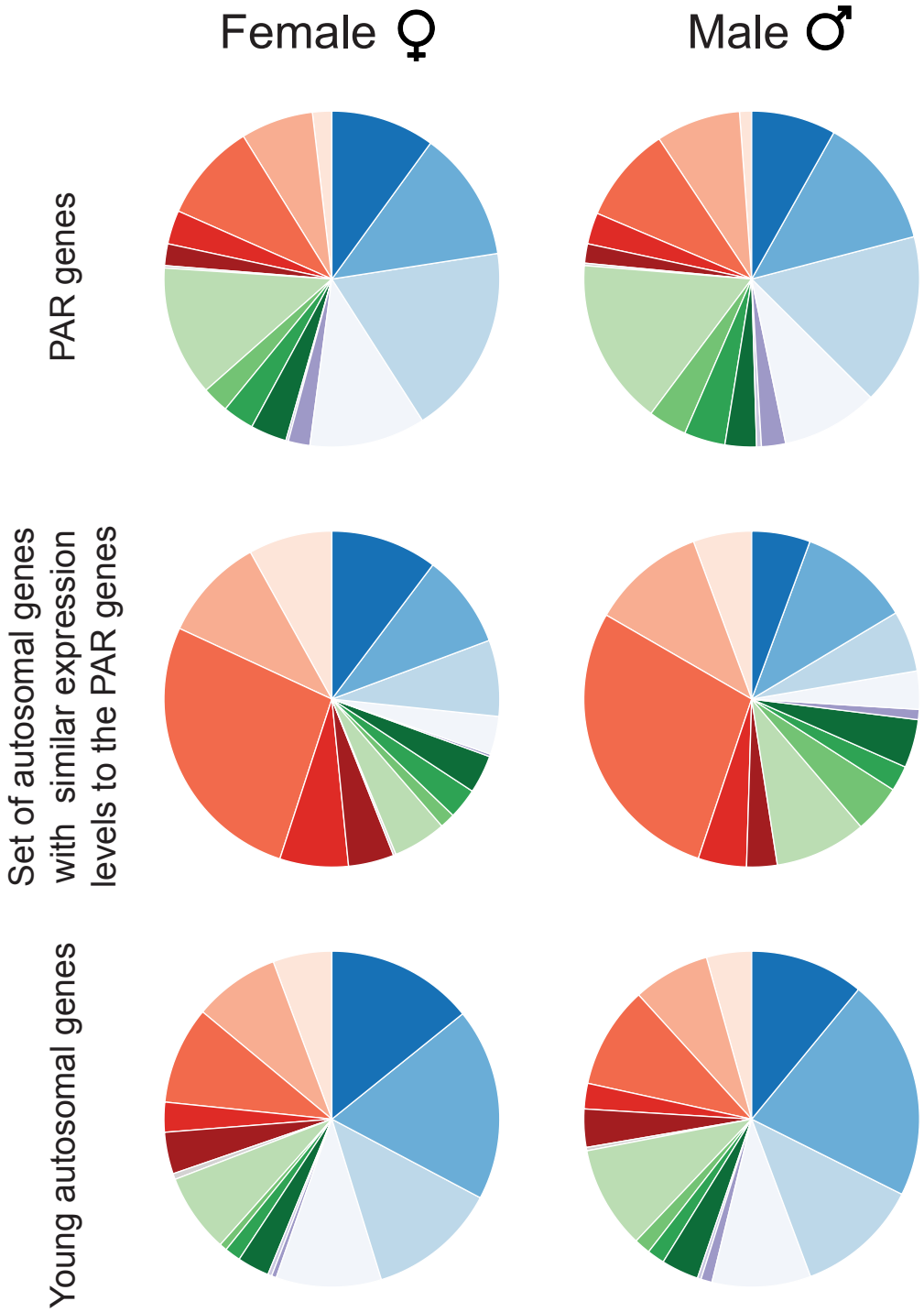

Fold Enrichment Female
Fold Enrichment Male
State (Emission order)
E1
E2
E3
E4
E5
E6
E7
E8
E9
E10
E11
E12
1
0.5
0
Genome
Genes
Exons
Introns
TSS_up_1kb
TSS
TSS_down_1kb
TEs
Autosomal region
Autosomal genes
PAR genes
PAR
SDR genes
SDR
Female−biased genes
Male−biased genes
Female transcribed
Male transcribed
Female silenced genes
Male silenced genes
Housekeeping genes
Narrowly−expressed genes


state (Emission order)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
7
8
9
10
11
12
13
14
15
16
Number of states in model


Female
Male
p=0.0004
p=1.47E−5
p=0.02
n.s.
p=0.02
n.s.
Transcript abundance
log2(TPM+1)
9
6
3
0
SBGs
S1-S5
(67)
SBGs
S6-S12
(43)
SBGs
S13-S16
(158)
SBGs
S1-S5
(59)
SBGs
S6-S12
(45)
SBGs
S13-S16
(164)


Female-biased genes
male
female
Transcript abundance
log2(TPM + 1)
0
3
6
9
S1 (9/7)
S2 (4)
S3 (16/23)
S4 (2/2)
S5 (16/18)
S7 (19/18)
S8F (2)
S9 (3/5)
S10 (7/7)
S11F (3)
S12F (2)
S13 (16/21)
S14 (54/42)
S15 (30/18)
S16 (17/20)


Male-biased genes
male
female
Transcript abundance
log2(TPM + 1)
0.0
2.5
5.0
7.5
10.0
S3 (10/8)
S4F (3)
S5M (4)
S7M (5)
S8M (3)
S9F (2)
S10 (3/2)
S13 (11/18)
S14 (22/26)
S15 (11/8)
S16 (3/5)


Female ♀
Male ♂
PAR genes
Set of autosomal genes with similar expression levels to the PAR genes
Young autosomal genes