



HAL
open science

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

► **To cite this version:**

Josselin Gueno, Michael Borg, Simon Bourdareau, Guillaume Cossard, Olivier Godfroy, et al.. Chromatin landscape associated with sexual differentiation in a UV sex determination system. *Nucleic Acids Research*, 2022, 50 (6), pp.3307-3322. 10.1093/nar/gkac145 . hal-03608518

HAL Id: hal-03608518

<https://cnrs.hal.science/hal-03608518v1>

Submitted on 14 Mar 2022

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

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



Distributed under a Creative Commons Attribution 4.0 International License

1 Chromatin landscape associated with sexual differentiation in a 2 UV sex determination system

3 Josselin Gueno¹, Michael Borg², Simon Bourdareau¹, Guillaume Cossard¹, Olivier Godfroy¹,
4 Agnieszka Lipinska^{1,2}, Leila Tirichine³, J. Mark Cock^{1*}, Susana M. Coelho^{1,2*}

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

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

10 Abstract

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

27 Introduction

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

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

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

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

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

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

1
2
3 98 pseudoautosomal regions (PAR) (the SDR occupies 1/10th of the sex chromosome), such that
4 99 the large majority of the sex chromosome recombines within the PAR (42, 46, 47). Note that
5
6 100 the female and male SDRs on the U and V chromosomes respectively have about the same
7
8 101 physical size and share similar number of albeit distinct genes (42). Given the lack of
9
10 102 chromosome-scale dosage compensation, the small size of the SDR, and the fact that the SDRs
11
12 103 displays only mild levels of degeneration, the U and V sex chromosomes in this system are not
13
14 104 expected to present a markedly distinct chromatin landscape compared with autosomes.
15 105 However, this prediction has never been tested.

16 106 The expression pattern of genes on the U and V sex chromosome differs from that of
17 107 autosomal genes (39). For example, most *Ectocarpus* genes located on the U and V SDRs are
18 108 upregulated in the haploid gametophyte phase of the life cycle (42, 48). Moreover, when
19 109 compared with autosomes, the sex chromosome PARs harbour an excess of evolutionary
20 110 young or taxonomically restricted genes (47) and are enriched in both life cycle-related genes
21 111 (sporophyte-biased genes) and female-biased genes (49). Nevertheless, what chromatin
22 112 states associates with this intriguing composition of genes and patterns of gene expression in
23 113 a UV sexual system still remains unclear.

24 114 Here, we investigated sex-specific chromatin landscapes of autosomes and sex chromosomes
25 115 in *Ectocarpus*, a model brown alga with a UV sexual system. We built on our recent chromatin
26 116 profiling in *Ectocarpus* by studying six different histone PTMs – four that are associated with
27 117 gene activation, namely H3K4me3, H3K9ac, H3K27ac and H3K36me3, and two that are
28 118 associated with reduced gene expression, namely H4K20me3 and H3K79me2 (38). H4K20me3
29 119 was associated with repeated sequences and H3K79me2 with genomic regions that often
30 120 extended over several genes (38). Note that *Ectocarpus* DNA is not methylated (44) so we have
31 121 not analysed this modification here. Moreover, *Ectocarpus* lacks polycomb complexes and
32 122 associated PTMs, including the repression-associated mark H3K27me3 mark (38). Similarly,
33 123 H3K9me2/3 have been detected in *Ectocarpus* but at very low abundance (38). Consequently,
34 124 none of these additional repression-associated methylation marks were analysed in this study.

35 125 Comparison of the profiles of these six histone PTMs with transcriptomic data showed that
36 126 chromatin states were predictive of transcript abundance. The chromatin landscapes across
37 127 the genomes of males and females were similar, overall. However, the chromatin signatures
38 128 of genes that exhibited sex-biased expression was markedly different in males and females
39 129 indicating that histone modifications may play an important role in mediating sexual
40 130 differentiation. Moreover, a substantial proportion of the PAR genes presented sex-specific
41 131 chromatin patterns. The U and V sex chromosomes were found to have very distinct
42 132 chromatin landscapes to autosomes, despite the absence of a requirement for chromosome-

1
2
3 133 scale dosage compensation in *Ectocarpus* and the fact that the U and V chromosomes do not
4
5 134 exhibit strong signs of genetic degeneration.
6
7

8 135 **Material and Methods**

10 136 **Biological Material**

11
12
13 137 The near-isogenic male (Ec457) and female (Ec460) *Ectocarpus* lines (Table S1) were generated
14
15 138 by crossing brother and sister gametophytes for either five or six generations, respectively
16
17 139 (42). The resulting male and female strains, therefore, had essentially identical genetic
18
19 140 backgrounds apart from the non-recombining SDR (Table S2). To verify the homogeneity in
20
21 141 terms of genetic background, we used the male and female input DNA from the ChIP-seq
22
23 142 experiments aligned to the *Ectocarpus* reference genome to assess SNP diversity. The SNPs
24
25 143 for male and female samples were called with bcftools mpileup and filtered for minimal
26
27 144 mapping quality (--minQ 30), depth of coverage (--minDP 10) and missing data (--max-missing
28
29 145 0.9). We found 2,862,827 valid sites out of which 2,995 were variants (either SNPs or INDELS),
30
31 146 differing from the reference genome (the SDR regions were excluded). We next compared the
32
33 147 distribution of variant sites between males and females. Only 121 of the 2,995 variant sites
34
35 148 were segregated between sexes, which accounts for 0.004% of all sites. Given this very low
36
37 149 level of female/male polymorphism, it is highly likely that any differences we observed
38
39 150 between the two strains are due to the presence of the female and male SDRs. Furthermore,
40
41 151 note that the level of genetic diversity within the SDRs (which represent 869,870 bp and
42
43 152 893,800 bp in the female and male, respectively) has been shown to be extremely low (46),
44
45 153 as is the case for the non-recombining regions for animals and plants (50). Therefore, the
46
47 154 results presented here are likely to be representative of any male and female strain of
48
49 155 *Ectocarpus* species, although they are based on only a single male V chromosome and a single
50
51 156 female U chromosome.

52
53 157 Male and female gametophytes were cultured until near-maturity for 13 days as previously
54
55 158 described (51) at 13°C in autoclaved natural sea water supplemented with 300 µl/L Provasoli
56
57 159 solution, with a light:dark cycle of 12:12 h (20 µmol photons.m⁻².s⁻¹) using daylight-type
58
59 160 fluorescent tubes. Note that we used between 400-600 male and female haploid individual
60
61 161 gametophytes in each replicate, although there was only one genotype per each sex. Ten
62
63 162 individual gametophytes were grown in each petri dish. The level of maturity of male and
64
65 163 female individual gametophytes was assessed under the microscope to ensure synchrony in
66
67 164 terms of developmental stage. All manipulations were performed in a laminar flow hood
68
69 165 under sterile conditions.

70 166 **Comparisons of male and female transcriptomes using RNA-seq**

1
2
3 167 RNA for transcriptome analysis was extracted from the same duplicate male and female
4 168 cultures as were used for the ChIP-seq analysis (see above). For each sex, total RNA was
5 169 extracted from a mix of 90 gametophytes each, using the Qiagen Mini kit
6 170 (<http://www.qiagen.com>). RNA quality and quantity were assessed using an Agilent 2100
7 171 bioanalyzer, associated with Qubit2.0 Fluorometer using the Qubit RNA BR assay kit
8 172 (Invitrogen, Life Technologies, Carlsbad, CA, USA), as described previously (48, 49).

9 173 For each replicate sample, cDNA was synthesized using an oligo-dT primer. The cDNA was
10 174 fragmented, cloned, and sequenced by Fasteris (CH-1228 Plan-les-Ouates, Switzerland) using
11 175 an Illumina HiSeq 4000 set to generate 150-bp single-end reads. See Table S1 for RNA-seq
12 176 accession numbers.

13 177 Data quality was assessed using FastQC
14 178 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>; accessed May 2019). Reads
15 179 were trimmed and filtered using Cutadapt (52) with a quality threshold of 33 (quality-cutoff)
16 180 and a minimal size of 30 bp.

17 181 Filtered reads were mapped to version v2 of the *Ectocarpus* sp. 7 reference genome (45, 53)
18 182 using TopHat2 with the Bowtie2 aligner (54). More than 85% of the sequencing reads from
19 183 each library could be mapped to the reference genome (Table S1). Note that the reference
20 184 genome is from a male strain but the female SDR scaffolds have been added. Consequently,
21 185 male and female data were mapped to the same reference genome.

22 186 The mapped sequencing data were then processed with featureCounts (55) to obtain counts
23 187 for sequencing reads mapped to genes. Gene expression levels were represented as
24 188 transcripts per million (TPMs). Genes with expression values below the fifth percentile of all
25 189 TPM values calculated per sample were considered not to be expressed and were removed
26 190 from the analysis. This resulted in a total of 18,462 genes that were considered to be
27 191 expressed.

28 192 Differential expression analysis was performed with the DESeq2 package (Bioconductor) (56).
29 193 Genes were considered to be male-biased or female-biased if they exhibited at least a twofold
30 194 difference (fold change; FC) in expression between sexes with a false discovery rate (FDR)
31 195 < 0.05. A list of the sex-biased genes can be found in Table S5.

32 196 To calculate breadth of expression we employed the tissue-specificity index tau (57) using
33 197 published expression data from nine tissues or stages of the life cycle (female and male
34 198 immature and mature gametophytes, mixed male and female gametophytes, partheno-
35 199 sporophytes, upright partheno-sporophyte filaments, basal partheno-sporophyte filaments,
36 200 diploid sporophytes) from *Ectocarpus* (45, 47–49, 58). This allowed us to define broadly
37 201 expressed (housekeeping) genes (with $\tau < 0.25$) and narrowly expressed genes ($\tau > 0.75$).

202 **Genome-wide detection of histone PTMs**

203 Male versus female *Ectocarpus* sp. gametophyte ChIP-seq experiments were carried for
204 H3K4me3, H3K9ac, H3K27ac, H3K36me3, H4K20me3, and H3K79me2 and three controls (an
205 input control corresponding to sonicated DNA, histone H3 and immunoglobulin G monoclonal
206 rabbit (IgG)) as in (38). RNA-seq data (see above) was generated from the same samples, to
207 ensure that the histone PTM and gene expression data were fully compatible. For ChIP-seq,
208 2.8 g (corresponding to 2800 individual gametophytes) of *Ectocarpus* tissue was fixed for five
209 minutes in seawater containing 1% formaldehyde and the formaldehyde eliminated by rapid
210 filtering followed by incubation in PBS containing 400 mM glycine. Nuclei were isolated by
211 grinding in liquid nitrogen and in a Tenbroeck Potter in nuclei isolation buffer (0.1% triton X-
212 100, 125 mM sorbitol, 20 mM potassium citrate, 30 mM MgCl₂, 5 mM EDTA, 5 mM β-
213 mercaptoethanol, 55 mM HEPES at pH 7.5 with complete ULTRA protease inhibitors), filtering
214 through Miracloth and then washing the precipitated nuclei in nuclei isolation buffer with and
215 then without triton X-100. Chromatin was fragmented by sonicating the purified nuclei in
216 nuclei lysis buffer (10 mM EDTA, 1% SDS, 50 mM Tris-HCl at pH 8 with cComplete ULTRA
217 protease inhibitors) in a Covaris M220 Focused-ultrasonicator (duty 25%, peak power 75,
218 cycles/burst 200, duration 900 seconds at 6°C). The chromatin was incubated with an anti-
219 histone PTM antibody (anti-H4K20me3, reference 5737S, anti-H3K4me3, reference 9751S and
220 anti-H3K9ac, reference 9649S, Cell Signal Technology; anti-H3K27ac, reference 07360,
221 Millipore; anti-H3K36me3, reference 9050, Abcam; anti-H3K79me2, reference D15E8, Cell
222 Signal Technology) overnight at 4°C and the immunoprecipitation carried out using Dynabeads
223 protein A and Dynabeads protein G. Following immunoprecipitation and washing, a reverse
224 cross-linking step was carried out by incubating for at least six hours at 65°C in 200 mM NaCl
225 and the samples were then digested with Proteinase K and RNAse A. Purified DNA was
226 analysed on an Illumina HiSeq 4000 platform with a single-end sequencing primer over 50
227 cycles and pair-end sequencing for H3K79me2. At least 20 million reads were generated for
228 each immunoprecipitation. The ChIP-seq dataset has been deposited in the NCBI Gene
229 Expression Omnibus database under the accession numbers described in Table S2.

230 Quality control of the sequence data was carried out using FastQC
231 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Poor quality sequences were
232 removed and the high quality sequences trimmed with Cutadapt (52, 59). Illumina reads were
233 mapped onto the *Ectocarpus* v2 genome (45) using Bowtie (60), which contains both male and
234 female SDR. Duplicates were removed using samtools markdup in the Samtools package (v
235 1.9) (61).

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

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

241 To identify peaks and regions of chromatin mark enrichment in a gene-by-gene basis, each
242 data set, after combining data for biological replicates, was analysed separately for the male
243 and female gametophyte. Peaks corresponding to regions enriched in H3K4me3, H3K9ac and
244 H3K27ac were identified using the MACS2 (version 2.1.1) callpeak module (minimum FDR of
245 0.01) and refined with the MACS2 bdgpeakcall and bdgbroadcall modules (64). H3K36me3 and
246 H4K20me3 were analysed using SICER (v1.1) (minimum FDR of 0.01) (65, 66) with a window
247 size of 200 bp and a gap size of 400 bp. Note that peaks associated with sex-biased, PAR and
248 SDR genes were manually inspected to validate reproducibility between replicates. The signal
249 was normalized using the Signal Extraction Scaling (SES) method (67).

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

252 **Detection of chromatin states and signatures**

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

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

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

1
2
3 272 mainly permissive TSS marks, 'Permissive 2' for states with mainly activation-associated mark
4 273 H3K36me3, 'Silent' for states with enriched in H4K20me3 and/or H3K79me2 marks, 'Mixed'
5 274 for states with a combination of permissive and repressive marks and 'Null' for absence of
6 275 marks (Figure 1A). We then associated these five major categories with each gene in the
7 276 *Ectocarpus* genome to assemble a series of unique combinations that we termed 'chromatin
8 277 signatures'. Genes were considered to be in the null signature (S16) only when associated with
9 278 no other emission state but E12. This resulted in a total of 16 distinct chromatin signatures
10 279 across the *Ectocarpus* genome (S1 to S16) (Figure 1C).

17 280 **Coverage for each histone PTM**

18
19 281 The coverage for each histone PTM per chromosome was calculated using bedtools coverage
20 282 where the coverage of each PTM was normalized by the size of the chromosome. The
21 283 pseudoautosomal regions (PAR) and the sex-specific, non-recombining regions (SDR) of the
22 284 sex chromosome were analysed separately, as in (12).

26 285 **Statistical analysis**

28 286 Statistical analysis was performed in R 3.6.3. Permutation tests were performed to study the
29 287 differences of proportions of chromatin states in PAR and SDR genes compared to autosomal
30 288 genes. We randomly subsampled 100,000 times a number of chromatin states equal to the
31 289 number of PAR genes, SDR genes or both, from autosomal genes in order to perform
32 290 proportion tests. We compared observed and simulated Pearson's Chi-square statistics to
33 291 assess whether the observed differences in chromatin state proportions between gene sets
34 292 (autosomal, SDR, PAR, SDR+PAR) were statistically due to chance. A significant p-value
35 293 indicates that the observed difference in proportion is not due to chance. In order to eliminate
36 294 any possible effect of transposable element (TEs) prevalence (which is different between PAR,
37 295 SDR and autosomal genes) we also performed these tests using a randomized set of autosomal
38 296 genes that displayed exact the same TE prevalence. Similarly, we performed permutation tests
39 297 (100,000 permutations) to determine whether the distribution of chromatin states of
40 298 evolutionary young genes was significantly different from that of evolutionary conserved
41 299 autosomal genes with similar expression levels (within 25% of the median), separately in
42 300 males and females. We then compared Pearson's Chi-square statistics between observed and
43 301 simulated datasets.

44 302 We performed linear models of $\log_2(\text{TPM}+1)$ as a function of chromatin signatures and the
45 303 interaction between chromatin state and genomic location (i.e., autosome or PAR). We report
46 304 significant p-values in bold when states significantly influence the level of expression (state S1

1
2
3 305 used as reference level). Interaction term is significant when the effect of chromatin state of
4 306 expression level is significantly different for an autosomal gene compared with a PAR gene.

6
7 307 We used the list of evolutionary young genes identified in (47, 58). In brief, evolutionary young
8 308 genes, i.e. genes that are taxonomically restricted to *Ectocarpus*, were defined as genes
9 309 present in the genome of only *Ectocarpus* and having no BLASTp match (10^{-4} e value cutoff)
10 310 with a range of other stramenopile genome-wide proteomes from public databases (indicating
11 311 that they are likely to have evolved since the split from the most recent common ancestor):
12 312 the brown algae *Cladosiphon okamuranus*, *Macrocystis pyrifera*, *Saccharina japonica*,
13 313 *Scytosiphon lomentaria*, the eustigmatophyte *Nannochloropsis gaditana*, the pelagophyte
14 314 *Aureococcus anophagefferene*, and the diatom *Thalassiosira pseudonana*.

20 315 **GO-term analysis**

21
22 316 Gene set enrichment analysis (GSEA) was carried out separately for each sex, grouping genes
23 317 with either activation-associated (S1-S5) or repression associated (S13-S16) signatures. We
24 318 used Fisher's exact Test implemented in the R package TopGO (71) to identify significantly
25 319 enriched terms in biological processes. Top 30 categories (p-value<0.01) were plotted using
26 320 ggplot2 package for R (Wickham, 2016) investigated enrichment in terms of molecular
27 321 function ontology and report significant GO-terms with p-value < 0.01.

32 322 **Results**

35 323 **Identification of chromatin states in males and females of *Ectocarpus***

36 324 Our previous work associated H3K4me3, H3K9ac and H3K27ac to the transcription start sites
37 325 (TSS) of active genes, whereas H3K36me3 was associated with gene bodies (38). H4K20me3
38 326 was associated with repeated sequences, particularly transposons, whereas H3K79me2 peaks
39 327 often covered several kilobases and included multiple genes. The presence of H3K79me2 and
40 328 H4K20me3 on gene bodies correlated with decreased transcript abundance (38)). Thus,
41 329 H3K4me3, H3K9ac, H3K27ac and H3K36me3 may be considered activation-associated marks
42 330 and H3K79me2 and H4K20me3 repression-associated marks in *Ectocarpus*. Together, these
43 331 six histone PTMs are therefore expected to provide a broad overview of the chromatin
44 332 landscape in male and female *Ectocarpus*.

45 333 Near-isogenic male and female gametophyte (haploid) lines (Table S1, Figure S1) were used
46 334 to generate sex-specific ChIP-seq profiles for the six histone PTMs (Table S2-3). The male and
47 335 female haploid lines were generated by inbreeding over five and six generations respectively,
48 336 and were virtually identical genetically except for the sex-specific region (SDR) of the sex
49 337 chromosome (see methods, Table S2). We profiled at least 400 individual, clonal
50
51
52
53
54
55
56
57
58
59
60

1
2
3 338 gametophytes for each male and female replicate line and confirmed high reproducibility
4 339 between our ChIP-seq replicates (Figure S2).

6
7 340 We used the ChromHMM algorithm to define twelve representative chromatin states
8 341 common to males and females based on distinct combinatorial patterns of the six histone
9 342 PTMs (Figure 1A, Figure S3). Emission states E1-E3 consisted of combinations of the TSS-
10 343 enriched permissive marks H3K4me3, H3K9ac and H3K27ac (designated as group 'permissive
11 344 1') while emission states E4-S5 (group 'permissive 2') corresponded to regions enriched in
12 345 HK36me3 (Figure 1A). Emission states E6-E8 corresponded to mixed states that all included
13 346 H4K20me3 or H3K79me2 together with one or more of the activation-associated marks (group
14 347 'mixed'). Finally, silent emission states (E9-E11) were all enriched with H3K79me2 and/or
15 348 H4K20me3 and H3K36me3 (group 'silent') while emission state E12 corresponded to a 'null'
16 349 state that was devoid of the histone PTMs we assayed (group 'Null'). An example of the
17 350 histone PTM coverage over a 602 kbp region of the *Ectocarpus* genome is shown in Figure 1B.

18 351 While ChromHMM provided a broad overview of chromatin states across the *Ectocarpus*
19 352 genome, our aim was to focus on chromatin changes between females and males at the gene
20 353 level. Because more than one emission state could be present over the length of a gene in a
21 354 multitude of combinations, we determined which of the five above groups of emission states
22 355 were represented at each gene in the *Ectocarpus* genome and then defined a total of 16
23 356 distinct combination of these groups that we herein refer to as chromatin signatures (see
24 357 methods) (Figure 1C). Based on the predominant histone PTMs represented in each signature,
25 358 the signatures were then classed into three groups: activation-associated, mixed and
26 359 repression-associated.

40 360 **Chromatin signatures of different categories of *Ectocarpus* genes**

41
42 361 To elucidate the relationship between chromatin signatures and gene expression in
43 362 *Ectocarpus*, we generated paired RNA-seq data using the same biological samples as those
44 363 used for the ChIP-seq analysis (see methods). Together with previously published datasets (48,
45 364 49, 72), we defined four categories of genes based on their expression patterns: transcribed
46 365 genes ($\text{TPM} \geq 1$), silent genes ($\text{TPM} < 1$), housekeeping genes with broad expression patterns in
47 366 multiple tissues and life cycle stages ($\tau < 0.25$; see methods) and narrowly expressed genes
48 367 (NEGs; $\tau > 0.75$; see methods).

49
50 368 The most common chromatin signature for the transcribed genes (38.4% and 38.2% in females
51 369 and males respectively) was S3, which corresponds to co-localisation of three or four of the
52 370 activation-associated histone PTMs (H3K36me3, H3K27ac, H3K9ac, H3K4me3; Figure 1D,
53 371 Table S4). At 'silent' genes, signature S15 containing H3K79me2 and H4K20me3 was the most
54 372 common (21.7% and 21.9% in females and males, respectively; Figure 1D, Table S4).

1
2
3 373 Consistent with their ubiquitous expression, the majority of the housekeeping genes in
4 374 *Ectocarpus* (49% and 48.8% in males and females, respectively) were associated with
5 375 activation-associated signature S3. In contrast, NEGs were associated with a much larger
6 376 proportion of mixed and repression-associated signatures compared with housekeeping
7 377 genes, consistent with the restricted expression of NEGs (Figure 1D, Table S4). Finally, the
8 378 relative proportion of chromatin signatures was broadly similar between males and females
9 379 for each of the four gene categories (transcribed, silent, housekeeping and NEG) (Figure 1D,
10 380 Table S4). Together, these data support our categorisation of activation-associated or
11 381 repression-associated chromatin at *Ectocarpus* genes across the genome and suggests that
12 382 the chromatin landscape remains largely stable during sexual differentiation.

20 383 **Identification of histone PTMs associated with gene activation and gene repression**

21
22 384 To further investigate the relationship between the observed chromatin signatures and gene
23 385 expression, we assessed the transcript abundances of genes corresponding to each chromatin
24 386 signature. Consistently, genes that were assigned to an activation-associated signature had
25 387 higher transcript levels than genes with mixed signatures, whereas genes assigned to a
26 388 repression-associated signature exhibited the lowest levels of expression overall (Figure 2A).
27 389 A clear trend towards increasingly higher transcript abundance was correlated with the
28 390 gradual acquisition of more activation-associated marks (H3K9ac, H3K27ac, H3K4me3 and
29 391 H3K36me3; Figure 2A; Table S5). These observations support the proposed association of
30 392 these four histone PTMs with gene activation (38) and further validate our assignment of
31 393 chromatin signatures.

32
33 394 Conversely, genes corresponding to chromatin signatures with H4K20me3 and/or H3K79me2
34 395 consistently exhibited lower transcript levels than genes with equivalent chromatin states
35 396 without H4K20me3 and H3K79me2 (Figure 2A; Table S5). For example, transcript abundance
36 397 for genes with signature S13 was significantly lower than genes with signature S3 (Wilcox test,
37 398 p -value $< 2.22 \times 10^{-16}$ Figure 2A; Table S5). These results are consistent with H4K20me3 and
38 399 H3K79me2 being associated with repressed gene expression in *Ectocarpus*. Note however that
39 400 because H4K20me3 is frequently associated with transposons (38), the observed association
40 401 with reduced gene expression could also be indirect through the silencing of intronic
41 402 transposon sequences. Finally, *Ectocarpus* genes associated with null signature S16, which
42 403 corresponded to regions devoid of any of the assayed histone PTMs, exhibited very low
43 404 transcript abundance (Figure 2A, Table S5).

44
45 405 Next, we compared the expression level of genes with activation-associated signatures (S1-
46 406 S5) in males and females to genes with repression-associated signatures (S13-S16). Note that
47 407 we did not include genes with mixed signatures in this analysis because they exhibited a

1
2
3 408 combination of both activation-associated and repression-associated marks and because they
4 409 were expressed at intermediate levels (Figure 2A). As expected, genes marked with active
5 410 signatures were expressed at higher levels in both sexes than those that were associated with
6 411 repressive-associated signatures (Figure 2B; pair-wise Wilcox test, $p\text{-value} < 2.2E-16$).
7 412 Importantly, levels of gene expression in males and females were also significantly different
8 413 for genes marked with active signatures in one sex but with repression-associated signatures
9 414 in the other (Figure 2B; pair-wise Wilcox test, $p\text{-value} = 0.004$ and $p\text{-value} = 0.02$). Thus, despite
10 415 the lack of global changes in chromatin landscape between males and females, our profiles
11 416 illustrate localised chromatin signature changes associated with sex-specific gene expression.
12
13 417 Gene ontology (GO) term enrichment analysis showed that genes with activation-associated
14 418 (S1-S5) signatures were enriched in functions related to metabolic process, whereas genes
15 419 with repressive signatures (S13-S16) were enriched in functions related to signalling (Figure
16 420 2C). Interestingly, GO term enrichment appeared more stable between sexes with repressive-
17 421 associated chromatin signatures, whereas sex-specific GO term enrichment was more
18 422 apparent for genes with active signatures (Figure 2C). This difference was not due to
19 423 repression-associated signatures being more stably maintained at genes in males and females
20 424 compared to activation-associated signatures because similar proportions of genes exhibited
21 425 stable maintenance of either activation-associated or repression-associated signatures in
22 426 females versus males (83.1% and 81.6%; respectively; Figure 2D). Therefore, it appears that
23 427 genes with activation-associated chromatin signatures exhibit more sex-specific functions
24 428 than those with repressive signatures.

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

40 430 To investigate the role of histone PTMs in sexual differentiation, we examined the chromatin
41 431 signatures of genes with sex-biased expression. A comparison of gene expression patterns in
42 432 the two near-isogenic male and female lines (Figure S1), based on RNA-seq data generated
43 433 using the same biological samples as were used for the ChIP-seq analysis, identified a total of
44 434 268 genes that exhibited sex-biased expression (adjusted $p\text{-value} < 0.05$, fold change > 2 , TPM
45 435 > 1 ; Table S5, S6).

46 436 The presence of the active signatures was associated with higher transcript abundance for
47 437 sex-biased genes in both males and females (Figure S4, S5). Sex-biased genes therefore display
48 438 a similar association between activation-associated chromatin and increased gene expression
49 439 levels as that observed genome-wide (Figure 2A). Interestingly, 38.2% of male-biased genes
50 440 (MBGs) and 37.5% of female-biased genes (FBGs) had a different chromatin signature
51 441 between males and females (Table S5), suggesting that chromatin dynamics underlie sex-
52 442 biased gene expression in males and females.

1
2
3 443 Sex-biased genes tend to have narrow expression patterns (49, 73) so we compared their
4 444 chromatin patterns with that of NEGs. Overall, the proportions of the different chromatin
5 445 signatures were significantly different compared with NEGs suggesting that their chromatin
6 446 landscape is not related to their narrower pattern of expression (Chi-square test, p-value =
7 447 4.937E-15 and p-value = 0.01608 in FBGs vs NEGs in females and males, respectively, and p-
8 448 value = 5.627E-4 and p-value = 3.333E-6 for MBGs vs NEGs in females and males, respectively;
9 449 Figure 1D and Figure 3A).

10
11 450 For the MBGs, there was a difference between the relative proportions of the different
12 451 chromatin signatures in males compared to females. In males, repression-associated
13 452 chromatin signatures were less frequent than in females, whereas signatures that included
14 453 activation-associated marks (H3K9ac, H3K27ac, H3K4me3 and/or H3K36me3) were more
15 454 common (Chi-square test p-value=0.05; Figure 3A-B; Table S4). A subset of MBG (14.5%)
16 455 changed from a repression-associated or mixed signature in females to an active signature in
17 456 males (Figure 3C; Table S7). A similar situation, albeit less clear, was observed for FBG, where
18 457 active and mixed signatures were more frequent in females (47.4%) compared with males
19 458 (38.9%), although not significantly (Chi-square test p=0.093). Conversely, a larger proportion
20 459 (61%) of female-biased genes were in a repressive configuration in males compared with
21 460 females (52.6%; Figure 3A-B; Table S4) but, again, not significantly (Chi-square test p=0.097).
22 461 Like MBG, 12.5% of the FBG had repression-associated or mixed signatures in males whilst the
23 462 same genes become associated with activation-associated in females (Figure 3C, Table S7).

24 463 Figure 3D-E shows genome browser tracks at representative FBG and MBG genes illustrating
25 464 histone PTM changes during sexual differentiation. The MBG Ec12_002810, which encodes a
26 465 conserved protein of unknown function, had an activation-associated chromatin signature in
27 466 males (Table S5), but accumulated both H3K79me2 and H3K20me3 in females where it also
28 467 exhibited decreased expression (Figure 3D). Conversely, the reduced male expression of the
29 468 FBG Ec-05_003380, which encodes a peroxidase enzyme, was associated with the
30 469 augmentation of an H3K79me2 domain downstream of the gene (Figure 3E). This observation
31 470 suggests that H3K79me2 might undergo differential deposition between males and females.
32 471 Indeed, we noted that sex-specific domains of H3K79me2 were present at 12.1% (632) and
33 472 9.1% (457) of genes in males and females, respectively (Table S8). However, the majority of
34 473 these loci (97.3% and 97.2%) were not differentially transcribed between males and females,
35 474 suggesting that H3K79me2 dynamics might have an indirect impact on sex-biased gene
36 475 expression.

37 476 We also noticed more frequent sex-specific deposition of H3K36me3, H3K79me2 and
38 477 H3K20me3 (Table S8) compared with H3K27ac H3K4me3 and H3K9ac, suggesting that the
39 478 former may drive the changes in chromatin signatures observed in males versus females.

1
2
3 479 In conclusion, our analysis revealed chromatin signatures modifications that were
4 480 concomitant with changes in sex-biased gene expression between males and females, with
5 481 MBGs undergoing more histone PTM transitions during sexual differentiation compared with
6 482 FBGs.
7
8
9

10 483 **The chromatin landscape of the *Ectocarpus* sex chromosomes**

11
12 484 In organisms with diploid sexual systems (XY or ZW), sex chromosomes exhibit different
13 485 patterns of histone PTMs to autosomes (12, 26, 31, 74, 75). Given the nature of the *Ectocarpus*
14 486 UV system, where most of the U and V sex chromosome recombines at the PAR, a markedly
15 487 different chromatin landscape in sex chromosomes compared with autosomes is not
16 488 expected. We investigated the chromatin signature of genes in the PAR, SDR and autosomes
17 489 of *Ectocarpus* to test this hypothesis.
18
19
20
21

22 490 Surprisingly, a marked difference between sex chromosomes and autosomes was observed in
23 491 *Ectocarpus* (Figure 4A, Table S5, S9-12; Figure S7, S8, S9). While the relative proportion of the
24 492 16 chromatin signatures showed some variance between autosomes, genes on the sex
25 493 chromosomes exhibited a strikingly different pattern compared to autosomal genes (Figure
26 494 4A). In particular, there was a significant underrepresentation of genes with activation-
27 495 associated chromatin signatures (specifically S3 and S4) on the sex chromosomes compared
28 496 to the autosomes (Table S12). Furthermore, the sex chromosome was significantly enriched
29 497 in signatures that included the histone PTMs H4K20me3 and H3K79me2 compared with
30 498 autosomes (Figure 4A-C, Table S12).
31
32
33
34
35
36

37 499 The significantly distinct chromatin patterns between the sex chromosome and the
38 500 autosomes were also observed when only the PAR was taken into account (Chi-square test p-
39 501 value <2.2E-16; Figure 4A-B). For example, 52.1% and 46.7% of the PAR genes in females and
40 502 males, respectively, were associated with repressive signatures compared with 25.3% and
41 503 23.8% of autosomal genes for females and males, respectively (Table S5, Table S10).
42 504 Interestingly, 32% of the genes located in the PAR were found to be associated with different
43 505 chromatin signatures in males and females (Table S5), indicating that a substantial proportion
44 506 of the PAR genes display sex-dependent chromatin signature transitions. Note that only 11 of
45 507 the 430 PAR genes were classed as sex-biased genes (Table S5), so these sex-related chromatin
46 508 patterns on the PAR do not appear to be correlated with sex-biased PAR gene expression.
47
48
49
50
51
52

53 509 Analysis of the sex-determining regions of the U and V chromosomes showed that the vast
54 510 majority (85%) of the genes within the female SDR (i.e., U-specific genes) were associated with
55 511 repression-associated signatures, whereas this proportion was significantly less prevalent for
56 512 the male SDR (where 53% of genes were associated with repressive signatures) (Chi-square
57
58
59
60

1
2
3 513 test p-value = 0.02; Figure 4B; Table S11). Therefore, male and female SDRs have distinct
4 514 chromatin landscapes.

7 515 **Sex chromosome features and chromatin patterns**

8
9 516 Previous work has shown that the *Ectocarpus* PAR is enriched in transposons compared with
10 517 autosomes (42, 47). Considering that in *Ectocarpus* H4K20me3 co-localizes with transposon
11 518 sequences in *Ectocarpus* (38), we asked if the presence of transposons in PAR genes could
12 519 explain the observed chromatin state distribution patterns. More PAR genes contained a
13 520 transposon sequence compared to autosomal genes (80% versus 36%, respectively) but this
14 521 did not correlate with an increased proportion of PAR genes marked with H4K20me3 (28-29%
15 522 for the PAR versus 25-27% for autosomes) (Table S13). Moreover, permutation tests using
16 523 subsets of autosomal genes, in which 80% of the genes were selected to contain transposons
17 524 (i.e., a similar proportion of genes with transposons to that observed for the PAR) indicated
18 525 that the unusual pattern of chromatin signatures in the PAR was not due simply due to the
19 526 presence of additional genes with transposon insertions (Table S13).

20 527 Overall, transcript abundances for genes located in the PAR were significantly lower than for
21 528 genes located on autosomes (Wilcoxon p-value<2.22E-16; Table S4, Figure 4C). This difference
22 529 in expression level may potentially be explained by the different chromatin environment of
23 530 the PAR and the autosomes. To test this hypothesis, we selected a subset of autosomal genes
24 531 that had similar transcript levels to those of the PAR genes (Table S14). Interestingly, the
25 532 distribution of chromatin signatures for this set of autosomal genes was different to that of
26 533 the PAR genes with similar expression levels (Figure 4E, Figure S6) indicating that gene
27 534 expression level was not the cause of the difference in chromatin signature patterns between
28 535 the PAR genes and the autosomes. Moreover, the lower transcript abundance for PAR genes
29 536 was consistent with a higher proportion of genes in repressive configurations compared with
30 537 autosomal genes (52.9% and 46.7% for the PAR compared with 25.2% and 23.8% for the
31 538 autosomes, in females and males respectively; Table S10, Table S5). Note however that even
32 539 PAR genes with a permissive chromatin signature had significantly lower expression levels
33 540 compared to autosomal genes with similar signatures (pairwise Wilcoxon test, p-value=1.9E-
34 541 8, p-value=4.5E-10 for female and male respectively; Figure 4D). Thus, our results indicate that
35 542 transcription level is not the sole cause for the striking chromatin differences between PAR
36 543 genes and autosomal genes.

37 544 The PAR of *Ectocarpus* is enriched in evolutionarily young, *Ectocarpus*-restricted genes
38 545 ('young' genes, see methods) (47). Young genes have unusual structural characteristics
39 546 including shorter coding regions, fewer exons, lower expression levels and weaker codon bias
40 547 compared with older genes (47, 76, 77). We thus asked if the presence of evolutionary young
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 548 genes might explain the distinctive chromatin configuration in the PAR compared with
4 549 autosomes. Genome-wide analysis of the 4534 *Ectocarpus* young genes (58) revealed a
5 550 significantly different distribution of chromatin signatures compared with more conserved
6 551 genes with similar expression levels (Table S15). Almost half of the PAR genes were classed as
7 552 young genes (235 out of 440), which is a significant enrichment compared to autosomes (Chi-
8 553 square test p -value $<2.2E-16$). Moreover, the distribution of chromatin signatures at young
9 554 genes within the PAR was significantly different to that of conserved genes (Chi-square test,
10 555 p -value=1.57E-8 and p -value=4.26E-11 in females and males, respectively) (Figure 4E, S6;
11 556 Table S10). When only evolutionarily conserved genes are included in the analysis the
12 557 differences in chromatin distribution between the PAR and autosomes was considerably less
13 558 marked (Table S15). The enrichment of young genes on the PAR may therefore contribute to
14 559 its unique chromatin distribution.

15
16 560 Taken together, our observations suggest that the sex chromosome exhibits significantly
17 561 different features in terms of its chromatin landscape to the autosomes, not only at the level
18 562 of the non-recombining SDR region but also for the PAR. The distinct chromatin features of
19 563 the PAR are not explained by the preponderance of intragenic transposons nor by lower levels
20 564 of gene expression but rather by the increased incidence of evolutionarily young genes.

21 565 **Chromatin signatures and expression of sex chromosome genes**

22
23 566 Gene expression levels and deposition of chromatin marks were highly correlated for the
24 567 complete set of *Ectocarpus* genes (see above, Figure 2A). For example, genes with an S3
25 568 signature enriched for all four activation-associated marks had significantly higher expression
26 569 than genes with mixed signature S7 that was distinguished by the added presence of
27 570 H3K79me2 and H4K20me3. However, when we analysed the association between expression
28 571 level and chromatin state for genes located on the PAR of the sex chromosomes the situation
29 572 was different. For example, the two chromatin signatures S4 and S6 in females had a
30 573 significantly weaker correlation with expression on the PAR compared with autosomes. In
31 574 males, a weaker correlation with gene expression was also observed for the PAR compared to
32 575 the autosomes for signatures S4, S6, S7 and S12 (Table S16, Figure S10). In other words,
33 576 depending on the location (PAR or autosomes) the correlation between chromatin signature
34 577 and gene expression level was not the same.

35
36 578 With respect to SDR genes, the only significant correlation between expression level and
37 579 chromatin signature was observed for chromatin signatures S3 and S4 in males and S15 and
38 580 S16 in females (Table S17), although the small sample size of SDR genes decreases the power
39 581 of our statistical testing. Note that activation-associated mark H3K36me3 was more often
40 582 present at male SDR genes (17/30) than at female SDR genes (4/22), with H3K36me3 coverage

1
2
3 583 also higher on the male SDR than for the female SDR (Figure S7, S8, Table S9). We also noticed
4 584 that median transcript levels for male SDR genes were higher than that of female SDR genes,
5 585 although the difference was not significant (Figure 4G). In conclusion, our chromatin profiles
6 586 suggest a different relationship between chromatin signature and expression levels for genes
7 587 on the sex chromosomes compared with autosomes, further highlighting the unique
8 588 chromatin configuration of UV sex chromosomes.

14 589 **Discussion**

17 590 **Chromatin regulation in a haploid UV sexual system**

19 591 Three types of genetic sex determination system exist in nature: XX/XY, ZZ/ZW systems and
20 592 U/V systems (1, 32). For UV systems, studies have focused on understanding sex
21 593 determination and sex-biased gene expression (e.g. (34, 35, 78) but we know considerably less
22 594 about chromatin patterns in males compared to females. Our study provides the first overview
23 595 of sex-specific differences in chromatin landscape in a haploid UV system and its relationship
24 596 with sex-biased gene expression, whilst also revealing the chromatin configuration of the U
25 597 and V sex chromosome.

31 598 Analysis of six histone PTMs in *Ectocarpus* males and females resulted in the definition of 16
32 599 distinct chromatin signatures associated with genes. Chromatin signatures that included
33 600 different combinations of H3K4me3, H3K9ac, H3K27ac and H3K36me3 were associated with
34 601 transcriptionally-permissive genes, whereas chromatin signatures that included H3K79me2
35 602 and/or H4K20me3 were associated with decrease gene expression compared to equivalent
36 603 states lacking these marks. Signatures with H3K36me3 were associated with broadly
37 604 expressed genes but were less prevalent on genes with narrow or tissue-specific expression,
38 605 which could be related to a lower sensitivity in detecting H3K36me3 accumulation in a
39 606 restricted subset of cells. The difference between housekeeping and NEG genes was
40 607 considerably more marked for H3K36me3-containing signatures than for those with TSS-
41 608 located marks (Table S4, S5), perhaps indicating a stronger association of this mark with gene
42 609 transcription. A similar association of H3K36me3 with broadly expressed genes has been
43 610 described for *Drosophila* (12, 79), indicating that this correlation has been conserved across
44 611 distantly related lineages. Overall, the *Ectocarpus* chromatin patterns described here are
45 612 consistent with H3K4me3, H3K9ac, H3K27ac and H3K36me3 having similar roles in brown
46 613 algae, land plants and animals (38, 80–82). The role of H4K20me3, in contrast, appears to be
47 614 less conserved across eukaryotic supergroups, being associated with low transcriptional levels
48 615 in both animals and brown algae but with euchromatin and transcriptional activation in land
49 616 plants (83, 84).

1
2
3 617 Transcriptional reprogramming during life cycle transitions often involves the loss of
4 618 repressive chromatin marks (85). Our recent profiling of the *Ectocarpus* sporophyte and
5 619 gametophyte indicated that H3K79me2-enriched domains, which are associated with
6 620 repressed genes in *Ectocarpus*, are stably maintained between generations (38). Interestingly,
7 621 here we identified several examples of genes that were associated with H3K79me2 in one sex
8 622 but not in another. However, the differential loss of this mark between sexes was not
9 623 correlated systematically with changes in sex-biased gene expression, suggesting that
10 624 H3K79me2 reprogramming might have only an indirect impact on transcription and sexual
11 625 differentiation.

18 626 **Chromatin signatures of *Ectocarpus* sex-biased genes**

19 627 When considered genome-wide, the proportion of genes associated with each chromatin
20 628 signature did not differ substantially in males compared with females. However, when
21 629 individual genes were compared, a considerable fraction was associated with different
22 630 chromatin states in the two sexes, including genes that did not exhibit sex-biased expression
23 631 patterns. Despite a lack of transcriptional changes in many cases, the strong correlation
24 632 between chromatin signatures and gene expression argues that the chromatin reconfiguration
25 633 we describe is biologically significant. One hypothesis is that the sex-specific alterations in
26 634 chromatin manifest prior to any significant sex differences in transcription and phenotypic
27 635 differentiation. In other words, differences in chromatin state may forecast sex-biased
28 636 differences in gene expression during later stages of development, as reported for mammalian
29 637 foetal germ cells (86). A more refined study using several stages during male and female
30 638 gametophyte development would be needed to gain further insights into this matter.

31 639 In males, FBG were more often associated with repressive signatures than in females.
32 640 Similarly, more MBGs in females were marked with repression-associated signatures
33 641 compared with males, whereas FBGs were more often in permissive or mixed signatures.
34 642 About 37% of sex-biased genes had different chromatin signatures in males or females, which
35 643 is significantly more than chromatin changes occurring at unbiased genes. These observations
36 644 support a link, at least partial, between chromatin signature, expression pattern and role of
37 645 sex-biased genes during sexual differentiation in *Ectocarpus*. Sex-specific chromatin states
38 646 appear not to explain the sex-biased expression patterns in *Drosophila* (12), and mouse (87).
39 647 It appears therefore that there is no absolute correlation between SBG and chromatin
40 648 landscape in animals, and sex-biased gene expression may be regulated by other mechanisms,
41 649 such as distal regulatory sites (87) or involving accessibility and 3D structure of chromatin
42 650 (88). Future work focusing on these alternative mechanisms in *Ectocarpus* may help to further
43 651 understand the regulation of sex-biased gene expression in the brown algae.

652 **Unique chromatin organisation features of the U and V sex chromosome**

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

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

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

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

1
2
3 687 differently on autosomes compared with sex chromosomes. It is possible that the expression
4 688 of genes on the U and V sex chromosomes is regulated by different chromatin processes than
5 689 those that regulate autosomal gene expression, perhaps involving histone PTMs we did not
6 690 assay in this study. Note that the PAR features are unlikely to be caused by linkage
7 691 disequilibrium with the SDR because the PAR is considerably large compared with the SDR and
8 692 recombines extensively (47). Further investigation of the chromatin landscape of UV
9 693 chromosomes in a diploid stage, and in several developmental stages during the haploid-
10 694 diploid life cycle, together with further profiling in other species with a UV sexual system,
11 695 promises to reveal more extraordinary features of these prevalent sex chromosomes.

18 19 696 **Figure legends**

20
21 697 **Figure 1.** Histone PTMs and chromatin signatures of female and male *Ectocarpus*. A) A model
22 698 of prevalent chromatin emission states found in *Ectocarpus* using ChromHMM. Permissive 1
23 699 and Permissive 2, activation-associated states; Mixed, States that mix activation-associated
24 700 and repression-associated chromatin PTMs; Silent, repression-associated chromatin states;
25 701 Null, absence of assayed histone PTMs. B) Representative region of the chromosome 19
26 702 showing profiles of mapped ChIP-seq reads for the six histone PTMs in females and males.
27 703 Coverage is represented as the ratio of IP DNA relative to H3 for H3K36me3, H4K20me3 and
28 704 H3K79me2 and input for TSS marks (H3K4me3, H3K9ac, H3K27ac). C) Chromatin signatures
29 705 assigned to genes based on ChromHMM states (see methods). Percentages of the total gene
30 706 set associated with each chromatin signature in males (M) and females (F) are shown to the
31 707 right. D) Proportions of transcribed ($\text{TPM} \geq 1\text{TPM}$), silent ($\text{TPM} < 1\text{TPM}$), housekeeping
32 708 ($\tau < 0.25$) and narrowly expressed genes ($\tau > 0.75$) associated with each chromatin signature
33 709 in males and females.

34
35
36
37
38
39
40
41
42 710 **Figure 2.** Gene expression and chromatin states. A) Transcript abundances for genes
43 711 associated with different chromatin signatures in males and females. The colour code is the
44 712 same as that used in Figure 1A. B) Transcript abundances for genes exhibiting either activation-
45 713 associated (S1 to S5) or repression-associated (S13 or S16) chromatin signatures in females
46 714 (dark pink) and males (grey). C) GO term enrichment for genes marked with activation-
47 715 associated or repression-associated chromatin signatures in males and females. D) Venn
48 716 diagrams representing the proportion of genes marked with activation-associated or
49 717 repression-associated chromatin signatures in males and females.

50
51
52
53
54
55 718 **Figure 3.** Histone PTM patterns at sex-biased genes in *Ectocarpus* males and females. A)
56 719 Proportions of the 16 chromatin signatures for female-biased, male-biased and unbiased
57 720 genes in females (left) and males (right). The number of genes in each category are inside

1
2
3 721 brackets. Chromatin signature colour codes correspond to Figure 1C. B) Proportions of genes
4 722 associated with each of the 16 chromatin signatures for female-biased (FBG) and male-biased
5 723 (MBG) genes in females (left) and males (right). The intensity of the grey squares is
6 724 proportional to the number of genes corresponding to each signature. Coloured squares
7 725 represent the different chromatin signatures (see Figure 1A). C) Chord diagrams comparing
8 726 chromatin signatures associated with female-biased (left) and male-biased (right) genes in
9 727 females and males. The colour code for the chromatin states is the same as that used in Figure
10 728 1C. Each chord represents a sex-biased gene and illustrates whether a gene changes from one
11 729 signature to another in the opposing sex. D) Representative chromatin profiles for a male-
12 730 biased gene on chromosome 12 in females and males. E) Representative chromatin profiles
13 731 for a female-biased gene in chromosome 5 in females and males.

14 732 **Figure 4.** Chromatin landscape of the U and V sex chromosomes compared with the
15 733 autosomes. A) Chromatin signature distribution for each autosome and for the SDR and PAR
16 734 regions of the sex chromosome in females (left panel) and in males (right panel). B)
17 735 Proportions of genes associated with each of the 16 chromatin signatures for all autosomes
18 736 and for the PAR and SDR regions of the sex chromosome in females and in males. The intensity
19 737 of the grey is proportional to the number of genes in each signature. The colour code for the
20 738 chromatin states is the same as that used in Figure 1A. C) Transcript abundances, measured
21 739 as $\log_2(\text{TPM}+1)$, for autosomal and for PAR genes in males and females. Significant differences
22 740 were assessed using pairwise Wilcoxon rank sum test. D) Transcript abundances for autosomal
23 741 and PAR genes associated with different chromatin signatures: activation-associated
24 742 signatures (S1-S5) and repression-associated signatures (S13-S16). Significant differences
25 743 were assessed using pairwise Wilcoxon rank sum test. E) Chromatin state distribution of
26 744 evolutionary young genes compared with autosomal conserved genes with similar expression
27 745 patterns. See also Table S13. F) Transcript abundances, measured as $\log_2(\text{TPM}+1)$, for
28 746 individual genes located in the female and male sex determining regions (SDRs). Coloured
29 747 plots represent chromatin signatures corresponding to the colour code indicated in Figure 1C
30 748 (see also Table S10). G) Transcript abundances of genes located within the female and male
31 749 sex-specific regions (SDRs). Significant differences were assessed using pairwise Wilcoxon rank
32 750 sum test.

33 751

34 752 **Acknowledgements**

35 753 We thank Carl Herrmann and Swann Floc'hlay for advice about ChIP-seq analysis, Thomas
36 754 Broquet for discussions on the statistical analysis and Remy Luthringer for assistance in the

1
2
3 755 preparation of the Ectocarpus cultures. We thank the Institut Français de Bioinformatique and
4 756 the Roscoff Analysis and Bioinformatics for Marine Science platform ABiMS ([http://abims.sb-](http://abims.sb-roscoff.fr)
5 757 [roscoff.fr](http://abims.sb-roscoff.fr)) for providing computing and data storage resources. This work was supported by
6 758 the ERC (projects 638240 and 864038), the CNRS, Sorbonne Université, the Max Planck Society
7 759 and the Agence Nationale de la Recherche project Epicycle (ANR-19-CE20-0028-01).

12 760 **Contributions**

15 761 S.C. conceived of the project. J.G. performed the main experimental work and bioinformatic
16 762 analyses. S.B. and L.T. developed protocols. M.B generated H3K79me2 ChIP-seq profiles. J.G.
17 763 and M.B. defined chromatin signatures. A.L. performed GO-enrichment and RNA-seq analyses.
18 764 G.C. performed statistical analyses. J.G., M.B., J.M.C. and S.C. interpreted the data. M.B. re-
19 765 plotted and re-assembled the main figures. S.C. and J.M.C supervised J.G. S.C. wrote the
20 766 manuscript with input from M.B. All authors read and commented the manuscript.

26 767 **References**

- 28 768 1. Bachtrog,D., Mank,J.E., Peichel,C.L., Kirkpatrick,M., Otto,S.P., Ashman,T.-L., Hahn,M.W.,
29 769 Kitano,J., Mayrose,I., Ming,R., *et al.* (2014) Sex determination: why so many ways of
30 770 doing it? *PLoS Biol*, **12**, e1001899.
- 33 771 2. Charlesworth,D. (2017) Evolution of recombination rates between sex chromosomes. *Philos*
34 772 *Trans R Soc Lond B Biol Sci*, **372**.
- 36 773 3. Umen,J. and Coelho,S. (2019) Algal Sex Determination and the Evolution of Anisogamy.
37 774 *Annu. Rev. Microbiol.*, 10.1146/annurev-micro-020518-120011.
- 40 775 4. Grath,S. and Parsch,J. (2016) Sex-Biased Gene Expression. *Annu Rev Genet*, **50**, 29–44.
- 42 776 5. Jones,P.A. (2012) Functions of DNA methylation: islands, start sites, gene bodies and
43 777 beyond. *Nat Rev Genet*, **13**, 484–492.
- 46 778 6. Nugent,B.M., Wright,C.L., Shetty,A.C., Hodes,G.E., Lenz,K.M., Mahurkar,A., Russo,S.J.,
47 779 Devine,S.E. and McCarthy,M.M. (2015) Brain feminization requires active repression
48 780 of masculinization via DNA methylation. *Nat Neurosci*, **18**, 690–697.
- 50 781 7. Elango,N., Hunt,B.G., Goodisman,M.A.D. and Yi,S.V. (2009) DNA methylation is widespread
51 782 and associated with differential gene expression in castes of the honeybee, *Apis*
52 783 *mellifera*. *Proc Natl Acad Sci U S A*, **106**, 11206–11211.
- 55 784 8. Schmid,M.W., Giraldo-Fonseca,A., Rovekamp,M., Smetanin,D., Bowman,J.L. and
56 785 Grossniklaus,U. (2018) Extensive epigenetic reprogramming during the life cycle of
57 786 *Marchantia polymorpha*. *Genome Biol*, **19**, 9.
- 59 787 9. Kouzarides,T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.

- 1
2
3 788 10. Lindeman,L.C., Winata,C.L., Aanes,H., Mathavan,S., Alestrom,P. and Collas,P. (2010)
4 789 Chromatin states of developmentally-regulated genes revealed by DNA and histone
5 790 methylation patterns in zebrafish embryos. *Int J Dev Biol*, **54**, 803–813.
6
7
8 791 11. Srivastava,S., Mishra,R.K. and Dhawan,J. (2010) Regulation of cellular chromatin state:
9 792 insights from quiescence and differentiation. *Organogenesis*, **6**, 37–47.
10
11 793 12. Brown,E.J. and Bachtrog,D. (2014) The chromatin landscape of Drosophila: comparisons
12 794 between species, sexes, and chromosomes. *Genome Res*, **24**, 1125–1137.
13
14 795 13. Brown,E.J., Nguyen,A.H. and Bachtrog,D. (2020) The Drosophila Y Chromosome Affects
15 796 Heterochromatin Integrity Genome-Wide. *Mol Biol Evol*, **37**, 2808–2824.
16
17
18 797 14. Navratilova,P., Danks,G.B., Long,A., Butcher,S., Manak,J.R. and Thompson,E.M. (2017) Sex-
19 798 specific chromatin landscapes in an ultra-compact chordate genome. *Epigenetics*
20 799 *Chromatin*, **10**, 3.
21
22
23 800 15. Zhou,Q. and Bachtrog,D. (2015) Ancestral Chromatin Configuration Constrains Chromatin
24 801 Evolution on Differentiating Sex Chromosomes in Drosophila. *PLoS Genet*, **11**,
25 802 e1005331.
26
27
28 803 16. Zhou,Q., Ellison,C.E., Kaiser,V.B., Alekseyenko,A.A., Gorchakov,A.A. and Bachtrog,D.
29 804 (2013) The epigenome of evolving Drosophila neo-sex chromosomes: dosage
30 805 compensation and heterochromatin formation. *PLoS Biol*, **11**, e1001711.
31
32 806 17. Baker,B.S., Gorman,M. and Marin,I. (1994) Dosage compensation in Drosophila. *Annu Rev*
33 807 *Genet*, **28**, 491–521.
34
35
36 808 18. Gelbart,M.E. and Kuroda,M.I. (2009) Drosophila dosage compensation: a complex voyage
37 809 to the X chromosome. *Development*, **136**, 1399–1410.
38
39 810 19. Girton,J.R. and Johansen,K.M. (2008) Chromatin structure and the regulation of gene
40 811 expression: the lessons of PEV in Drosophila. *Adv Genet*, **61**, 1–43.
41
42 812 20. Lemos,B., Branco,A.T. and Hartl,D.L. (2010) Epigenetic effects of polymorphic Y
43 813 chromosomes modulate chromatin components, immune response, and sexual
44 814 conflict. *Proc Natl Acad Sci U S A*, **107**, 15826–15831.
45
46
47 815 21. Straub,T. and Becker,P.B. (2007) Dosage compensation: the beginning and end of
48 816 generalization. *Nat Rev Genet*, **8**, 47–57.
49
50
51 817 22. Vicoso,B. and Charlesworth,B. (2006) Evolution on the X chromosome: unusual patterns
52 818 and processes. *Nat Rev Genet*, **7**, 645–653.
53
54 819 23. Furman,B.L.S., Metzger,D.C.H., Darolti,I., Wright,A.E., Sandkam,B.A., Almeida,P., Shu,J.J.
55 820 and Mank,J.E. (2020) Sex Chromosome Evolution: So Many Exceptions to the Rules.
56 821 *Genome Biol Evol*, **12**, 750–763.
57
58
59 822 24. Bachtrog,D. (2013) Y chromosome evolution: emerging insights into processes of Y
60 823 chromosome degeneration. *Nature reviews. Genetics*, **14**, 113–124.

- 1
2
3 824 25. Charlesworth,B. and Charlesworth,D. (2000) The degeneration of Y chromosomes. *Philos*
4 825 *Trans R Soc Lond B Biol Sci*, **355**, 1563–1572.
- 6 826 26. Lucchesi,J.C., Kelly,W.G. and Panning,B. (2005) Chromatin remodeling in dosage
8 827 compensation. *Annu Rev Genet*, **39**, 615–651.
- 10 828 27. Picard,M.A.L., Cosseau,C., Ferre,S., Quack,T., Grevelding,C.G., Coute,Y. and Vicoso,B.
11 829 (2018) Evolution of gene dosage on the Z-chromosome of schistosome parasites. *Elife*,
12 830 **7**.
- 15 831 28. Vicoso,B. and Charlesworth,B. (2009) Effective population size and the faster-X effect: an
16 832 extended model. *Evolution*, **63**, 2413–2426.
- 18 833 29. Chandler,C.H. (2017) When and why does sex chromosome dosage compensation evolve?
19 834 *Ann N Y Acad Sci*, **1389**, 37–51.
- 22 835 30. Yasuhara,J.C. and Wakimoto,B.T. (2008) Molecular landscape of modified histones in
23 836 *Drosophila* heterochromatic genes and euchromatin-heterochromatin transition
24 837 zones. *PLoS Genet*, **4**, e16.
- 26 838 31. Picard,M.A.L., Vicoso,B., Roquis,D., Bulla,I., Augusto,R.C., Arancibia,N., Grunau,C.,
27 839 Boissier,J. and Cosseau,C. (2019) Dosage Compensation throughout the *Schistosoma*
28 840 *mansoni* Lifecycle: Specific Chromatin Landscape of the Z Chromosome. *Genome Biol*
30 841 *Evol*, **11**, 1909–1922.
- 32 842 32. Coelho,S.M., Gueno,J., Lipinska,A.P., Cock,J.M. and Umen,J.G. (2018) UV chromosomes
33 843 and haploid sexual systems. *Trends Plant Sci*, **23**, 794–807.
- 35 844 33. Coelho,S.M. and Cock,J.M. (2020) Brown Algal Model Organisms. *Annu. Rev. Genet.*, **54**,
37 845 71–92.
- 39 846 34. Carey,S.B., Jenkins,J., Lovell,J.T., Maumus,F., Sreedasyam,A., Payton,A.C., Shu,S.,
40 847 Tiley,G.P., Fernandez-Pozo,N., Healey,A., *et al.* (2021) Gene-rich UV sex chromosomes
41 848 harbor conserved regulators of sexual development. *Sci Adv*, **7**.
- 44 849 35. Iwasaki,M., Kajiwara,T., Yasui,Y., Yoshitake,Y., Miyazaki,M., Kawamura,S., Suetsugu,N.,
45 850 Nishihama,R., Yamaoka,S., Wanke,D., *et al.* (2021) Identification of the sex-
46 851 determining factor in the liverwort *Marchantia polymorpha* reveals unique evolution
47 852 of sex chromosomes in a haploid system. *Curr Biol*, 10.1016/j.cub.2021.10.023.
- 49 853 36. Montgomery,S.A., Tanizawa,Y., Galik,B., Wang,N., Ito,T., Mochizuki,T., Akimcheva,S.,
50 854 Bowman,J.L., Cognat,V., Maréchal-Drouard,L., *et al.* (2020) Chromatin Organization in
52 855 Early Land Plants Reveals an Ancestral Association between H3K27me3, Transposons,
53 856 and Constitutive Heterochromatin. *Curr Biol*, **30**, 573-588.e7.
- 55 857 37. Bowman,J.L., Kohchi,T., Yamato,K.T., Jenkins,J., Shu,S., Ishizaki,K., Yamaoka,S.,
56 858 Nishihama,R., Nakamura,Y., Berger,F., *et al.* (2017) Insights into Land Plant Evolution
58 859 Garnered from the *Marchantia polymorpha* Genome. *Cell*, **171**, 287-304.e15.
- 59
60

- 1
2
3 860 38. Bourdareau,S., Tirichine,L., Lombard,B., Loew,D., Scornet,D., Wu,Y., Coelho,S.M. and
4 861 Cock,J.M. (2021) Histone modifications during the life cycle of the brown alga
5 862 *Ectocarpus*. *Genome Biology*, **22**, 12.
- 7
8 863 39. Coelho,S.M., Mignerot,L. and Cock,J.M. (2019) Origin and evolution of sex-determination
9 864 systems in the brown algae. *New Phytologist*, **10.1111/nph.15694**.
- 11 865 40. Bull,J.J. (1978) Sex Chromosomes in Haploid Dioecy: A Unique Contrast to Muller's Theory
12 866 for Diploid Dioecy. *The American Naturalist*, **112**, 245–250.
- 14
15 867 41. Immler,S. and Otto,S.P. (2015) The evolution of sex chromosomes in organisms with
16 868 separate haploid sexes. *Evolution*, **69**, 694–708.
- 18 869 42. Ahmed,S., Cock,J.M., Pessia,E., Luthringer,R., Cormier,A., Robuchon,M., Sterck,L.,
19 870 Peters,A.F., Dittami,S.M., Corre,E., *et al.* (2014) A haploid system of sex determination
20 871 in the brown alga *Ectocarpus* sp. *Curr. Biol.*, **24**, 1945–1957.
- 23 872 43. Baudry,L., Guiglielmoni,N., Marie-Nelly,H., Cormier,A., Marbouty,M., Avia,K., Mie,Y.L.,
24 873 Godfroy,O., Sterck,L., Cock,J.M., *et al.* (2020) instaGRAAL: chromosome-level quality
25 874 scaffolding of genomes using a proximity ligation-based scaffold. *Genome Biol*, **21**,
26 875 148.
- 28
29 876 44. Cock,J.M., Sterck,L., Rouzé,P., Scornet,D., Allen,A.E., Amoutzias,G., Anthouard,V.,
30 877 Artiguenave,F., Aury,J.-M., Badger,J.H., *et al.* (2010) The *Ectocarpus* genome and the
31 878 independent evolution of multicellularity in brown algae. *Nature*, **465**, 617–621.
- 33 879 45. Cormier,A., Avia,K., Sterck,L., Derrien,T., Wucher,V., Andres,G., Monsoor,M., Godfroy,O.,
34 880 Lipinska,A., Perrineau,M.-M., *et al.* (2017) Re-annotation, improved large-scale
35 881 assembly and establishment of a catalogue of noncoding loci for the genome of the
36 882 model brown alga *Ectocarpus*. *New Phytol*, **214**, 219–232.
- 39 883 46. Avia,K., Lipinska,A.P., Mignerot,L., Montecinos,A.E., Jamy,M., Ahmed,S., Valero,M.,
40 884 Peters,A.F., Cock,J.M., Roze,D., *et al.* (2018) Genetic diversity in the UV sex
41 885 chromosomes of the brown alga *Ectocarpus*. *Genes (Basel)*, **9**.
- 44 886 47. Luthringer,R., Lipinska,A.P., Roze,D., Cormier,A., Macaisne,N., Peters,A.F., Cock,J.M. and
45 887 Coelho,S.M. (2015) The pseudoautosomal regions of the U/V sex chromosomes of the
46 888 brown alga *Ectocarpus* exhibit unusual features. *Mol Biol Evol*, **32**, 2973–2985.
- 48
49 889 48. Lipinska,A.P., Toda,N.R.T., Heesch,S., Peters,A.F., Cock,J.M. and Coelho,S.M. (2017)
50 890 Multiple gene movements into and out of haploid sex chromosomes. *Genome Biol*, **18**,
51 891 104.
- 53 892 49. Lipinska, Cormier,A., Luthringer,R., Peters,A.F., Corre,E., Gachon,C.M.M., Cock,J.M. and
54 893 Coelho,S.M. (2015) Sexual dimorphism and the evolution of sex-biased gene
55 894 expression in the brown alga *Ectocarpus*. *Mol Biol Evol*, **32**, 1581–1597.
- 57
58 895 50. Wilson Sayres,M.A. (2018) Genetic Diversity on the Sex Chromosomes. *Genome Biology*
59 896 *and Evolution*, **10**, 1064–1078.
- 60

- 1
2
3 897 51. Coelho,S.M., Scornet,D., Rousvoal,S., Peters,N.T., Dartevelle,L., Peters,A.F. and Cock,J.M.
4 898 (2012) How to cultivate *Ectocarpus*. *Cold Spring Harb Protoc*, **2012**, 258–261.
- 5
6 899 52. Martin,M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing
7 900 reads. *EMBNet*, <https://doi.org/10.14806/ej.17.1.200>.
- 8
9 901 53. Montecinos,A., Valero,M., Guillemin,M.-L., Peters,A., Desrut,A. and Couceiro,L. (2017)
10 902 Species delimitation and phylogeographic analyses in the *Ectocarpus* subgroup
11 903 *siliculosi* (Ectocarpales, Phaeophyceae). *Journal of Phycology*, **53**, 17–31.
- 12
13 904 54. Kim,D., Pertea,G., Trapnell,C., Pimentel,H., Kelley,R. and Salzberg,S.L. (2013) TopHat2:
14 905 accurate alignment of transcriptomes in the presence of insertions, deletions and gene
15 906 fusions. *Genome Biol.*, **14**, R36.
- 16
17 907 55. Liao,Y., Smyth,G.K. and Shi,W. (2014) featureCounts: an efficient general purpose program
18 908 for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930.
- 19
20 909 56. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and
21 910 dispersion for RNA-seq data with DESeq2. *Genome Biol*, **15**, 550.
- 22
23 911 57. Yanai,I., Benjamin,H., Shmoish,M., Chalifa-Caspi,V., Shklar,M., Ophir,R., Bar-Even,A., Horn-
24 912 Saban,S., Safran,M., Domany,E., *et al.* (2005) Genome-wide midrange transcription
25 913 profiles reveal expression level relationships in human tissue specification.
26 914 *Bioinformatics*, **21**, 650–659.
- 27
28 915 58. Lipinska,A.P., Serrano-Serrano,M.L., Cormier,A., Peters,A.F., Kogame,K., Cock,J.M. and
29 916 Coelho,S.M. (2019) Rapid turnover of life-cycle-related genes in the brown algae.
30 917 *Genome Biol*, **20**, 35.
- 31
32 918 59. Hansen,P., Hecht,J., Ibn-Salem,J., Menkuec,B.S., Roskosch,S., Truss,M. and Robinson,P.N.
33 919 (2016) Q-nexus: a comprehensive and efficient analysis pipeline designed for ChIP-
34 920 nexus. *BMC Genomics*, **17**, 873.
- 35
36 921 60. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient
37 922 alignment of short DNA sequences to the human genome. *Genome Biol*, **10**, R25.
- 38
39 923 61. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G.,
40 924 Durbin,R., and 1000 Genome Project Data Processing Subgroup (2009) The Sequence
41 925 Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
- 42
43 926 62. Landt,S.G., Marinov,G.K., Kundaje,A., Kheradpour,P., Pauli,F., Batzoglou,S., Bernstein,B.E.,
44 927 Bickel,P., Brown,J.B., Cayting,P., *et al.* (2012) ChIP-seq guidelines and practices of the
45 928 ENCODE and modENCODE consortia. *Genome Research*, **22**, 1813–1831.
- 46
47 929 63. Ramirez,F., Dundar,F., Diehl,S., Gruning,B.A. and Manke,T. (2014) deepTools: a flexible
48 930 platform for exploring deep-sequencing data. *Nucleic Acids Res*, **42**, W187-191.
- 49
50 931 64. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C.,
51 932 Myers,R.M., Brown,M., Li,W., *et al.* (2008) Model-based analysis of ChIP-Seq (MACS).
52 933 *Genome Biol.*, **9**, R137.

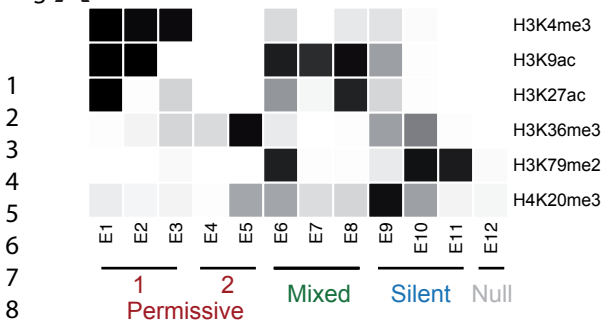
- 1
2
3 934 65. Xu,S., Grullon,S., Ge,K. and Peng,W. (2014) Spatial clustering for identification of ChIP-
4 935 enriched regions (SICER) to map regions of histone methylation patterns in embryonic
5 936 stem cells. *Methods Mol Biol*, **1150**, 97–111.
- 7
8 937 66. Zang,C., Schones,D.E., Zeng,C., Cui,K., Zhao,K. and Peng,W. (2009) A clustering approach
9 938 for identification of enriched domains from histone modification ChIP-Seq data.
10 939 *Bioinformatics*, **25**, 1952–1958.
- 12
13 940 67. Diaz,A., Park,K., Lim,D.A. and Song,J.S. (2012) Normalization, bias correction, and peak
14 941 calling for ChIP-seq. *Stat Appl Genet Mol Biol*, **11**, Article 9.
- 16
17 942 68. Lerdrup,M., Johansen,J.V., Agrawal-Singh,S. and Hansen,K. (2016) An interactive
18 943 environment for agile analysis and visualization of ChIP-sequencing data. *Nat. Struct.*
19 944 *Mol. Biol.*, **23**, 349–357.
- 21
22 945 69. Krzywinski,M., Schein,J., Birol,I., Connors,J., Gascoyne,R., Horsman,D., Jones,S.J. and
23 946 Marra,M.A. (2009) Circos: an information aesthetic for comparative genomics.
24 947 *Genome Res.*, **19**, 1639–1645.
- 25
26 948 70. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing
27 949 genomic features. *Bioinformatics*, **26**, 841–842.
- 28
29 950 71. Alexa,A. and Rahnenfuhrer,J. (2020) Enrichment Analysis for Gene Ontology.
- 31
32 951 72. Lipinska,A.P., D’hondt,S., Van Damme,E.J. and De Clerck,O. (2013) Uncovering the genetic
33 952 basis for early isogamete differentiation: a case study of *Ectocarpus siliculosus*. *BMC*
34 953 *Genomics*, **14**, 909–909.
- 35
36 954 73. Meisel,R.P., Malone,J.H. and Clark,A.G. (2012) Disentangling the relationship between sex-
37 955 biased gene expression and X-linkage. *Genome Res*, **22**, 1255–1265.
- 38
39 956 74. Bisoni,L., Batlle-Morera,L., Bird,A.P., Suzuki,M. and McQueen,H.A. (2005) Female-specific
40 957 hyperacetylation of histone H4 in the chicken Z chromosome. *Chromosome Res*, **13**,
41 958 205–214.
- 43
44 959 75. Brockdorff,N. and Turner,B.M. (2015) Dosage compensation in mammals. *Cold Spring Harb*
45 960 *Perspect Biol*, **7**, a019406.
- 46
47 961 76. Arendsee,Z.W., Li,L. and Wurtele,E.S. (2014) Coming of age: orphan genes in plants. *Trends*
48 962 *Plant Sci*, **19**, 698–708.
- 49
50 963 77. Palmieri,N., Kosiol,C. and Schlötterer,C. (2014) The life cycle of *Drosophila* orphan genes.
51 964 *Elife*, **3**, e01311.
- 53
54 965 78. McDaniel,S.F., Neubig,K.M., Payton,A.C., Quatrano,R.S. and Cove,D.J. (2013) Recent gene-
55 966 capture on the UV sex chromosomes of the moss *Ceratodon purpureus*. *Evolution*, **67**,
56 967 2811–2822.
- 58
59 968 79. Filion,G.J., van Bemmelen,J.G., Braunschweig,U., Talhout,W., Kind,J., Ward,L.D.,
60 969 Brugman,W., de Castro,I.J., Kerkhoven,R.M., Bussemaker,H.J., *et al.* (2010) Systematic

- 1
2
3 970 protein location mapping reveals five principal chromatin types in *Drosophila* cells.
4 971 *Cell*, **143**, 212–224.
5
6
7 972 80. Baroux,C., Raissig,M.T. and Grossniklaus,U. (2011) Epigenetic regulation and
8 973 reprogramming during gamete formation in plants. *Current Opinion in Genetics &*
9 974 *Development*, **21**, 124–133.
10
11 975 81. Margueron,R. and Reinberg,D. (2010) Chromatin structure and the inheritance of
12 976 epigenetic information. *Nat Rev Genet*, **11**, 285–296.
13
14
15 977 82. She,W. and Baroux,C. (2015) Chromatin dynamics in pollen mother cells underpin a
16 978 common scenario at the somatic-to-reproductive fate transition of both the male and
17 979 female lineages in *Arabidopsis*. *Front Plant Sci*, **6**, 294.
18
19 980 83. de la Paz Sanchez,M. and Gutierrez,C. (2009) *Arabidopsis* ORC1 is a PHD-containing
20 981 H3K4me3 effector that regulates transcription. *Proc Natl Acad Sci USA*, **106**, 2065.
22
23 982 84. Fischer,A., Hofmann,I., Naumann,K. and Reuter,G. (2006) Heterochromatin proteins and
24 983 the control of heterochromatic gene silencing in *Arabidopsis*. *J Plant Physiol*, **163**, 358–
25 984 368.
26
27 985 85. Vigneau,J. and Borg,M. (2021) The epigenetic origin of life history transitions in plants and
28 986 algae. *Plant Reprod*, **34**, 267–285.
29
30
31 987 86. Lesch,B.J. and Page,D.C. (2013) Sex-specific chromatin states in mammalian fetal germ
32 988 cells. *Epigenetics Chromatin*, **6**, P45–P45.
33
34 989 87. Sugathan,A. and Waxman,D.J. (2013) Genome-wide analysis of chromatin states reveals
35 990 distinct mechanisms of sex-dependent gene regulation in male and female mouse
36 991 liver. *Mol Cell Biol*, **33**, 3594–3610.
38
39 992 88. Kukurba,K.R., Parsana,P., Balliu,B., Smith,K.S., Zappala,Z., Knowles,D.A., Favé,M.-J.,
40 993 Davis,J.R., Li,X., Zhu,X., *et al.* (2016) Impact of the X Chromosome and sex on regulatory
41 994 variation. *Genome Res*, **26**, 768–777.
43
44 995 89. Mignerot,L. and Coelho,S.M. (2016) The origin and evolution of the sexes: Novel insights
45 996 from a distant eukaryotic lineage. *C R Biol*, **339**, 252–257.
46
47 997 90. Ferris,P., Olson,B.J., De Hoff,P.L., Douglass,S., Diaz-Cano,D.C., Prochnik,S., Geng,S., Rai,R.,
48 998 Grimwood,J., Schmutz,J., *et al.* (2010) Evolution of an Expanded Sex Determining Locus
49 999 in *Volvox*. *Science (New York, N.Y.)*, **328**, 351–354.
51
52 1000 91. Yamato,K.T., Ishizaki,K., Fujisawa,M., Okada,S., Nakayama,S., Fujishita,M., Bando,H.,
53 1001 Yodoya,K., Hayashi,K., Bando,T., *et al.* (2007) Gene organization of the liverwort Y
54 1002 chromosome reveals distinct sex chromosome evolution in a haploid system. *Proc Natl*
55 1003 *Acad Sci U S A*, **104**, 6472–6477.
56
57
58 1004 92. Okada,S., Sone,T., Fujisawa,M., Nakayama,S., Takenaka,M., Ishizaki,K., Kono,K., Shimizu-
59 1005 Ueda,Y., Hanajiri,T., Yamato,K.T., *et al.* (2001) The Y chromosome in the liverwort

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

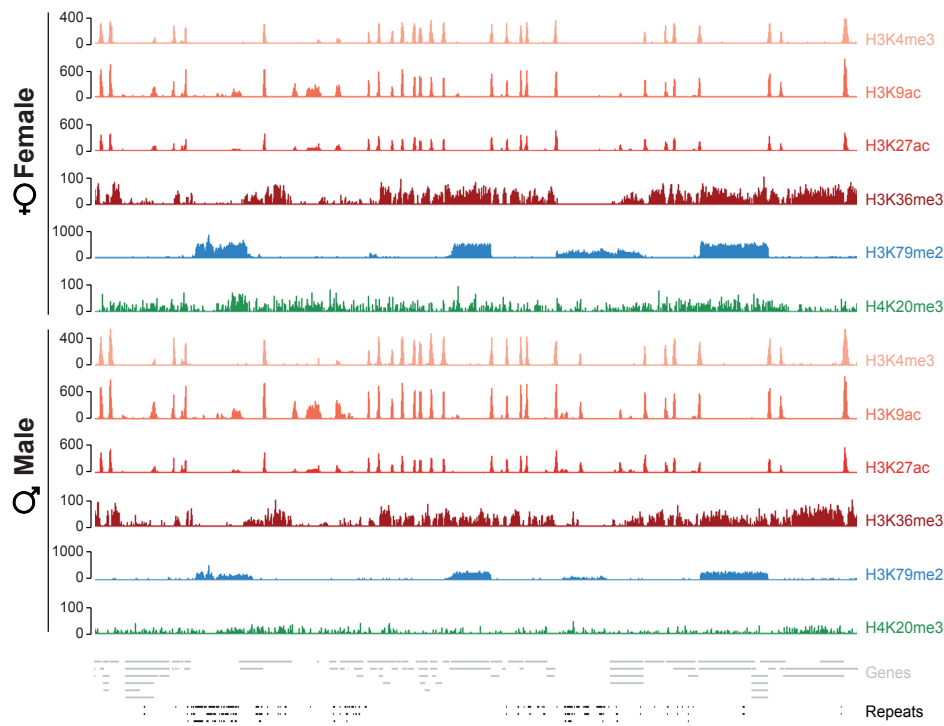
- 1006 Marchantia polymorpha has accumulated unique repeat sequences harboring a male-
1007 specific gene. *Proceedings of the National Academy of Sciences*, **98**, 9454–9459.
- 1008 93. Bell,O., Conrad,T., Kind,J., Wirbelauer,C., Akhtar,A. and Schubeler,D. (2008) Transcription-
1009 coupled methylation of histone H3 at lysine 36 regulates dosage compensation by
1010 enhancing recruitment of the MSL complex in *Drosophila melanogaster*. *Mol Cell Biol*,
1011 **28**, 3401–3409.
- 1012 94. Roudier,F., Ahmed,I., Bérard,C., Sarazin,A., Mary-Huard,T., Cortijo,S., Bouyer,D.,
1013 Caillieux,E., Duvernois-Berthet,E., Al-Shikhley,L., *et al.* (2011) Integrative epigenomic
1014 mapping defines four main chromatin states in *Arabidopsis*. *The EMBO Journal*, **30**,
1015 1928–1938.
- 1016 95. Shilatifard,A. (2006) Chromatin modifications by methylation and ubiquitination:
1017 implications in the regulation of gene expression. *Annu Rev Biochem*, **75**, 243–269.
- 1018 96. Werner,M.S., Sieriebriennikov,B., Prabh,N., Loschko,T., Lanz,C. and Sommer,R.J. (2018)
1019 Young genes have distinct gene structure, epigenetic profiles, and transcriptional
1020 regulation. *Genome Research*, **28**, 1675–1687.

Emission States

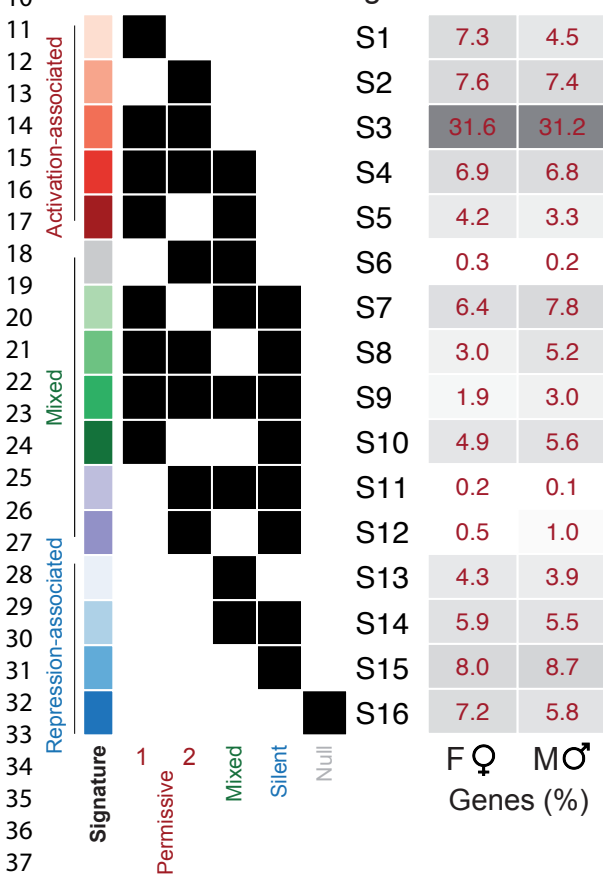


Nucleic Acids Research

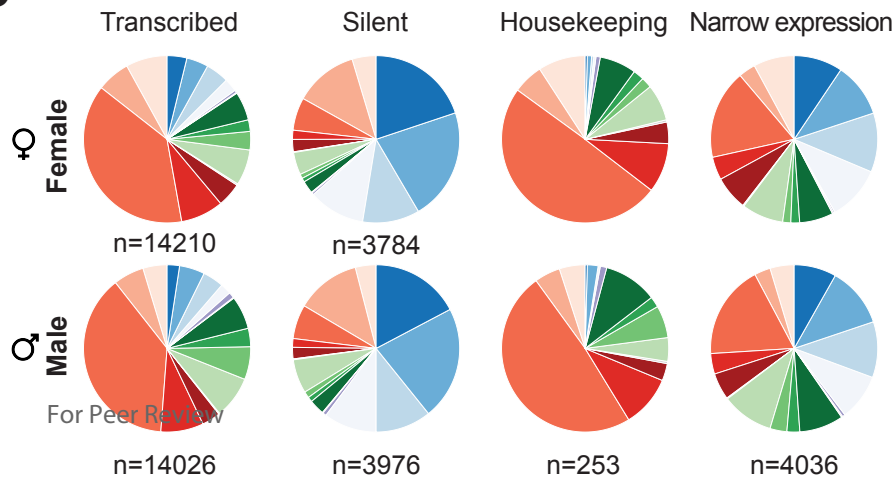
Chr 19: 2335100 - 2937100

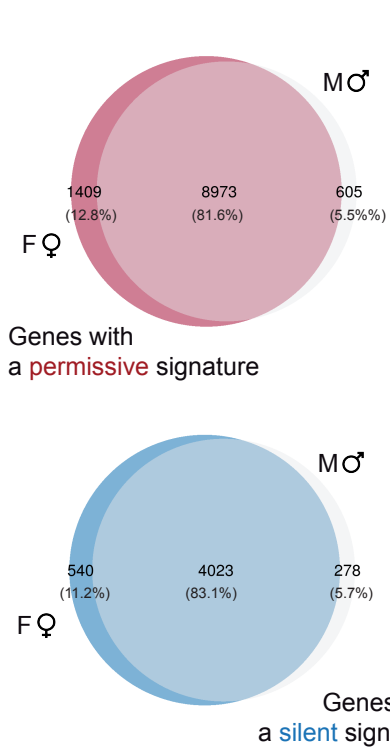
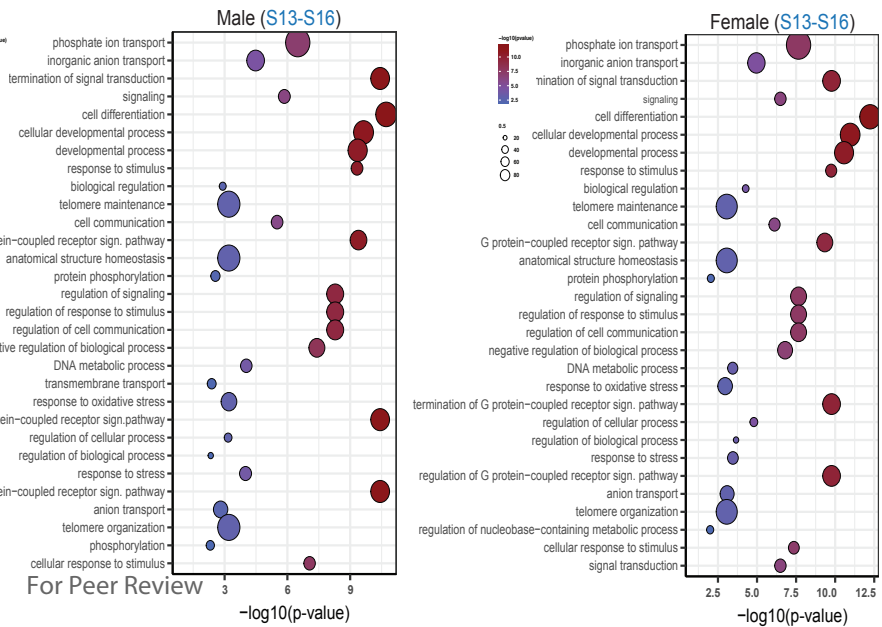
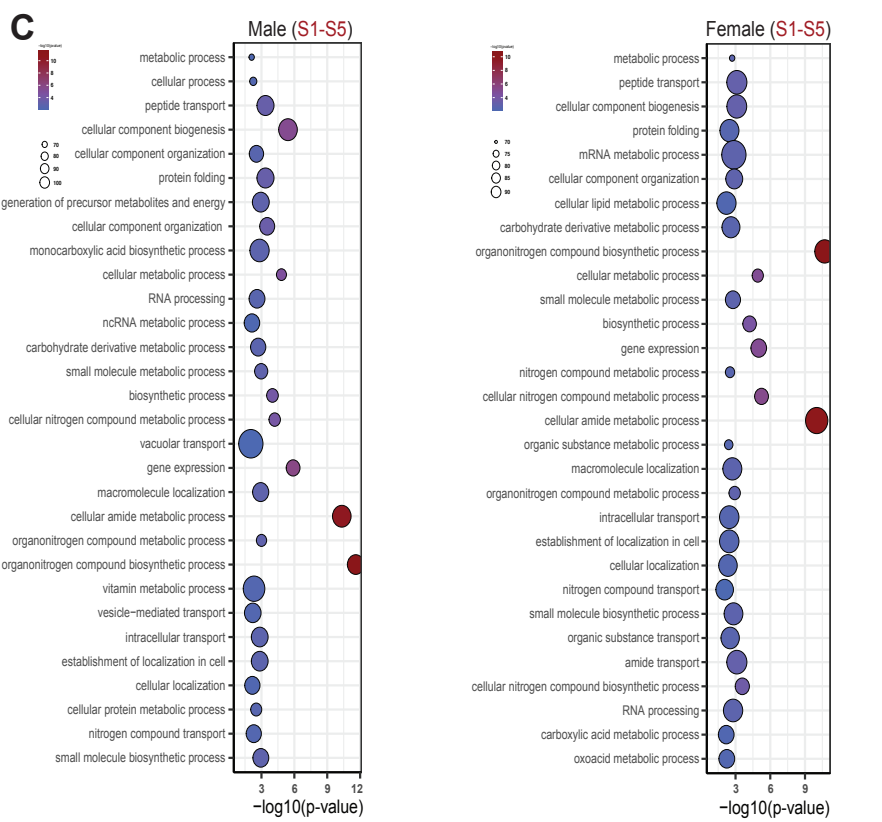
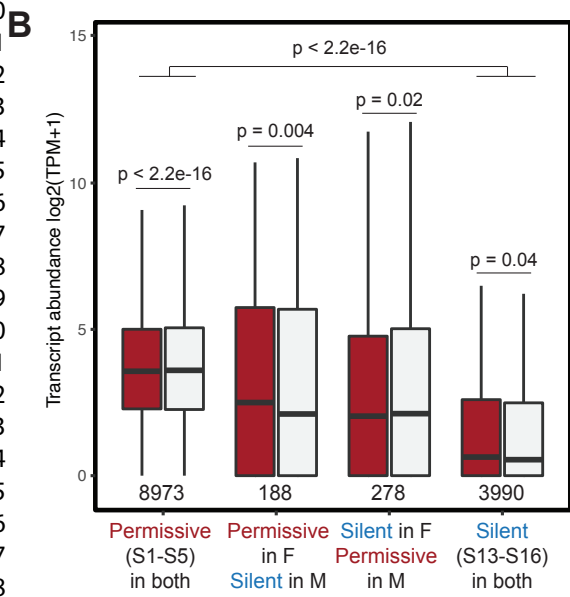
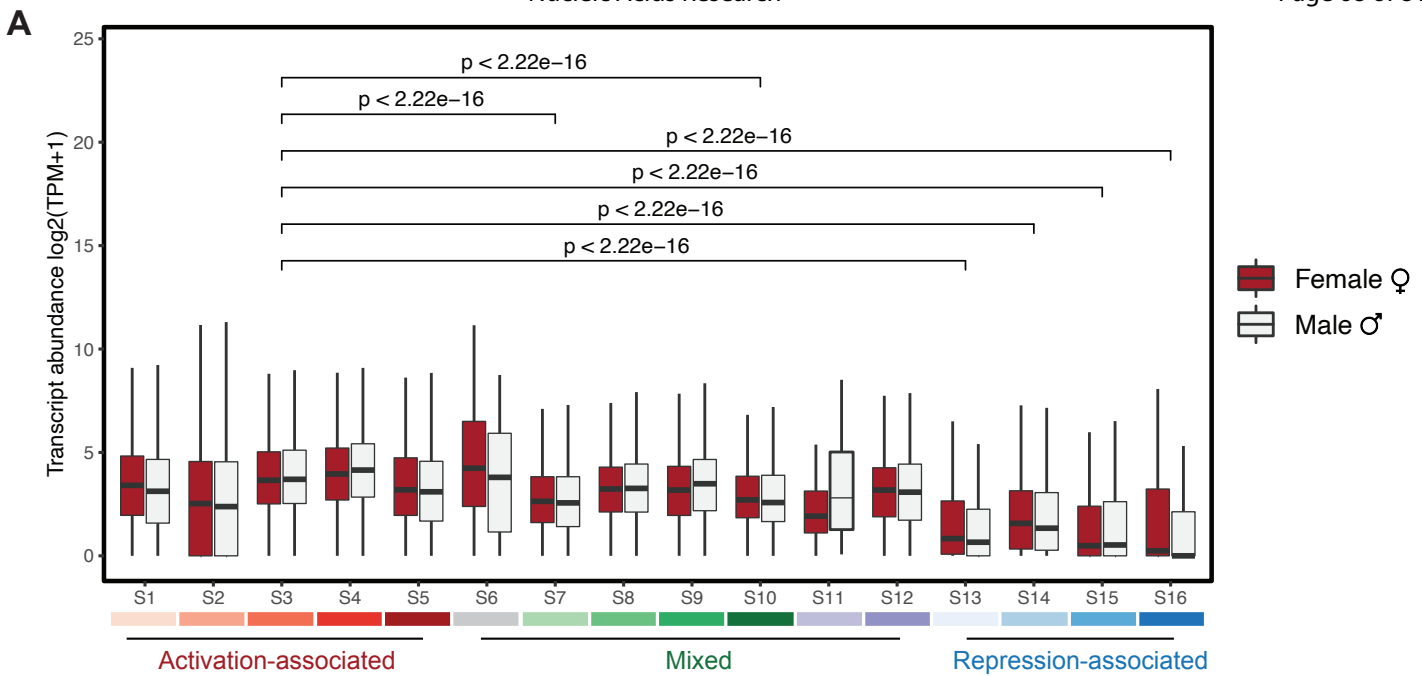


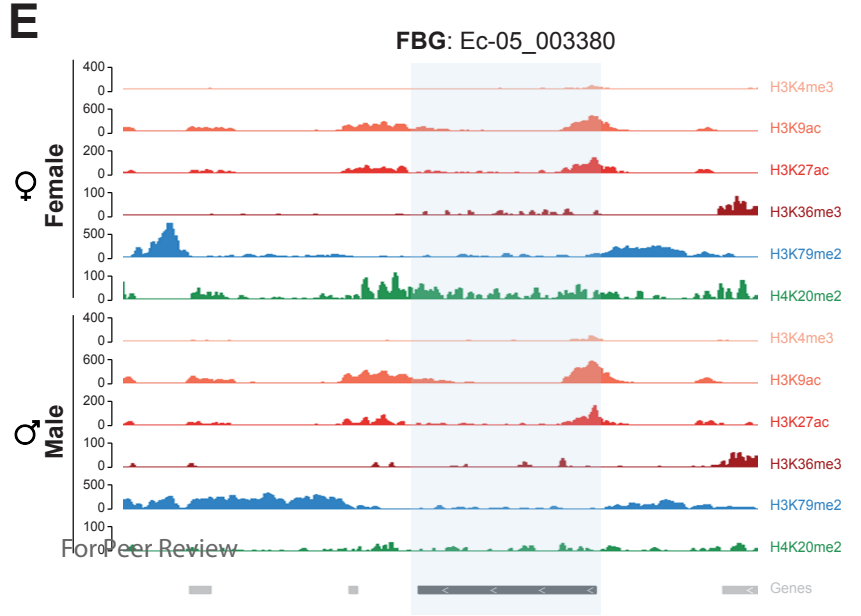
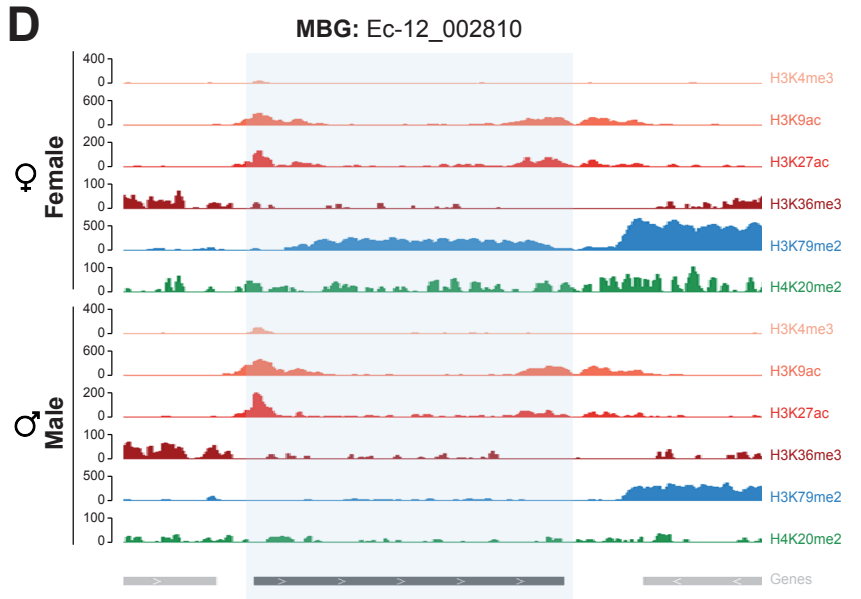
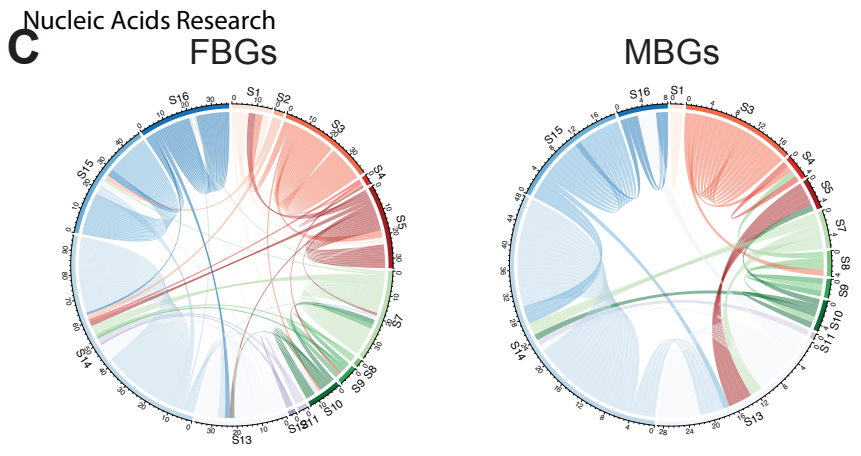
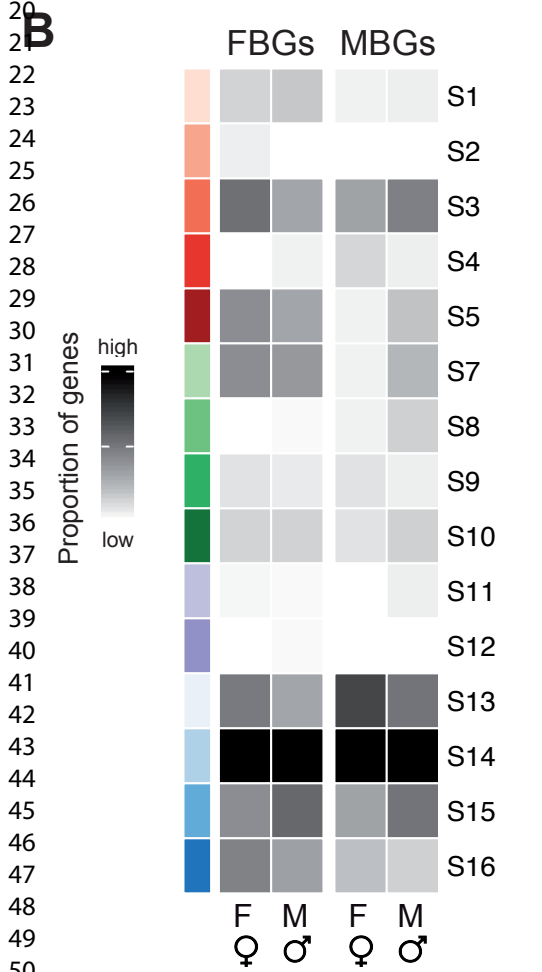
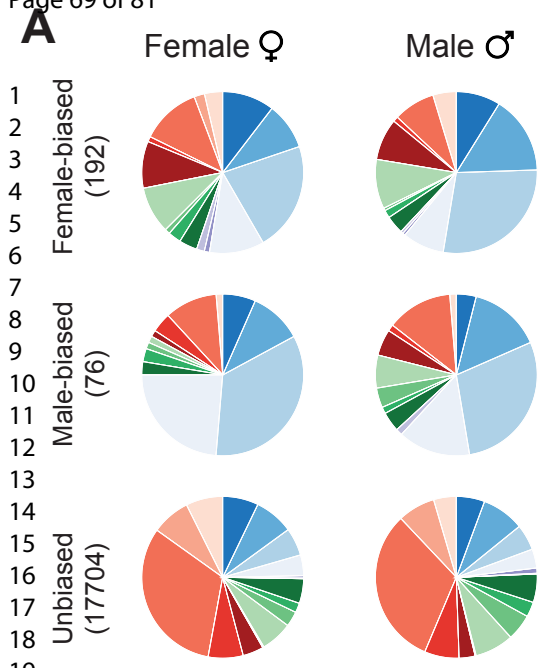
Chromatin Signatures

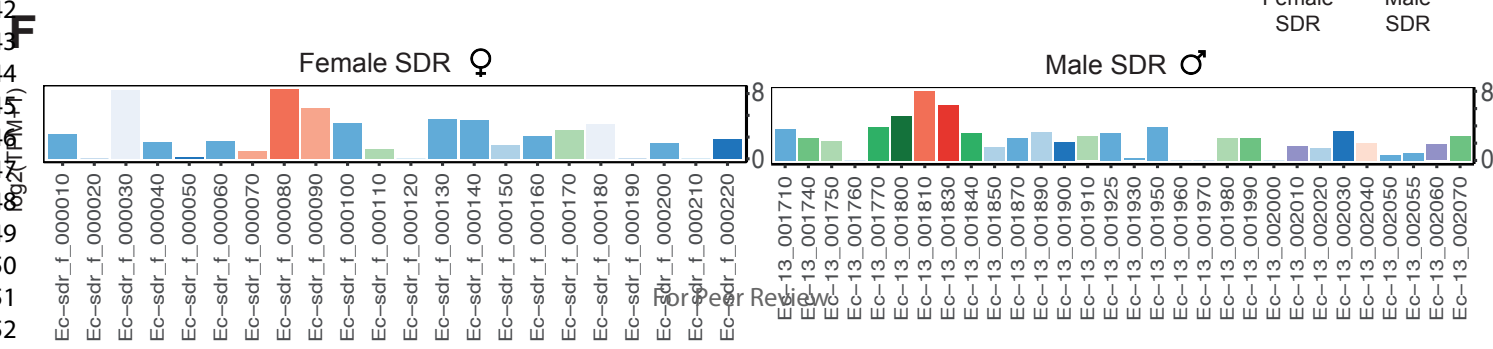
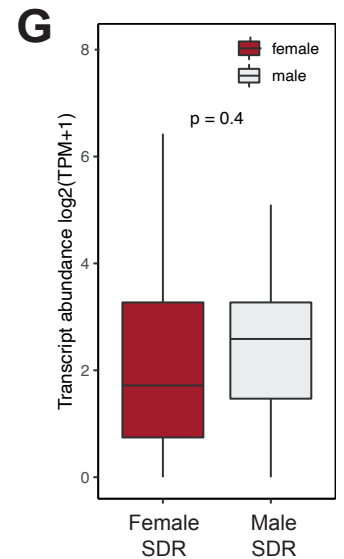
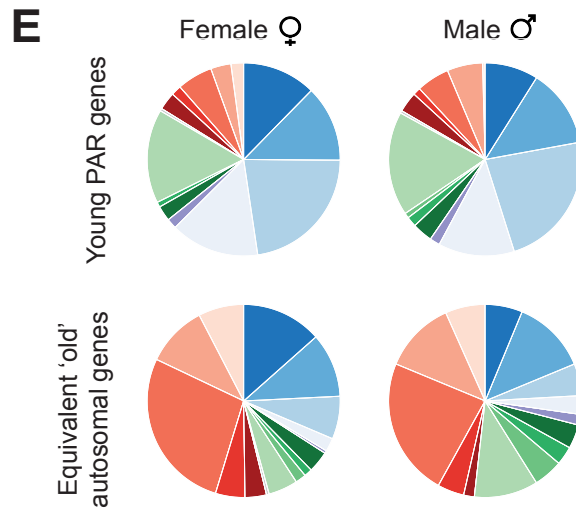
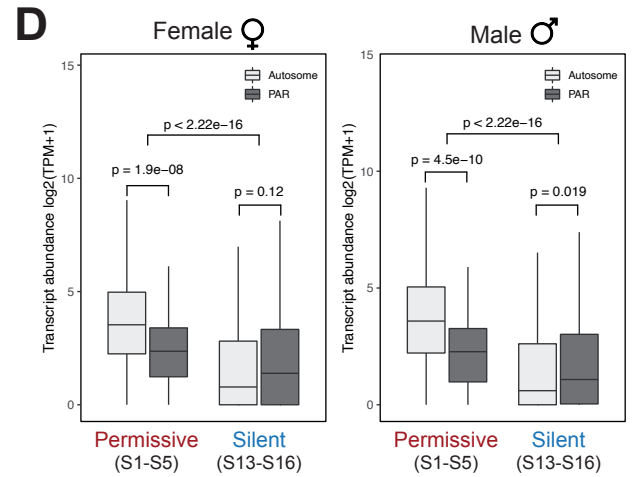
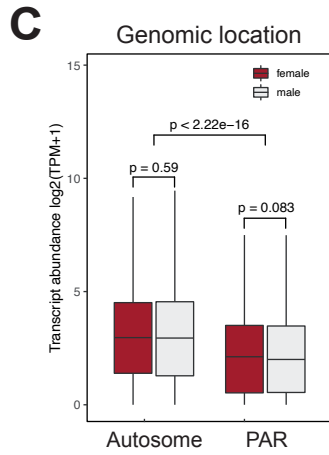
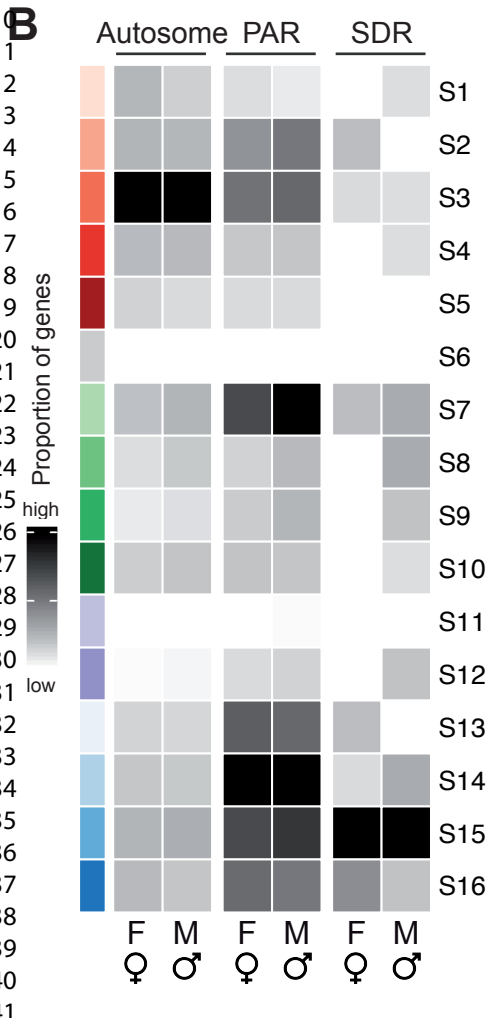
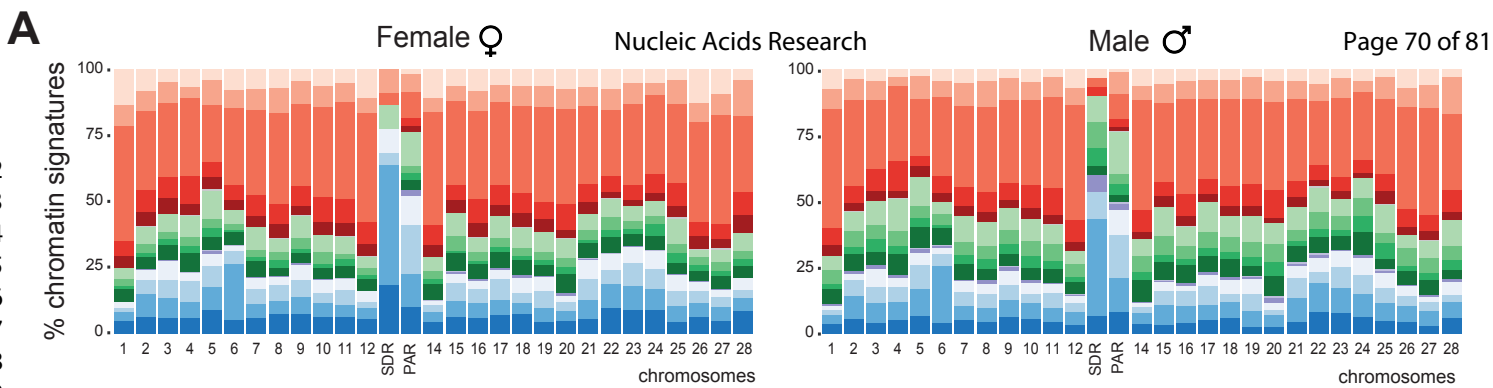


D









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.](#)

1
2
3 **Figure S10.** Transcript abundances, measured as $\log_2(\text{TPM}+1)$, for PAR genes associated with
4 different chromatin signatures in males and females. The colour code is the same as that used in
5 Figure 1C.
6
7

8 9 Supplemental Tables Legends

10
11
12 **Table S1.** *Ectocarpus* strains used, RNA-seq sequencing statistics and SRA accession numbers.
13

14 **Table S2.** SNPs between male and female lines.
15

16
17 **Table S3.** Sequencing statistics for the ChIP-seq analysis and GEO reference for the dataset. N,
18 peaks, number of peaks; FRiP, fraction of reads in peaks.

19
20 **Table S4.** Percentages of genes associated with each of the 16 chromatin signatures for different
21 gene sets in males and females. Global, all genes in the genome; Transcribed genes, genes with
22 TPM > 1; Silent genes, genes with TPM <1; Housekeeping and Narrowly-expressed, genes with
23 tau <0.75 and tau >0.75, respectively; Unbiased, no sex-biased expression. For the chromatin
24 signatures, refer to Figure 1A.
25
26

27
28 **Table S5.** Chromatin signatures (S1-S16) and transcript abundances (measured as TPM) for all
29 *Ectocarpus* genes in males and females. FBG, female-biased gene; MBG, male-biased gene. For
30 the chromatin states, refer to Figure 1C.
31
32

33 **Table S6.** Number of sex-biased genes in each of the chromatin signatures S1-S16 in males and
34 females. FBG, female-biased gene, MBG, male-biased gene.
35

36
37 **Table S7.** Transitions between chromatin signatures observed for male-biased and female-biased
38 genes in males compared with females. For chromatin signatures, refer to Figure 1C.
39

40
41 **Table S8.** Proportion of genes marked with each of the histone PTMs in males in females in
42 different gene categories. Histone PTM peaks in each gene were analysed based on peak callers
43 MACS2 and Sicer (see methods for details).
44

45
46 **Table S9.** Coverage of the six histone PTMs across male and female genomes. The sex
47 chromosome (chromosome 13) is divided into PAR1 (pseudo-autosomal region 1), SDR (sex-
48 determining region) and PAR2 (pseudo-autosomal region 2).
49

50
51 **Table S10.** Chromatin signatures of PAR genes in males and females.
52

53 **Table S11.** Chromatin signatures and transcript abundances ($\log_2\text{TPM}+1$) for SDR genes (see also
54 Figure 4F).
55
56
57
58
59
60

1
2
3 **Table S12.** Permutation tests performed to determine whether the relative proportions of the
4 different chromatin signatures were statistically different in different regions of the genome. We
5 randomized the genomic location of autosomal genes 100,000 times and tested the difference
6 between the observed proportions for the SDR, the PAR or the entire sex chromosome and the
7 permuted gene sets using Chi-square statistics. Tests were performed independently for each
8 chromatin state. Significant p-values (<0.01) are highlighted in bold.
9
10
11

12
13 **Table S13.** The presence of transposon sequences in the majority (80%) of PAR genes does not
14 explain the distinct chromatin landscape of the PAR. Correlation between the presence of
15 transposable elements within introns and the presence of H4K20me3 in PAR genes and
16 autosomal genes (left table). Permutation tests comparing the proportion of each chromatin
17 signature in the PAR with the proportion of that signature in 100,000 samples of 430 autosomal
18 genes with transposon sequences in 80% of the genes. For most chromatin signatures, the
19 proportion on the PAR was significantly different from those of the autosomal gene samples
20 indicating that transposon content does not explain the unusual pattern of chromatin states
21 observed for the PAR. Significant p-values (<0.01) are highlighted in bold (right table).
22
23
24
25
26
27

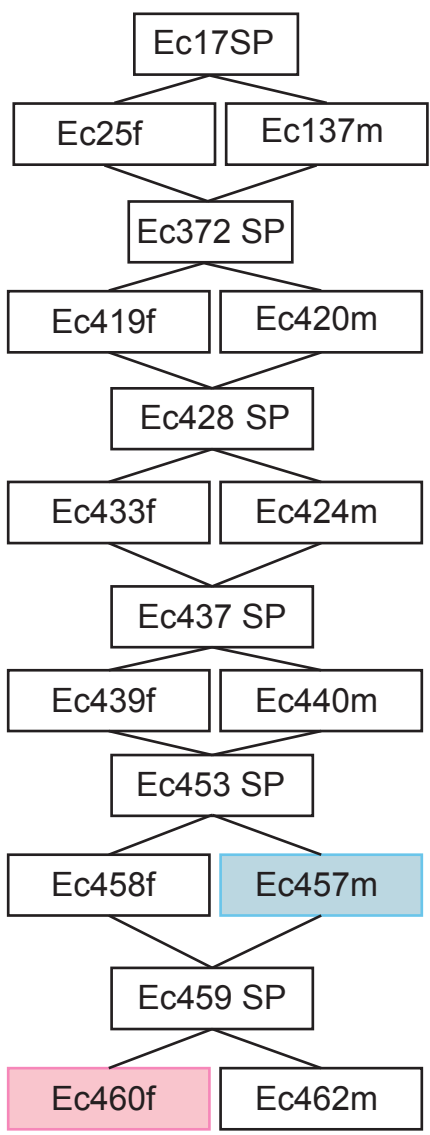
28 **Table S14.** Comparison of chromatin signatures of the PAR genes with those of a set of autosomal
29 genes with a similar pattern of gene expression levels. To establish the autosomal gene set, for
30 each PAR gene, the full set of autosomal genes was searched for the gene that had the most
31 similar level of expression. When the TPM of the PAR gene was zero, an autosomal gene with a
32 similar level of expression. When the TPM of the PAR gene was zero, an autosomal gene with a
33 TPM of zero was selected at random. Figure S6 presents the proportions of chromatin signatures
34 associated with the two gene sets.
35
36
37

38 **Table S15.** Permutation analysis to test whether there was a significant difference between the
39 distribution of chromatin signatures in young genes compared with evolutionary conserved
40 genes with similar expression levels (left table) and to test whether the distribution of signatures
41 of PAR genes that are evolutionary conserved are different from autosomal genes (right table).
42 Significant differences are highlighted in bold (p-value <0.01).
43
44
45

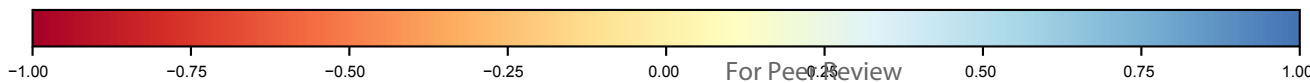
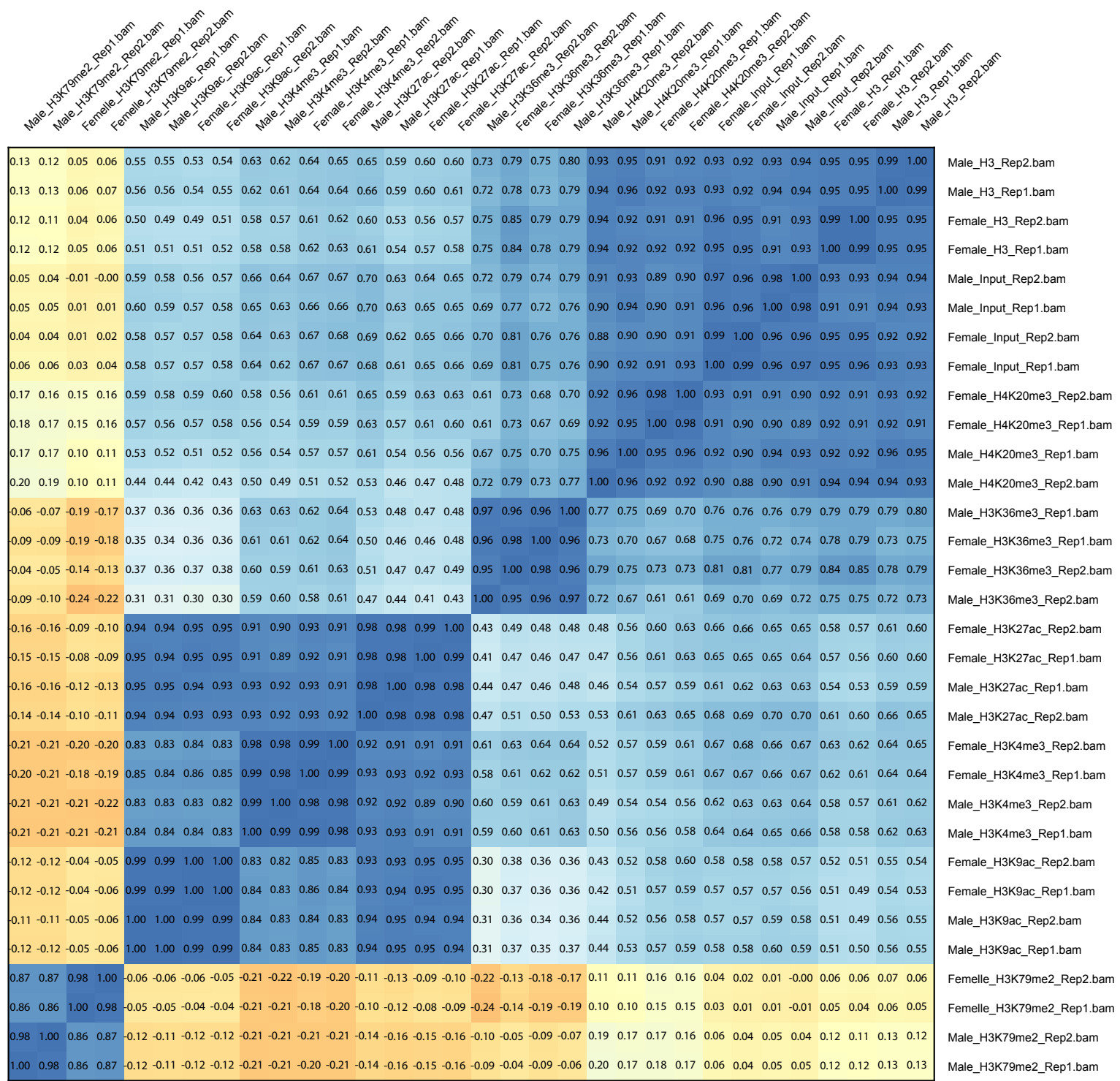
46 **Table S16.** Linear models to test whether there was a significant correlation between expression
47 level ($\log_2(\text{TPM} + 1)$) and chromatin state (upper table) or to test whether location of a gene on
48 the PAR or on an autosome significantly influenced the expression level associated with each
49 chromatin signature (bottom table). Significant interaction terms, in bold, represent a
50 significantly different effect of the chromatin signature on gene expression level in the PAR region
51 compared to autosomal genes (p-value <0.05).
52
53
54
55
56
57
58
59
60

1
2
3 **Table S17.** Linear models to test whether there was a significant correlation between expression
4 level ($\log_2(\text{TPM} + 1)$) and chromatin signature for SDR genes.
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

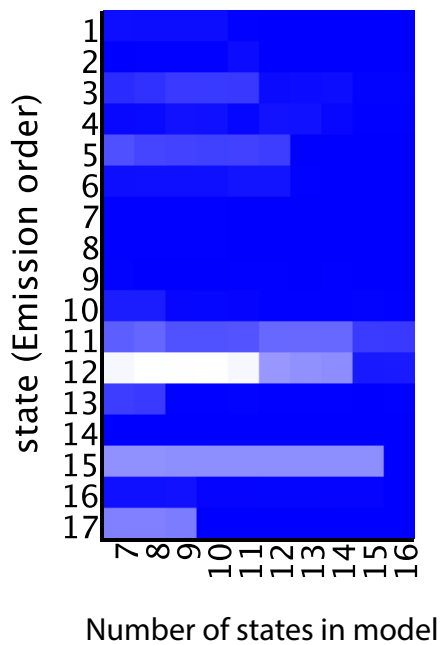
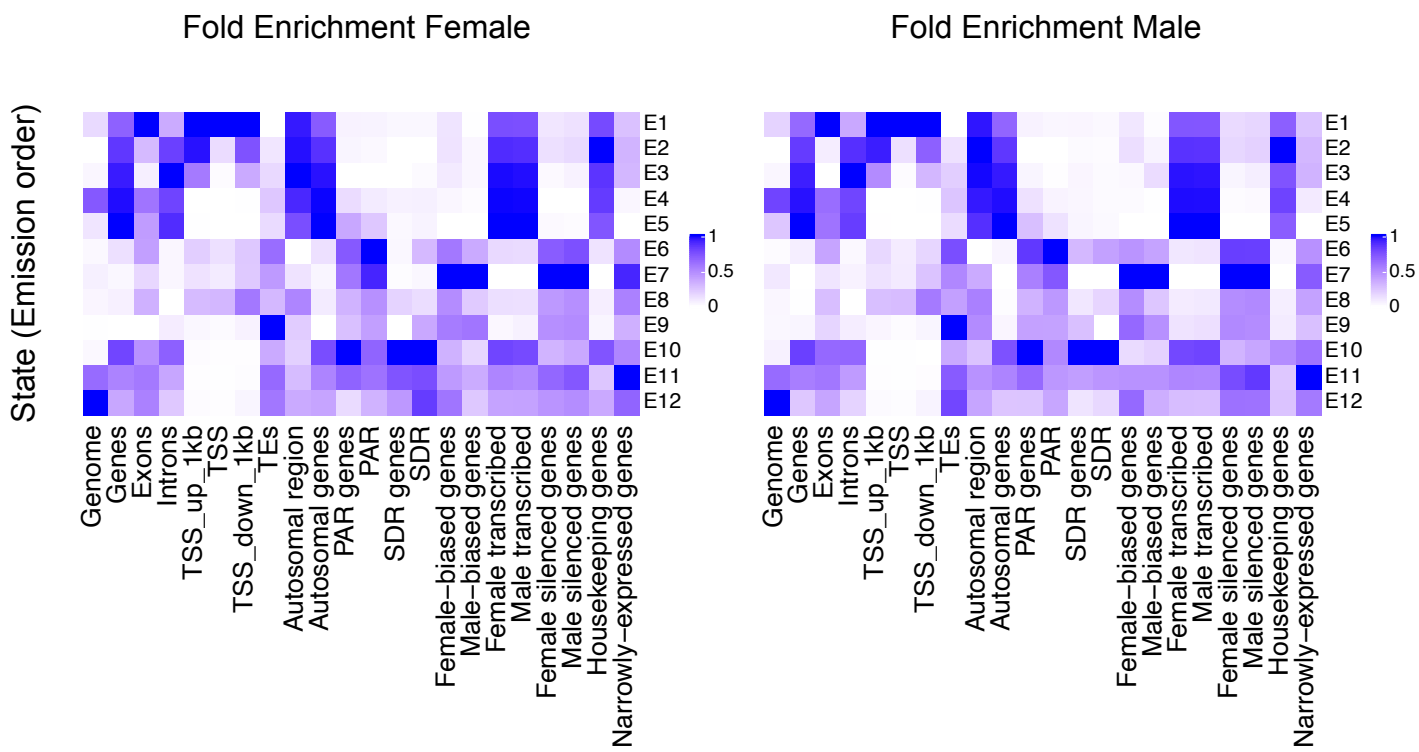
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

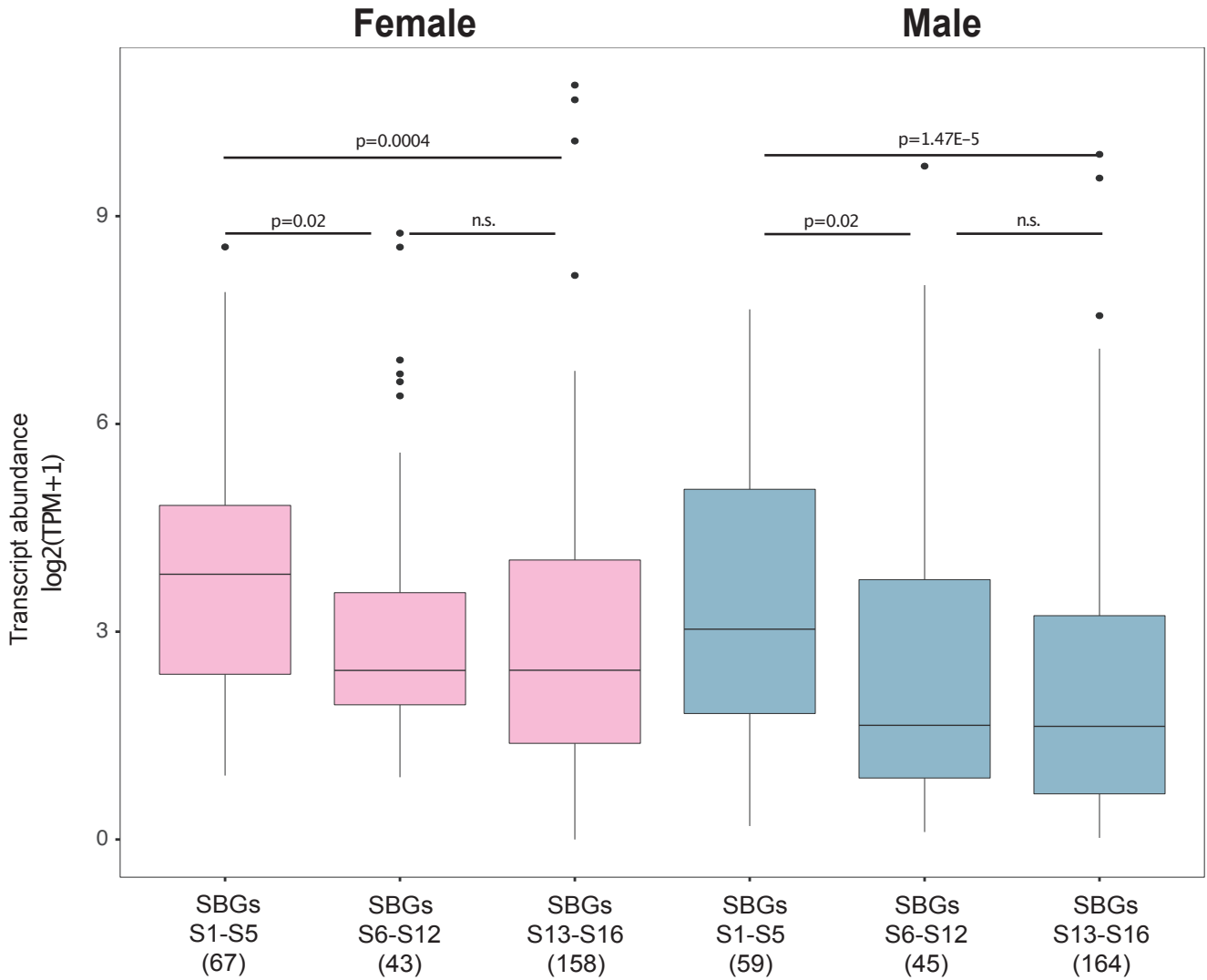


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

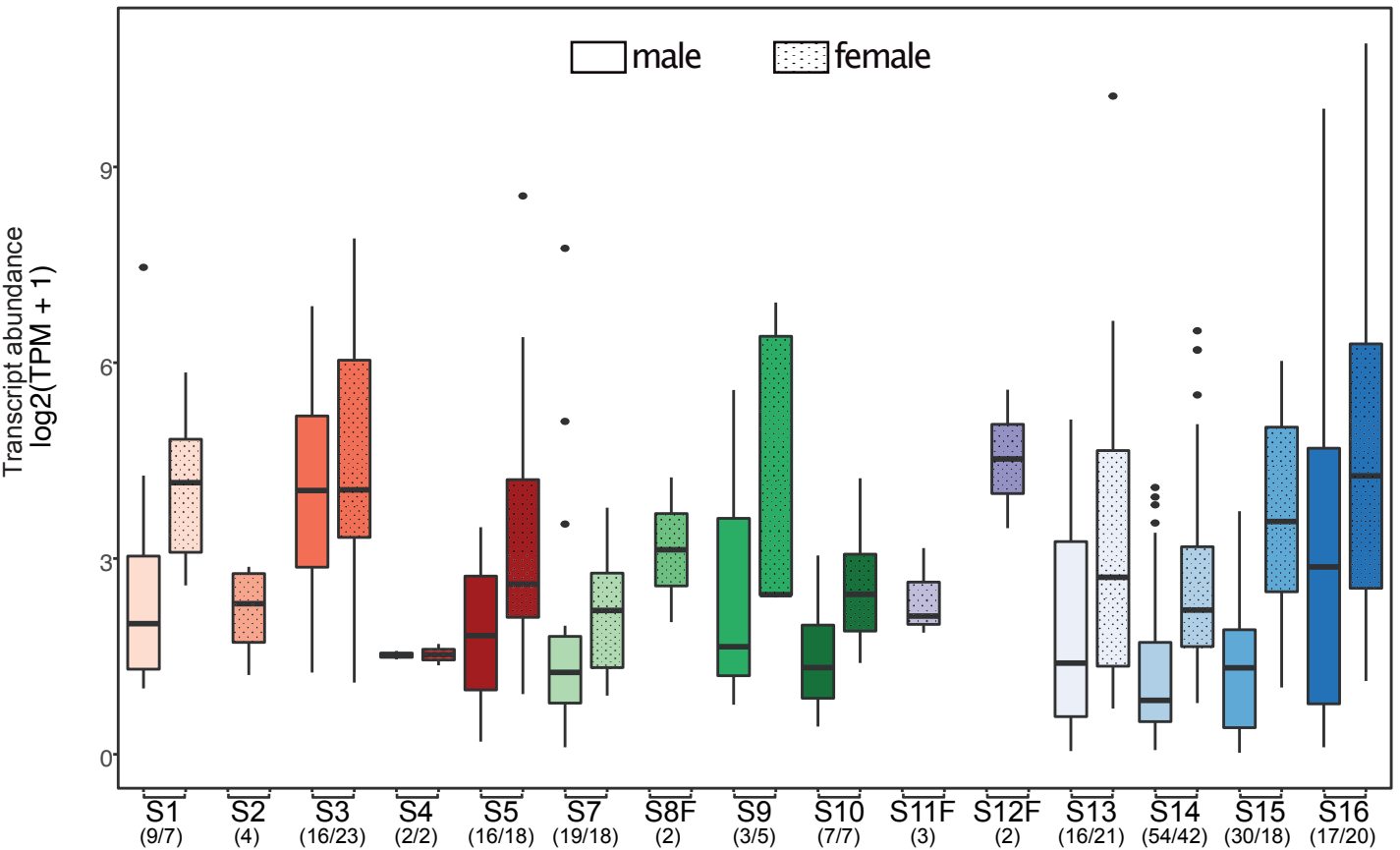


Male_H3_Rep2.bam
 Male_H3_Rep1.bam
 Female_H3_Rep2.bam
 Female_H3_Rep1.bam
 Male_Input_Rep2.bam
 Male_Input_Rep1.bam
 Female_Input_Rep2.bam
 Female_Input_Rep1.bam
 Female_H4K20me3_Rep2.bam
 Female_H4K20me3_Rep1.bam
 Male_H4K20me3_Rep1.bam
 Male_H4K20me3_Rep2.bam
 Male_H3K36me3_Rep1.bam
 Female_H3K36me3_Rep1.bam
 Female_H3K36me3_Rep2.bam
 Male_H3K36me3_Rep2.bam
 Female_H3K27ac_Rep2.bam
 Female_H3K27ac_Rep1.bam
 Male_H3K27ac_Rep1.bam
 Male_H3K27ac_Rep2.bam
 Female_H3K4me3_Rep2.bam
 Female_H3K4me3_Rep1.bam
 Male_H3K4me3_Rep2.bam
 Male_H3K4me3_Rep1.bam
 Female_H3K9ac_Rep2.bam
 Female_H3K9ac_Rep1.bam
 Male_H3K9ac_Rep2.bam
 Male_H3K9ac_Rep1.bam
 Femelle_H3K79me2_Rep2.bam
 Femelle_H3K79me2_Rep1.bam
 Male_H3K79me2_Rep2.bam
 Male_H3K79me2_Rep1.bam

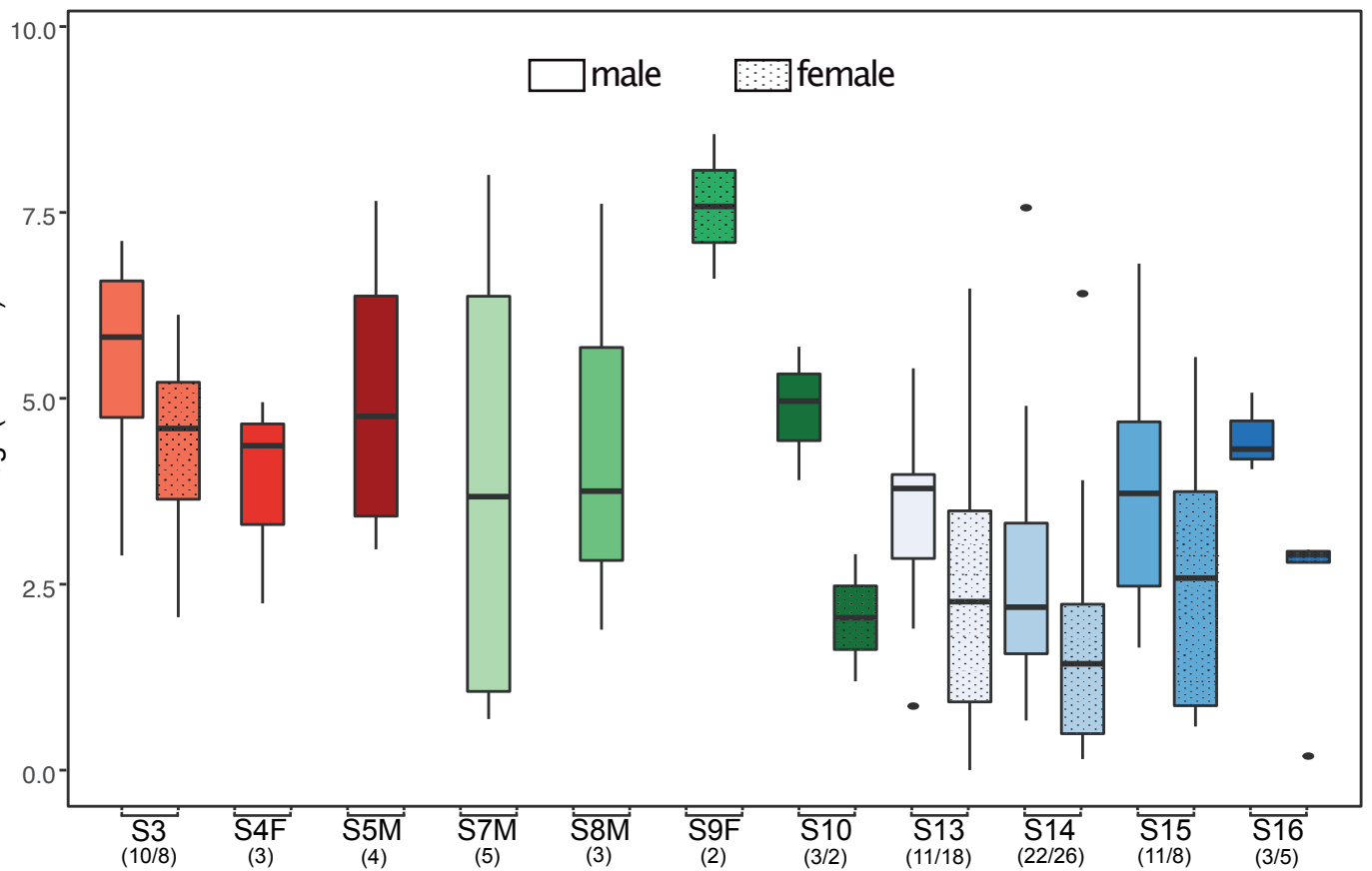




Female-biased genes



Male-biased genes

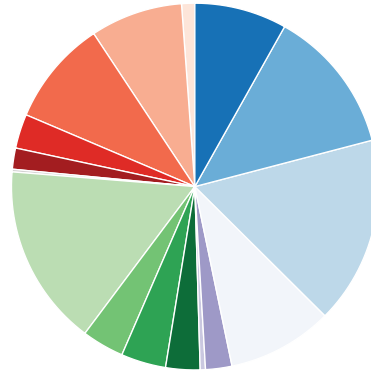
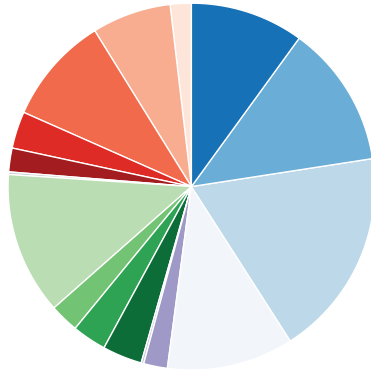


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

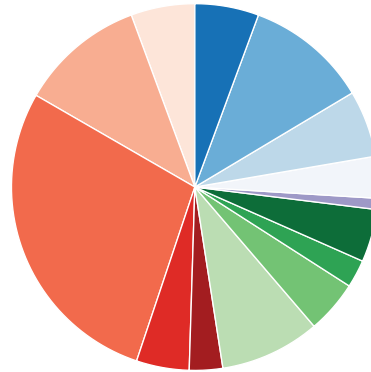
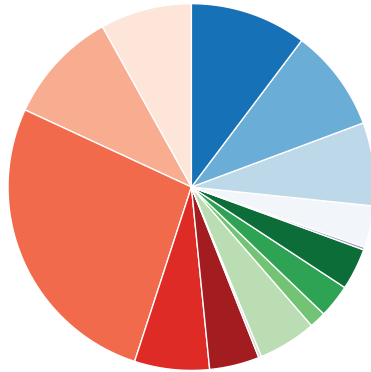
Female ♀

Male ♂

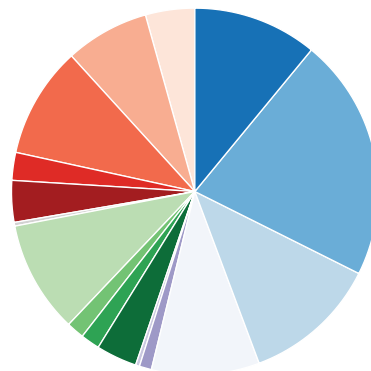
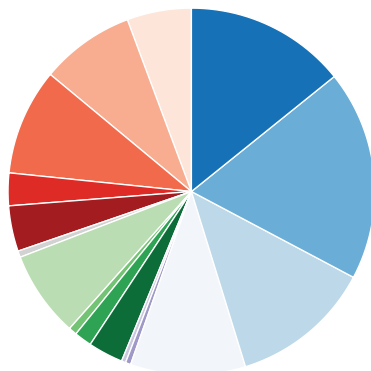
PAR genes

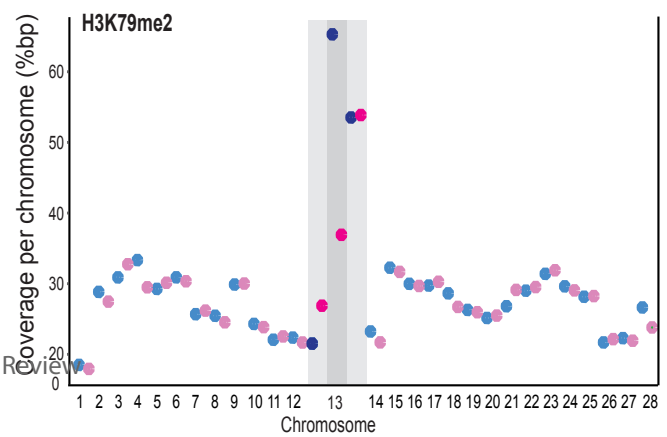
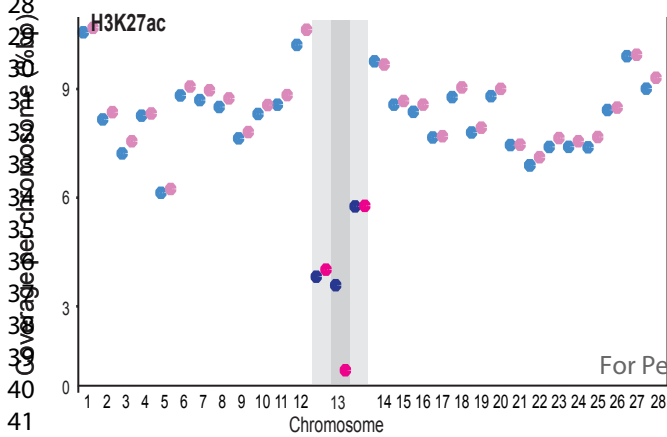
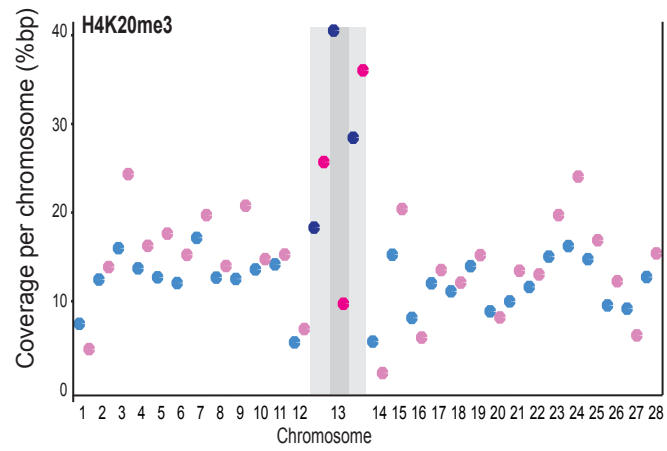
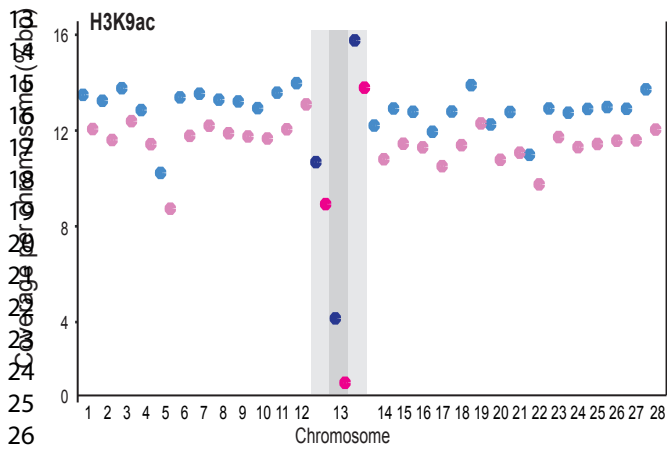
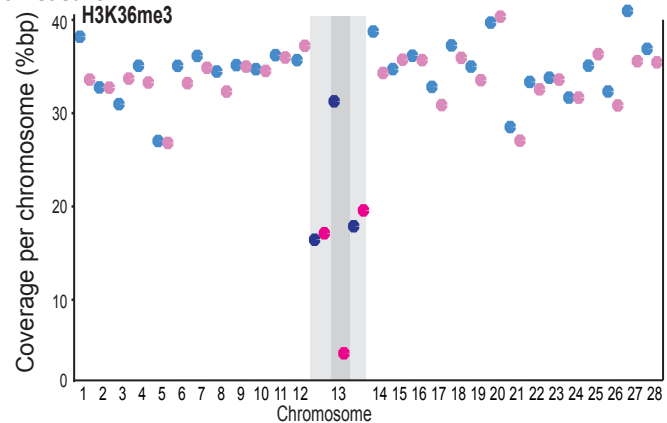
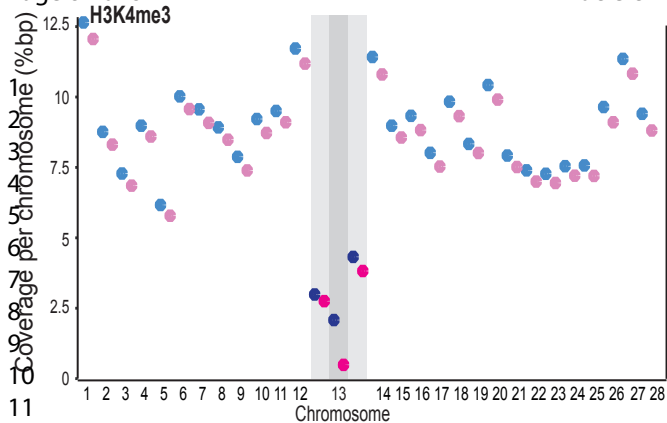


Set of autosomal genes with similar expression levels to the PAR genes



Young autosomal genes





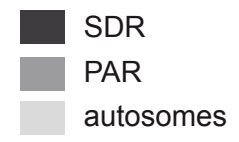
Males

Females

Nucleic Acids Research

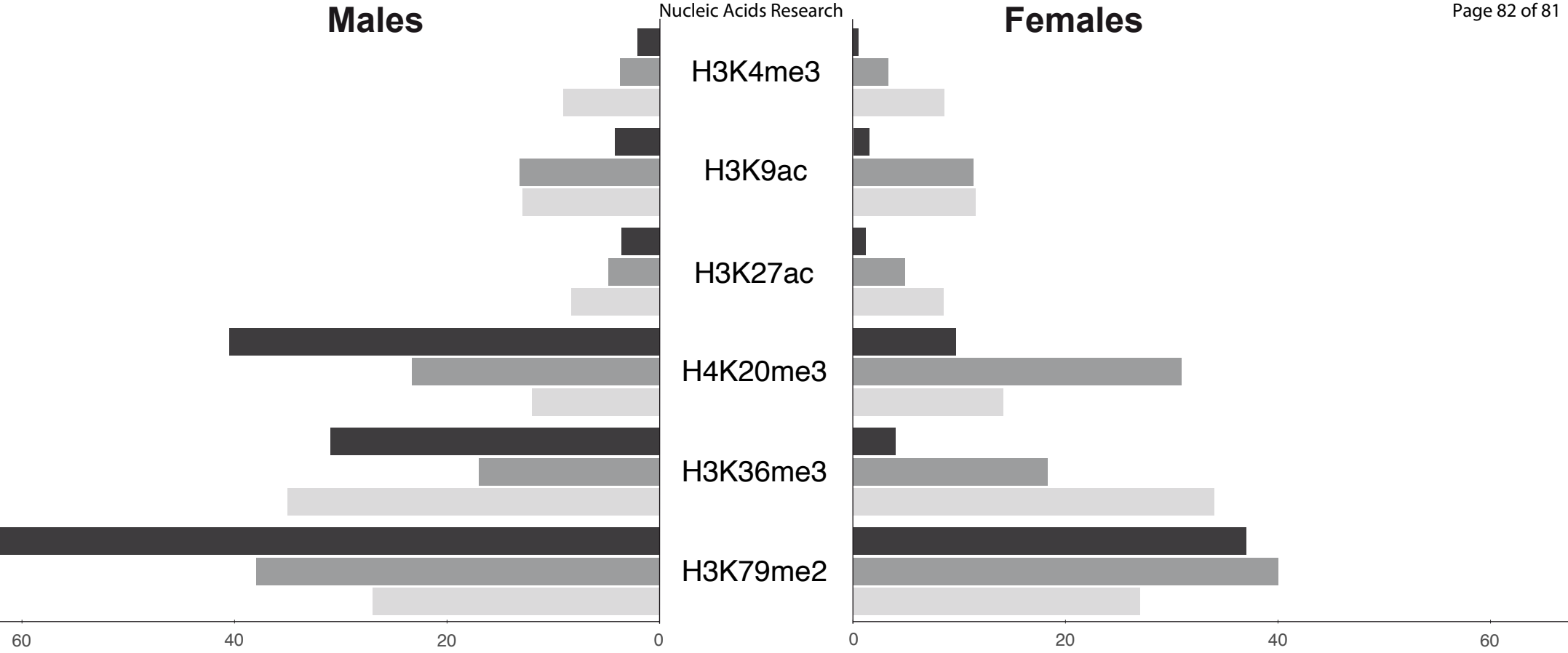
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

H3K4me3
H3K9ac
H3K27ac
H4K20me3
H3K36me3
H3K79me2



%Coverage

For Peer Review



PAR

SDR

PAR

